Annals of Internal Medicine | ARTICLE

The Sensitivity and Specificity of a Simple Test To Distinguish between Urge and Stress Urinary Incontinence

Jeanette S. Brown, MD; Catherine S. Bradley, MD, MSCE; Leslee L. Subak, MD; Holly E. Richter, MD, PhD; Stephen R. Kraus, MD; Linda Brubaker, MD, MS; Feng Lin, MS; Eric Vittinghoff, PhD; and Deborah Grady, MD, MPH, for the Diagnostic Aspects of Incontinence Study (DAISy) Research Group

Background: Urinary incontinence is common in women. Because treatments differ, urge incontinence should be distinguished from stress incontinence. To make this distinction, current guidelines recommend an extensive evaluation that is too time-consuming for primary care practice.

Objective: To test the accuracy of a simple questionnaire to categorize type of urinary incontinence in women.

Design: Multicenter, prospective study of the accuracy of the 3 Incontinence Questions (3IQ) compared with an extended evaluation to distinguish between urge incontinence and stress incontinence.

Setting: 5 academic medical centers in the United States.

Participants: 301 women enrolled from April to December 2004 who were older than 40 years of age (mean age, 56 years [SD, 11]) with untreated incontinence for an average of 7 years (SD, 7) and a broad range of incontinence severity.

Measurements: All participants included in the analyses answered the 3IQ questionnaire, and a urologist or urogynecologist who was blinded to the responses performed the extended evaluation. Sensitivity, specificity, and likelihood ratios were determined for the 3IQ.

Results: For classification of urge incontinence and with the extended evaluation as the gold standard, the 3IQ had a sensitivity of 0.75 (95% CI, 0.68 to 0.81), a specificity of 0.77 (CI, 0.69 to 0.84), and a positive likelihood ratio of 3.29 (CI, 2.39 to 4.51). For classification of stress incontinence, the sensitivity was 0.86 (CI, 0.79 to 0.90), the specificity was 0.60 (CI, 0.51 to 0.68), and the positive likelihood ratio was 2.13 (CI, 1.71 to 2.66).

Limitations: Participants were enrolled by urologists and urogynecologists at academic medical centers.

Conclusions: The 3IQ questionnaire is a simple, quick, and noninvasive test with acceptable accuracy for classifying urge and stress incontinence and may be appropriate for use in primary care settings. Similar studies are needed in other populations. We also need a clinical trial comparing the outcomes of treatments based on the 3IQ and the extended evaluation.

Ann Intern Med. 2006;144:715-723. www.annals.org
For author affiliations, see end of text.

Nearly 35% of women older than 40 years of age have urinary incontinence, which is associated with increased social isolation, falls, fractures, and admissions to specialized nursing units (1–3). Diagnosis and treatment of this common condition by primary care physicians would be optimal.

Incontinence is generally classified as urge (urine leakage with the urge to urinate), stress (urine leakage when straining, coughing, or exercising), mixed (both types), or other uncommon types of incontinence (such as neurogenic and overflow). While both urge and stress incontinence may improve with behavioral interventions, such as bladder training, urge incontinence is effectively treated with antimuscarinic or anticholinergic medications (4, 5) and stress incontinence is treated with pelvic muscle exercises and surgery (6, 7). Because treatment differs, it is important to distinguish urge incontinence from stress incontinence.

To distinguish urge from stress incontinence, current guidelines recommend a history, voiding diary, test for urinary tract infection, neurologic and pelvic examination, measurement of postvoid residual urine volume, and a cough stress test (8, 9). Completion of these tests is time-consuming, invasive, and expensive and is generally not feasible in primary care practice.

On the basis of previous research (10, 11) and expert clinical opinion, we developed a brief, self-administered questionnaire to distinguish urge from stress incontinence that includes 3 questions (the 3 Incontinence Questions [3IQ]) and requires about 30 seconds to complete. To estimate the accuracy of the 3IQ, we conducted a prospective study among ambulatory women with incontinence at 5 academic medical centers in the United States.

METHODS

The Diagnostic Aspects of Incontinence Study (DAISy) was a prospective multicenter study. We enrolled participants from April 2004 to December 2004 at the Loyola University of Chicago; University of Alabama at Birmingham; University of California, San Francisco

See also:

Print
Editors' Notes 716
Summary for Patients I-30

Web-Only
CME quiz
Conversion of figures and tables into slides

Context

Clinicians and patients need simple, feasible ways to distinguish urge urinary incontinence from stress urinary incontinence.

Contribution

In this multicenter study, specialists evaluated and diagnosed incontinence in 301 middle-aged and older women who reported 3 or more episodes per week for at least 3 months. The authors developed a 3-item, self-administered questionnaire to classify women as having urge incontinence or stress incontinence. The answers to the questionnaire increased the likelihood of a final diagnosis of urge incontinence (positive likelihood ratio, 3.29) and stress incontinence (positive likelihood ratio, 2.13).

Implications

Asking patients some simple and quick questions may help clinicians distinguish urge urinary incontinence from stress urinary incontinence.

—The Editors

(UCSF); University of Iowa, Iowa City; and University of Texas Health Science Center at San Antonio. We selected these 5 U.S. clinical sites because of their broad experience with diagnosis and treatment of urinary incontinence and the availability of a clinically active urologist or urogynecologist. We designed the study to assess the reproducibility and accuracy of the 3IQ, with an extended evaluation as the gold standard, in classifying a broad spectrum of urinary incontinence by type. The local investigational review boards of the 5 clinical sites and of UCSF, where the study was coordinated, approved the study protocol. All participants provided written informed consent.

Participants

We recruited women through newspaper advertisements and flyers (93.7%) and from urology and gynecology clinics (6.3%). Interested participants called trained research assistants, who screened the women over the telephone by using a standardized script. We chose eligibility criteria to define a community-dwelling sample of women with incontinence who were appropriate for evaluation and treatment in primary care settings. Eligible women were ambulatory, were 40 years of age or older, reported 3 or more episodes of incontinence per week for at least 3 months, did not have urinary tract infection, and were bothered enough by their incontinence to seek treatment. We excluded women with incontinence who had complex problems that were more appropriate for specialist referral, including 4 or more urinary tract infections in the preceding year; pregnancy within 6 months; previous anti-incontinence or urethral surgery or procedures; previous major pelvic or abdominal surgery; pelvic radiation within 6 months; or known diseases of the genitourinary tract, such

Figure 1. **The 3 Incontinence Questions (*3IQ*).**

1. During the last 3 months, have you leaked urine (even a small amount)?

❑ Yes ❑ No → *Questionnaire completed.*

2. During the last 3 months, did you leak urine:
(Check all that apply.)

❑ a. When you were performing some physical activity, such as coughing, sneezing, lifting, or exercise?
❑ b. When you had the urge or the feeling that you needed to empty your bladder, but you could not get to the toilet fast enough?
❑ c. Without physical activity and without a sense of urgency?

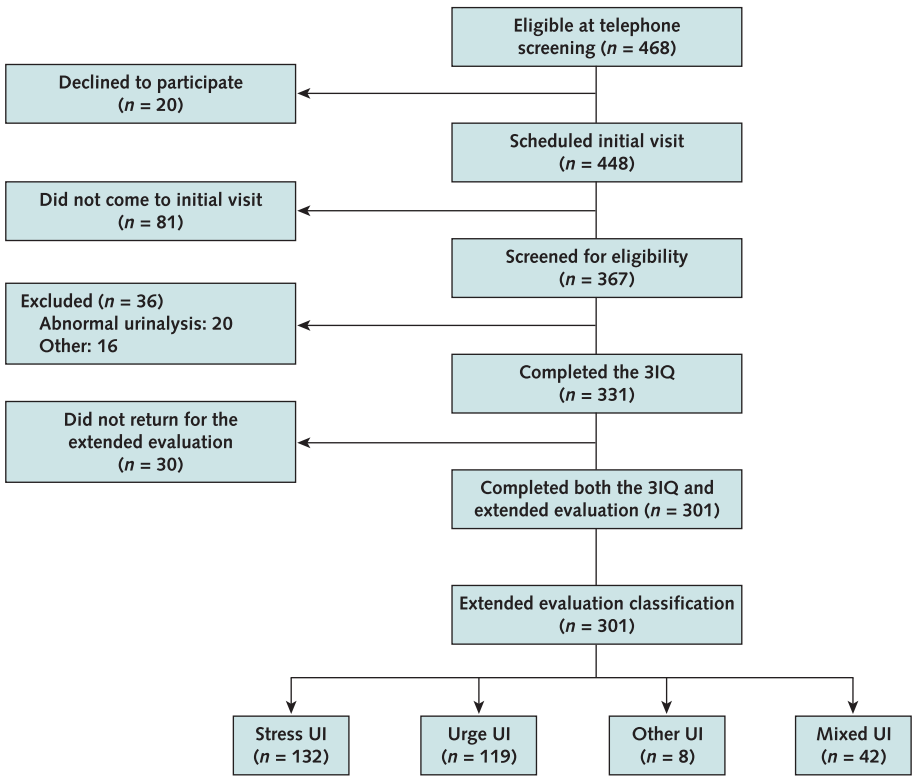
3. During the last 3 months, did you leak urine *most often*:
(Check only one.)

❑ a. When you were performing some physical activity, such as coughing, sneezing, lifting, or exercise?
❑ b. When you had the urge or the feeling that you needed to empty your bladder, but you could not get to the toilet fast enough?
❑ c. Without physical activity and without a sense of urgency?
❑ d. About equally as often with physical activity as with a sense of urgency?

Definitions of type of urinary incontinence are based on responses to question 3:

Response to Question 3	Type of Incontinence
a. Most often with physical activity	Stress only or stress predominant
b. Most often with the urge to empty the bladder	Urge only or urge predominant
c. Without physical activity or sense of urgency	Other cause only or other cause predominant
d. About equally with physical activity and sense of urgency	Mixed

Figure 2. **Flow diagram of study recruitment.**



3IQ = 3 Incontinence Questions; UI = urinary incontinence.

as lower urinary tract or rectal fistula, congenital abnormality leading to incontinence, interstitial cystitis, severe symptomatic pelvic prolapse, current or past urogenital cancer, spinal cord lesions, multiple sclerosis, stroke with clinically significant residual disability, Parkinson disease, or other major central nervous system abnormality affecting the lower urinary tract. Since we wanted to test the accuracy of the 3IQ in distinguishing between urge and stress incontinence in women with newly identified incontinence and no previous evaluation, we also excluded women who had been treated for incontinence in the previous 3 months.

Measurements

At the initial visit, participants completed the 3IQ questionnaire without assistance (**Figure 1**). The first question establishes that incontinence has occurred in the past 3 months. The second question familiarizes the women with the types of incontinence: stress (associated with physical activity), urge (associated with the feeling of urge or the need to empty the bladder), or other (occurs without activity or urge to empty the bladder). The third question determines the category of incontinence: stress, urge, other, or mixed (stress and urge equally).

Participants also completed questionnaires on age, race or ethnicity, education, menopause status, parity, previous hysterectomy, smoking habits, alcohol consumption, overall health status, age of onset of incontinence, and duration (in years) of incontinence. A dipstick urinalysis was performed. To determine the reproducibility of the 3IQ, participants completed a second 3IQ questionnaire at home, 7 to 10 days after the initial administration. After completing the second questionnaire, participants completed a 3-day voiding diary at home.

To assess the severity of incontinence and effect on quality of life, we asked participants to complete the Sandvik Severity Scale and a condition-specific quality-of-life instrument, the Incontinence Impact Questionnaire Short Form (IIQ-7), at a second visit. These instruments have been shown to be reliable, responsive to change over time, and valid and are widely used in incontinence research (12–14). The Sandvik Severity Scale is calculated from frequency and amount of urine loss on a scale of 1 to 12 (mild, 1 to 2; moderate, 3 to 6; severe, 8 to 9; or very severe, 12). The IIQ-7 is scored on a scale of 0 to 100, with a higher score representing greater negative effect on quality of life.

Each participant also underwent an extended evaluation that included a medical, surgical, reproductive, and incontinence history; review of all medications; physical examination, including evaluation of sacral nerves 2 to 4

Table 1. **Characteristics of 301 Participants in the Diagnostic Aspects of Incontinence Study (DAISy)**

Variable	Value
Mean (SD) age, *y*	56.4 (11.4)
Race or ethnicity, *n (%)*	
White	207 (68.8)
African American	38 (12.6)
Latina	36 (12.0)
Asian or Pacific Islander	7 (2.3)
Native American or other	13 (4.3)
Education, *n (%)*	
≤High school	65 (21.9)
Some college or bachelor's degree	177 (59.6)
Graduate degree	55 (18.5)
Parity, *n (%)*	
None	17 (6.4)
1–2	139 (52.3)
3–4	86 (32.3)
>4	24 (9.0)
Postmenopausal, *n (%)*	98 (32.7)
Hysterectomy, *n (%)*	103 (34.2)
Current smoking, *n (%)*	41 (13.7)
Alcohol consumption, *n (%)*	
None	132 (44.0)
<Weekly	88 (29.3)
≥Weekly	80 (26.7)
Health, *n (%)*	
Excellent	101 (33.6)
Good	150 (49.8)
Fair, poor, or very poor	50 (16.6)
Urinary incontinence	
Mean (SD) total incontinence episodes per wk, *n**	30.2 (27.4)
Mean (SD) duration of incontinence, *y*	7.0 (7.2)
Sandvik Severity Scale, *n (%)*	
Mild	16 (5.3)
Moderate	177 (58.8)
Severe	74 (24.6)
Very severe	34 (11.3)
Mean (SD) Incontinence Impact Questionnaire-7 score†	30.9 (25.1)
Urinary tract infections (any in past year), *n (%)*	46 (15.3)
Pelvic organ prolapse, *n (%)*‡	4 (1.3)
Any neurologic abnormality in sacral nerves 2 to 4, *n (%)*	20 (6.6)

* Based on a 3-day voiding diary.
† The Incontinence Impact Questionnaire-7 is scored on a scale of 0 to 100, with a higher score representing greater negative effect on quality of life.
‡ Any prolapse ≥1 cm beyond the hymen at pelvic examination.

(deep tendon reflexes, anal wink, perineal sensation, and bulbocavernosus reflex); pelvic examination; cough stress test; measurement of postvoid residual volume (by catheter or ultrasonography); and review of the participant's completed 3-day voiding diary. A urologist or urogynecologist at each clinical site supervised or performed the extended evaluation by using the site's own standard approach. One specialist performed all extended evaluations at 3 sites, 1 site used 2 specialists (University of Iowa), and 1 site used 3 specialists (Loyola University of Chicago). The urologist, urogynecologist, and supervised clinicians were blinded to the results of the 3IQ. The urologist or urogynecologist reviewed all elements of the extended evaluation and used clinical judgment to classify the participant by using the typical clinical categories of urge-only, urge-predominant, stress-only, stress-predominant, mixed, or other incontinence.

We sent the complete records of the extended evaluation (without the final clinical classification) to a randomly selected second specialist at another DAISy clinical site for review. The second specialist independently classified the type of incontinence. If disagreement about type of incontinence occurred, the 2 specialists discussed the case and arrived at a consensus diagnosis. Eighty-three disagreements in diagnosis required discussion for a consensus diagnosis. To be consistent with the categories based on the 3IQ (stress, urge, other, or mixed), we combined the extended evaluation categories of urge-only and urge-predominant incontinence as "urge incontinence" and the stress-only and stress-predominant incontinence categories as "stress incontinence."

Investigators and study staff monitored serious adverse events (such as life-threatening condition, cancer, or inpatient hospitalization). They asked participants about serious adverse events, recorded any reported event on a standardized form, and faxed the form to the UCSF coordinating center within 24 hours for review.

Treatment to improve both urge and stress incontinence is typically given to women with mixed incontinence. Thus, to test the accuracy of the 3IQ for selection of appropriate treatment, we assessed the accuracy of the questionnaire with respect to the extended evaluation (or gold standard test) to identify women with urge or mixed incontinence (vs. any other type of incontinence) and those with stress or mixed incontinence (vs. any other type of incontinence).

Statistical Analysis

We used the κ statistic, comparing the first and second set of 3IQ results to assess reproducibility (15). A κ statistic of 0.75 or greater indicates excellent agreement, and values between 0.40 to 0.75 indicates fair to good agreement (16). We calculated sensitivity, specificity, and positive and negative predictive values for the diagnosis of urge incontinence and stress incontinence by using appropriate sample proportions and computed 95% CIs by using the score method (17). We computed the positive and negative likelihood ratios as functions of the sample sensitivity and specificity, with 95% CIs computed on the log scale (18). We also estimated post-test probabilities of urge and stress incontinence with 95% CIs that account for uncertainty in the estimated sensitivity and specificity of the 3IQ (19), and we regarded pretest probabilities as fixed at 25%, 50%, and 75% depending on age (<40 years, 40 to 60 years,

Table 2. **Classification of Type of Urinary Incontinence by Extended Evaluation and 3IQ***

3IQ Classification	Extended Evaluation Classification				
	Stress UI, *n* (%)	Urge UI, *n* (%)	Other UI, *n* (%)	Mixed UI, *n* (%)	Total, *n*
Stress UI	102 (77.3)	17 (14.3)	2 (25.0)	17 (40.5)	138
Urge UI	12 (9.1)	68 (57.1)	1 (12.5)	10 (23.8)	91
Other UI	3 (2.3)	6 (5.0)	1 (12.5)	0	10
Mixed UI	15 (11.4)	28 (23.5)	4 (50.0)	15 (35.7)	62
Total	132	119	8	42	301

* Percentages may not add up to 100 because of rounding. 3IQ = 3 Incontinence Questions; UI = urinary incontinence.

and >60 years, respectively). We planned a sample size of 300 participants to achieve margins of sampling error of approximately 6 percentage points for both sensitivity and specificity for urge and stress incontinence, assuming that the sample prevalence of each would be approximately 50% and that sensitivity and specificity would be 80%. We performed all analyses by using SAS, version 9.1 (SAS Institute Inc., Cary, North Carolina).

Role of the Funding Source

Astellas Pharma US (formerly Yamanouchi Pharma America) funded the study through a research contract to UCSF. The funding source had no role in the design, conduct, analysis, or interpretation of the data or in the decision to submit the manuscript for publication.

RESULTS

During the telephone screening, 468 women were eligible for the study (**Figure 2**). Of these women, 331 provided informed consent. All 331 women completed the 3IQ questionnaire and were scheduled for an extended evaluation. Thirty of these women did not return for the extended evaluation (18 declined further participation, and 12 were lost to follow-up). The proportion of participants who did not return for the extended evaluation was similar (range, 8% to 12%) in each category of incontinence based on the first set of 3IQ results ($P = 0.47$). The 30 women who did not complete the study were more likely to be African American or Native American ($P = 0.034$), to report more alcohol use ($P = 0.087$), and to rate themselves as being in poor or very poor health ($P = 0.113$). All analyses included the 301 women who completed the 3IQ questionnaire and underwent the extended evaluation. No questionnaire or extended evaluation classification was incomplete or indeterminate. The 301 women were from the University of Texas Health Science Center at San Antonio ($n = 74$), University of Alabama at Birmingham ($n = 69$), UCSF ($n = 62$), University of Iowa ($n = 52$), and Loyola University of Chicago ($n = 44$).

The mean age of participants was 56.4 years (SD, 11.4; range, 40 to 94 years). The women were racially diverse (69% white, 13% African American, 12% Latina, and 2% Asian), and more than 78% had attended college or had an advanced degree. Most women reported good or excellent health (83%). The mean duration of incontinence was 7 years (SD, 7), and the range of incontinence severity was broad (5% mild, 59% moderate, 25% severe, and 11% very severe). **Table 1** lists other characteristics of the DAISy participants.

Participants completed the 3IQ questionnaire at the initial visit and repeated the questionnaire a mean of 7.4 days (SD, 1.0) later. The κ statistic for the reproducibility of the 3IQ is 0.69 (95% CI, 0.61 to 0.77) for urge incontinence (vs. not urge incontinence) and 0.65 (CI, 0.56 to 0.74) for stress incontinence (vs. not stress incontinence). The extended evaluation was completed a mean of 25.0 days (SD, 7.7) after the initial 3IQ. The extended evaluation was completed by using each site's own standard approach to perform the component tests. All participants had a medical, surgical, reproductive, and incontinence history; review of medications; physical examination, including a pelvic examination; and review of the 3-day voiding diary. Evaluation of sacral nerves 2 to 4 was completed in 99.7% of participants, cough stress test was performed in 98.7%, and postvoid residual volume was measured in 96.4%.

Table 2 presents the prevalence of urge (urge-only plus urge-predominant), stress (stress-only plus stress-predominant), mixed, and other incontinence on the basis of the results of the 3IQ and the extended evaluation. On the basis of the extended evaluation, 119 (39.5%) women had urge incontinence, 132 (43.9%) women had stress incontinence, 42 (14%) women had mixed incontinence, and 8 (2.7%) women had other types of incontinence.

Table 3 presents the accuracy estimates for classifica-

Table 3. **Accuracy of the 3IQ Compared with the Extended Evaluation***

Variable	Urge Incontinence	Stress Incontinence
Sensitivity (95% CI)	0.75 (0.68–0.81)	0.86 (0.79–0.90)
Specificity (95% CI)	0.77 (0.69–0.84)	0.60 (0.51–0.68)
Positive likelihood ratio (95% CI)	3.29 (2.39–4.51)	2.13 (1.71–2.66)
Negative likelihood ratio (95% CI)	0.32 (0.24–0.43)	0.24 (0.16–0.35)

* 3IQ = 3 Incontinence Questions.

Table 4. **Post-Test Probability of Stress or Urge Incontinence among Women with Incontinence and Positive or Negative 3IQ Results***

Age	Prevalence of Incontinence (Pretest Probability), %		Post-Test Probability (3IQ Response), %	
	Urge UI	Stress UI	Positive for Urge Incontinence (95% CI)	Positive for Stress Incontinence (95% CI)
<40 y	25	75	52.1 (44.2–59.9)	86.6 (83.8–88.9)
40–60 y	50	50	76.5 (70.4–81.7)	68.3 (63.3–72.8)
>60 y	75	25	90.7 (87.7–93.1)	41.7 (36.5–47.2)

* The post-test probabilities and 95% CIs were estimated assuming fixed pretest probabilities and account for uncertain sensitivities and specificities. The post-test probability was estimated by using the relationship between post-test odds and pretest odds (post-test odds = likelihood ratio × pretest odds). 3IQ = 3 Incontinence Questions; UI = urinary incontinence.

tion of urge and stress incontinence on the basis of the results of the 3IQ compared with the gold standard, extended evaluation. For urge incontinence, the 3IQ had a sensitivity of 0.75 (CI, 0.68 to 0.81), a specificity of 0.77 (CI, 0.69 to 0.84), and a positive likelihood ratio of 3.29 (CI, 2.39 to 4.51). For classification of stress incontinence, the 3IQ had a sensitivity of 0.86 (CI, 0.79 to 0.90), a specificity of 0.60 (CI, 0.51 to 0.68), and a positive likelihood ratio of 2.13 (CI, 1.71 to 2.66). Using logistic models to evaluate heterogeneity in the sensitivity and specificity by site, we found no evidence for heterogeneity in sensitivity for either urge (range across 5 sites, 0.71 to 0.79; $P = 0.93$) or stress (range, 0.80 to 0.91; $P = 0.66$) incontinence, and no heterogeneity in specificity for urge incontinence (range, 0.71 to 0.82; $P = 0.65$). The estimates of specificity for stress incontinence suggested some heterogeneity (estimates for the 5 sites were 0.41, 0.50, 0.61, 0.71, and 0.76, respectively; $P = 0.076$). The range of type classification by study site was 48% to 73% for urge plus mixed incontinence, as well as for stress plus mixed incontinence.

On the extended evaluation, 8 (2.7%) participants were classified as having other type of incontinence. Of these participants, 1 participant was classified as having other incontinence, 4 were classified as having mixed incontinence, 2 were classified as having stress incontinence, and 1 was classified as having urge incontinence on the 3IQ. We queried the 2 urologists or urogynecologists who made the original diagnosis of other incontinence on the extended evaluation to determine whether basing treatment on the 3IQ results could have endangered the women because of a delay in correct diagnosis. For the 7 participants who were not identified on the 3IQ as having other incontinence, all specialists agreed that treatment based on the 3IQ results and a delay in diagnosis for 6 to 12 months were not dangerous.

Postvoid residual volumes were normal (<150 mL) in 283 (97.6%) participants. Seven participants had postvoid residual volumes greater than 150 mL. Repeated measurement of postvoid residual volume in 4 of these participants was normal. Three participants with an abnormal postvoid residual volume chose not to return for repeated measurement. The initial postvoid residual volumes in these participants were 155 mL, 157 mL, and 220 mL, respectively. All specialists agreed that treatment based on the 3IQ results and a delay in diagnosis were not dangerous in the participants with mildly abnormal postvoid residual volumes of 155 mL and 157 mL, and 1 specialist felt that the woman with a postvoid residual volume of 220 mL might be at increased risk for urinary tract infection.

We monitored serious adverse events during the study. Five adverse events were reported, and we assessed them as unrelated to the study.

DISCUSSION

Despite the availability of effective treatments, primary care clinicians often do not inquire about incontinence or recommend treatment (8, 20, 21). We believe that this is partly because national guidelines (8, 9) recommend an extended evaluation for classifying type of incontinence that is not practical in primary care settings. The 3IQ questionnaire is a simple, quick, and reproducible test with acceptable accuracy for classifying urge and stress incontinence among women who are appropriate for evaluation and treatment in primary care settings. Using the 3IQ questionnaire might encourage primary care physicians to classify and treat incontinence in women.

The reproducibility of the 3IQ was fair to good, with κ statistics ranging from 0.65 to 0.69. For urge incontinence, the sensitivity was about 75%, which suggests that approximately 25% of women with urge incontinence will be missed by using the 3IQ; the specificity was 77%, which suggests that about 23% of women with other types of incontinence will be inappropriately treated for urge incontinence. For stress incontinence, the sensitivity was higher (86%), suggesting that only approximately 14% of women with stress incontinence will be missed by using the 3IQ, but the specificity was 60%, which suggests that about 40% of women with other types of incontinence may be inappropriately treated for stress incontinence.

The probability that a woman with incontinence has

urge rather than stress incontinence increases with age (2, 22, 23). Using this information, we estimated pretest probabilities of urge or stress incontinence according to age and calculated the post-test probability of urge or stress incontinence depending on the 3IQ results (**Table 4**). Among young women with incontinence, approximately three quarters have stress incontinence (2, 24). Among those classified by the 3IQ as having urge incontinence, the probability of having urge incontinence is increased from 25% (pretest) to 52% (post-test). In contrast, among those classified as having stress incontinence by the 3IQ, the probability of having stress incontinence is increased slightly from 75% (pretest) to 87% (post-test). Among women of middle age, prevalence of urge and stress incontinence are about equal. In this group, probabilities for both urge and stress incontinence are increased similarly from 50% (pretest) to 77% (post-test) for urge incontinence and to 68% (post-test) for stress incontinence, according to the test result. Among older women, where urge incontinence accounts for three quarters of total prevalence, estimated post-test probability of urge and stress incontinence is 91% and 42%, respectively.

The accuracy of the 3IQ is modest but acceptable given that the risk for misclassification and inappropriate treatment by primary care physicians is low. Urge-suppression training for urge incontinence and pelvic floor muscle–strengthening exercises for stress incontinence are safe and effective treatments that can be learned by using a self-help booklet. Misclassifying a patient, however, may result in patient inconvenience and unnecessary expenditure of resources. Treatment with antimuscarinic or anticholinergic drugs for urge incontinence can cause dry mouth but is rarely associated with more clinically significant adverse effects, such as urinary retention. Stress incontinence that is unresponsive to behavioral therapy may be treated with surgery, but surgery would not be performed without first completing an extended evaluation. Given the simplicity and ease of use of the 3IQ questionnaire and the fact that it avoids an invasive and expensive evaluation, we believe that the accuracy documented in our study is acceptable.

Seven participants in our study who were classified by the extended evaluation as having other incontinence were not identified by the 3IQ, and 7 participants had abnormal postvoid residual volumes. Urologists and urogynecologists reviewed these cases and determined that no danger to the participants was related to the missed diagnosis, abnormal postvoid residual volume, or delay in appropriate treatment. These women would have been referred for specialty care under current guidelines. In addition, women who do not improve after several months of therapy should be referred. This proportion is difficult to estimate but will be markedly lower than the 100% that would be recommended for referral to specialty care under current guidelines.

On the basis of a MEDLINE search from January 1965 through October 2005, we identified 7 published studies that evaluated the accuracy of questionnaires to classify type of incontinence (**Table 5**). The questionnaires used in 6 of the studies were substantially longer (5 to 18 questions) than the 3IQ, classification required calculation of a score, and all were evaluated in specialty clinic populations at a single site. Sandvik and colleagues' study (10) included 250 women referred to a Norwegian gynecology clinic for evaluation of incontinence. A nurse asked 2 questions to classify incontinence as stress-only, urge-only, or mixed. The test results were compared with a gynecologist's classification that was based on an extensive evaluation (10). The accuracy of the test was fair, with especially good specificity. However, the study was conducted at a single gynecology site, and a trained interviewer administered the questions. In addition, women with predomi-

Table 5. **Published Studies Evaluating the Accuracy of Questionnaires To Classify Type of Urinary Incontinence in Women***

Author, Year (Reference)	Location	Mean Age, *y*	Patients, *n*	Setting	Questions, *n*	Question Administration	Study Design	"Gold Standard"	Blinded Evaluation	Sensitivity/Specificity	
										Stress UI	Urge UI
Sandvik et al., 1995 (10)	Norway	≥20 (range, NR)	250	Gynecology clinic	2	Interviewer	Prospective	Extensive evaluation†	Yes	0.66/0.88	0.56/0.96
Klovning et al., 1996 (25)	Norway	49.2 (SD, 0.9) (range, 15–83)	250	Gynecology clinic	10	Interviewer	Prospective	Extensive evaluation†	Yes	0.77/0.52	Not determined
Lemack and Zimmern, 1999 (26)	United States	61 (range, 27–86)	128	Urology clinic	6	Self-reported	Retrospective; chart abstraction	Urodynamic diagnosis	No	0.85/0.63	0.83/0.50
Kirschner-Hermanns et al., 1998 (27)	United States	75 (range, 65–98)	132	Geriatric clinic	5	Interviewer	Prospective	Urodynamic diagnosis	Yes	0.60/0.96	0.84/0.21
Diokno et al., 1999 (28)	United States	NR	118	Urology clinic	18	Self-reported	Retrospective; chart abstraction	Extensive evaluation†	Yes	0.62/0.1	Not determined
Ishiko et al., 2000 (29)	Japan	59 (SD, 12) (range, 27–73)	191	Urinary dysfunction clinic	15	Self-reported	Prospective	Urodynamic diagnosis	No	0.83/not determined	0.86/not determined
Bradley et al., 2005 (30)	United States	56 (range, 22–87)	117	Urogynecology clinic	6	Self-reported	Prospective	Extensive evaluation†	Yes	0.85/0.71	0.79/0.79

* NR = not reported; UI = urinary incontinence.
† Extensive evaluation contained most of these components in all studies: medical, surgical, and incontinence history; physical evaluation, including evaluation of sacral nerves 2 to 4; pelvic examination; cough stress test; measurement of postvoid residual volume; voiding diary; urodynamic studies; and/or cystoscopy.

nantly urge or stress incontinence were classified as having mixed incontinence, which does not provide direction for treatment.

We aimed to improve the questionnaire used by Sandvik and colleagues and to replicate the findings in a more diverse population of women. For the 3IQ questionnaire, we added an initial question to establish that incontinence occurred in the previous 3 months. The second question incorporates both questions used in the Sandvik questionnaire. We also added a third question to classify incontinence in categories to direct initial treatment: urge (urge-only, urge-predominant, or mixed) or stress (stress-only, stress-predominant, or mixed) incontinence. We designed DAISy to determine the accuracy of the 3IQ among women with incontinence who were most appropriate for treatment by primary care providers, with a broad spectrum of type and severity of incontinence, and we tested the 3IQ questionnaire at 5 sites across the United States.

The accuracy of Sandvik and colleagues' questionnaire and the 3IQ was similar, although the 3IQ questionnaire was more sensitive but less specific. These differences could be because a nurse administered Sandvik and colleagues' questionnaire and the questionnaire limited the categorization of incontinence to urge-only, stress-only, or mixed. These categories, although less clinically useful, are more specific. Participants were enrolled at academic medical centers, and urologists and urogynecologists led the study. In contrast, only 23 women in our study (6%) were recruited from specialty clinics, and we designed our inclusion and exclusion criteria to reflect patients typically seen in primary care settings. We excluded women with active urinary tract infections because infection may cause or worsen all types of incontinence and treatment often markedly improves incontinence. While our results are generalizable to racially diverse women with a broad spectrum of incontinence severity, our study did not include women with complex urinary incontinence (such as those with neurologic problems or those whose surgery failed), women who did not speak English, or men. The 3IQ may not be accurate in these populations. In addition, our study does not provide information on the number of extended evaluations that would occur if treatment were based on the results of the 3IQ or on the clinical outcomes that would result from using the 3IQ.

In the primary care setting, evaluation of incontinence should include a urine screening to exclude urinary tract infection and hematuria and the 3IQ to classify type of incontinence. If urinalysis results are normal, treatment of urge, stress, or mixed incontinence can be based on the results of the 3IQ. Patients with other incontinence and those with complex urinary tract problems should be referred to a specialist for an extended evaluation. The initial treatment for both urge and stress incontinence is behavioral (reduced oral fluid intake, regular voiding, pelvic muscle–strengthening exercises for stress incontinence, and urge-suppression exercises for urge incontinence). Antimuscarinic or anticholinergic medications are effective for treating urge incontinence. If primary care management does not adequately control incontinence after 6 to 12 months, patients should be referred for evaluation and treatment by an incontinence specialist.

In summary, the 3IQ questionnaire is a simple, quick, and noninvasive test with acceptable accuracy for classifying urge and stress incontinence among the mostly middle-aged women included in our study. Our findings should be replicated in other primary care clinical settings. In addition, clinical outcomes should be assessed in a trial comparing treatments based on the 3IQ and the extended evaluation.

From the University of California, San Francisco, and San Francisco Veterans Affairs Medical Center, San Francisco, California; University of Iowa, Iowa City, Iowa; University of Alabama at Birmingham, Birmingham, Alabama; University of Texas Health Science Center at San Antonio, San Antonio, Texas; and Loyola University of Chicago, Maywood, Illinois.

Grant Support: By a UCSF research contract from Astellas Pharma US (formerly Yamanouchi Pharma America).

Potential Financial Conflicts of Interest: *Honoraria:* C.S. Bradley (Pfizer Inc.), S.R. Kraus (Pfizer Inc., Astellas, Novartis); *Grants received:* C.S. Bradley (Yamanouchi), S.R. Kraus (National Institutes of Health, Astellas, GlaxoSmithKline).

Requests for Single Reprints: Jeanette S. Brown, MD, University of California, San Francisco, 1635 Divisadero Street, Suite 600, San Francisco, CA 94115; e-mail, brownj@obgyn.ucsf.edu.

Current author addresses and author contributions are available at www.annals.org.

References

1. **Brown JS, Vittinghoff E, Wyman JF, Stone KL, Nevitt MC, Ensrud KE, et al.** Urinary incontinence: does it increase risk for falls and fractures? Study of Osteoporotic Fractures Research Group. J Am Geriatr Soc. 2000;48:721-5. [PMID: 10894308]

2. **Thom D.** Variation in estimates of urinary incontinence prevalence in the community: effects of differences in definition, population characteristics, and study type. J Am Geriatr Soc. 1998;46:473-80. [PMID: 9560071]

3. **Thom DH, Haan MN, Van Den Eeden SK.** Medically recognized urinary incontinence and risks of hospitalization, nursing home admission and mortality. Age Ageing. 1997;26:367-74. [PMID: 9351481]

4. **Roe B, Williams K, Palmer M.** Bladder training for urinary incontinence in adults. Cochrane Database Syst Rev. 2000:CD001308. [PMID: 10796768]

5. **Haeusler G, Leitich H, van Trotsenburg M, Kaider A, Tempfer CB.** Drug therapy of urinary urge incontinence: a systematic review. Obstet Gynecol. 2002; 100:1003-16. [PMID: 12423868]

6. **Hay-Smith EJ, Bø Berghmans LC, Hendriks HJ, de Bie RA, van Waalwijk van Doorn ES.** Pelvic floor muscle training for urinary incontinence in women. Cochrane Database Syst Rev. 2001:CD001407. [PMID: 11279716]

7. **Holroyd-Leduc JM, Straus SE.** Management of urinary incontinence in women: scientific review. JAMA. 2004;291:986-95. [PMID: 14982915]

8. Managing acute and chronic urinary incontinence. AHCPR Urinary Incontinence in Adults Guideline Update Panel. Am Fam Physician. 1996;54:1661-72. [PMID: 8857788]

9. Assessment and treatment of urinary incontinence. Scientific Committee of the First International Consultation on Incontinence. Lancet. 2000;355:2153-8.

[PMID: 10902644]
10. **Sandvik H, Hunskaar S, Vanvik A, Bratt H, Seim A, Hermstad R.** Diagnostic classification of female urinary incontinence: an epidemiological survey corrected for validity. J Clin Epidemiol. 1995;48:339-43. [PMID: 7897455]
11. **Lagro-Janssen AL, Debruyne FM, van Weel C.** Value of the patient's case history in diagnosing urinary incontinence in general practice. Br J Urol. 1991; 67:569-72. [PMID: 2070199]
12. **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. [PMID: 8120507]
13. **Hanley J, Capewell A, Hagen S.** Validity study of the severity index, a simple measure of urinary incontinence in women. BMJ. 2001;322:1096-7. [PMID: 11337439]
14. **Uebersax JS, Wyman JF, Shumaker SA, McClish DK, Fantl JA.** Short forms to assess life quality and symptom distress for urinary incontinence in women: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Continence Program for Women Research Group. Neurourol Urodyn. 1995;14:131-9. [PMID: 7780440]
15. **Fleiss JL.** Statistical Methods for Rates and Proportions. New York: J Wiley; 1981.
16. **Landis JR, Koch GG.** The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-74. [PMID: 843571]
17. **Newcombe RG.** Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med. 1998;17:857-72. [PMID: 9595616]
18. **Simel DL, Samsa GP, Matchar DB.** Likelihood ratios with confidence: sample size estimation for diagnostic test studies. J Clin Epidemiol. 1991;44:763-70. [PMID: 1941027]
19. **Monsour MJ, Evans AT, Kupper LL.** Confidence intervals for post-test probability. Stat Med. 1991;10:443-56. [PMID: 2028127]
20. **Jones TV, Bunner SH.** Approaches to urinary incontinence in a rural population: a comparison of physician assistants, nurse practitioners, and family physicians. J Am Board Fam Pract. 1998;11:207-15. [PMID: 9625512]
21. **Shaw C, Tansey R, Jackson C, Hyde C, Allan R.** Barriers to help seeking in people with urinary symptoms. Fam Pract. 2001;18:48-52. [PMID: 11145628]
22. **Hannestad YS, Rortveit G, Sandvik H, Hunskaar S.** A community-based epidemiological survey of female urinary incontinence: the Norwegian EPINCONT study. Epidemiology of Incontinence in the County of Nord-Trøndelag. J Clin Epidemiol. 2000;53:1150-7. [PMID: 11106889]
23. **Brown JS, Grady D, Ouslander JG, Herzog AR, Varner RE, Posner SF.** Prevalence of urinary incontinence and associated risk factors in postmenopausal women. Heart & Estrogen/Progestin Replacement Study (HERS) Research Group. Obstet Gynecol. 1999;94:66-70. [PMID: 10389720]
24. **Hunskaar S, Arnold EP, Burgio K, Diokno AC, Herzog AR, Mallett VT.** Epidemiology and natural history of urinary incontinence. Int Urogynecol J Pelvic Floor Dysfunct. 2000;11:301-19. [PMID: 11052566]
25. **Klovning A, Hunskaar S, Eriksen BC.** Validity of a scored urological history in detecting detrusor instability in female urinary incontinence. Acta Obstet Gynecol Scand. 1996;75:941-5. [PMID: 9003097]
26. **Lemack GE, Zimmern PE.** Predictability of urodynamic findings based on the Urogenital Distress Inventory-6 questionnaire. Urology. 1999;54:461-6. [PMID: 10475355]
27. **Kirschner-Hermanns R, Scherr PA, Branch LG, Wetle T, Resnick NM.** Accuracy of survey questions for geriatric urinary incontinence. J Urol. 1998;159: 1903-8. [PMID: 9598484]
28. **Diokno AC, Dimaculangan RR, Lim EU, Steinert BW.** Office based criteria for predicting type II stress incontinence without further evaluation studies. J Urol. 1999;161:1263-7. [PMID: 10081882]
29. **Ishiko O, Hirai K, Sumi T, Nishimura S, Ogita S.** The urinary incontinence score in the diagnosis of female urinary incontinence. Int J Gynaecol Obstet. 2000;68:131-7. [PMID: 10717817]
30. **Bradley CS, Rovner ES, Morgan MA, Berlin M, Novi JM, Shea JA, et al.** A new questionnaire for urinary incontinence diagnosis in women: development and testing. Am J Obstet Gynecol. 2005;192:66-73. [PMID: 15672005]

Current Author Addresses: Drs. Brown, Subak, Vittinghoff, and Grady: University of California, San Francisco, 1635 Divisadero Street, Suite 600, San Francisco, CA 94115.
Dr. Bradley: University of Iowa Hospitals and Clinics, 200 Hawkins Drive, Iowa City, IA 52242.
Dr. Richter: University of Alabama at Birmingham, 618 20th Street South, NHB 219, Birmingham, AL 35233.
Dr. Kraus: Department of Urology, University of Texas Health Science Center at San Antonio, MC 7845, 7703 Floyd Curl Drive, San Antonio, TX 78229.
Dr. Brubaker: Department of Obstetrics-Gynecology and Urology, Loyola University Medical Center, 2160 South First Avenue, Maywood, IL 60153.
Ms. Lin: University of California, San Francisco, 185 Berry Street, Lobby 4, 5th Floor, Suite 5700, San Francisco, CA 94107.

Author Contributions: Conception and design: J.S. Brown, C.S. Bradley, L.L. Subak, S.R. Kraus, E. Vittinghoff, D. Grady.
Analysis and interpretation of the data: J.S. Brown, C.S. Bradley, L.L. Subak, S.R. Kraus, L. Brubaker, F. Lin, E. Vittinghoff, D. Grady.
Drafting of the article: J.S. Brown, S.R. Kraus, E. Vittinghoff.
Critical revision of the article for important intellectual content: J.S. Brown, C.S. Bradley, H.E. Richter, S.R. Kraus, L. Brubaker, E. Vittinghoff, D. Grady.
Final approval of the article: J.S. Brown, C.S. Bradley, L.L. Subak, H.E. Richter, S.R. Kraus, L. Brubaker, F. Lin, E. Vittinghoff, D. Grady.
Provision of study materials or patients: C.S. Bradley, L.L. Subak, H.E. Richter, L. Brubaker.
Statistical expertise: J.S. Brown, F. Lin, E. Vittinghoff, D. Grady.
Obtaining of funding: J.S. Brown, D. Grady.
Administrative, technical, or logistic support: J.S. Brown, D. Grady.
Collection and assembly of data: J.S. Brown, C.S. Bradley, S.R. Kraus, L. Brubaker, D. Grady.