Annals of Internal Medicine REVIEW

Screening for Urinary Incontinence in Women: A Systematic Review for the Women's Preventive Services Initiative

Heidi D. Nelson, MD, MPH; Amy Cantor, MD, MPH; Miranda Pappas, MA; and Liev Miller, BA

Background: Urinary incontinence is infrequently addressed during routine health care despite its high prevalence and adverse effects on health.

Purpose: To evaluate whether screening for urinary incontinence in women not previously diagnosed improves outcomes (symptoms, quality of life, and function) and to evaluate the accuracy of screening methods and potential harms of screening.

Data Sources: English-language searches of Ovid MEDLINE, Cochrane Central Register of Controlled Trials, and Cochrane Database of Systematic Reviews (1 January 1996 to 30 March 2018); ClinicalTrials.gov (April 2018); and reference lists of studies and reviews.

Study Selection: Randomized trials, cohort studies, systematic reviews of studies that enrolled nonpregnant women without previously diagnosed urinary incontinence and compared clinical outcomes and adverse effects between women who were and were not screened, and diagnostic accuracy studies that reported performance measures of screening tests.

Data Extraction: Dual extraction and quality assessment of individual studies.

Data Synthesis: No studies evaluated the overall effectiveness or harms of screening. Seventeen studies evaluated the diagnostic accuracy of 18 screening questionnaires against a clinical diagnosis or results of diagnostic tests. Of these, 14 poor-quality studies were based in referral clinics, enrolled only symptomatic women, or had other limitations. One good-quality and 2 fair-quality studies (evaluating 4 methods) enrolled women not recruited on the basis of symptoms. Areas under the receiver-operating characteristic curve for stress, urge, and any type of incontinence in these studies were 0.79, 0.88, and 0.88 for the Michigan Incontinence Symptom Index; 0.85, 0.83, and 0.87 for the Bladder Control Self-Assessment Questionnaire; and 0.68, 0.82, and 0.75 for the Overactive Bladder Awareness Tool. The Incontinence Screening Questionnaire had a sensitivity of 66% and specificity of 80% for any type of incontinence.

Limitation: Studies enrolled few participants, often from symptomatic referral populations; used various reference standards; and infrequently reported CIs.

Conclusion: Evidence is insufficient on the overall effectiveness and harms of screening for urinary incontinence in women. Limited evidence in general populations suggests fairly high accuracy for some screening methods.

Primary Funding Source: Health Resources and Services Administration.

Ann Intern Med. 2018;169:311-319. doi:10.7326/M18-0225 **Annals.org**
For author affiliations, see end of text.
This article was published at Annals.org on 14 August 2018.

Urinary incontinence adversely affects women's health through increased risks for urinary tract infections, skin ulceration, falls, and fractures (1). It interferes with work and social activities, sexual function, quality of life, and independence (2). Urinary incontinence, the involuntary loss of urine (3), is characterized by different types (4). Stress incontinence is the inability to retain urine during physical exertion or activities that increase intra-abdominal pressure, such as coughing or sneezing, and results from impaired sphincter function (4). Urge incontinence is associated with the sensation of a sudden urge to void and usually results from a rise in bladder pressure caused by contraction, overactivity, or dysfunction of the detrusor muscle (4). The term "overactive bladder" refers to urinary urgency with or without incontinence, usually accompanied by frequent and nighttime voiding (4). Mixed urinary incontinence includes both stress and urge incontinence.

Approximately 25% of reproductive-aged women (5), 44% to 57% of middle-aged and postmenopausal women (6), and 75% of older women experience some involuntary urine loss (7). Stress incontinence is more common in younger women with pelvic floor trauma and uterine prolapse related to previous vaginal delivery (8). Urgency and mixed incontinence are more common in older women in association with overactive bladder (3, 8). Rates are higher for women with specific risk factors, particularly obesity (9, 10) and previous vaginal delivery (11), whereas age alone may not be a risk factor independent of other comorbid conditions (12).

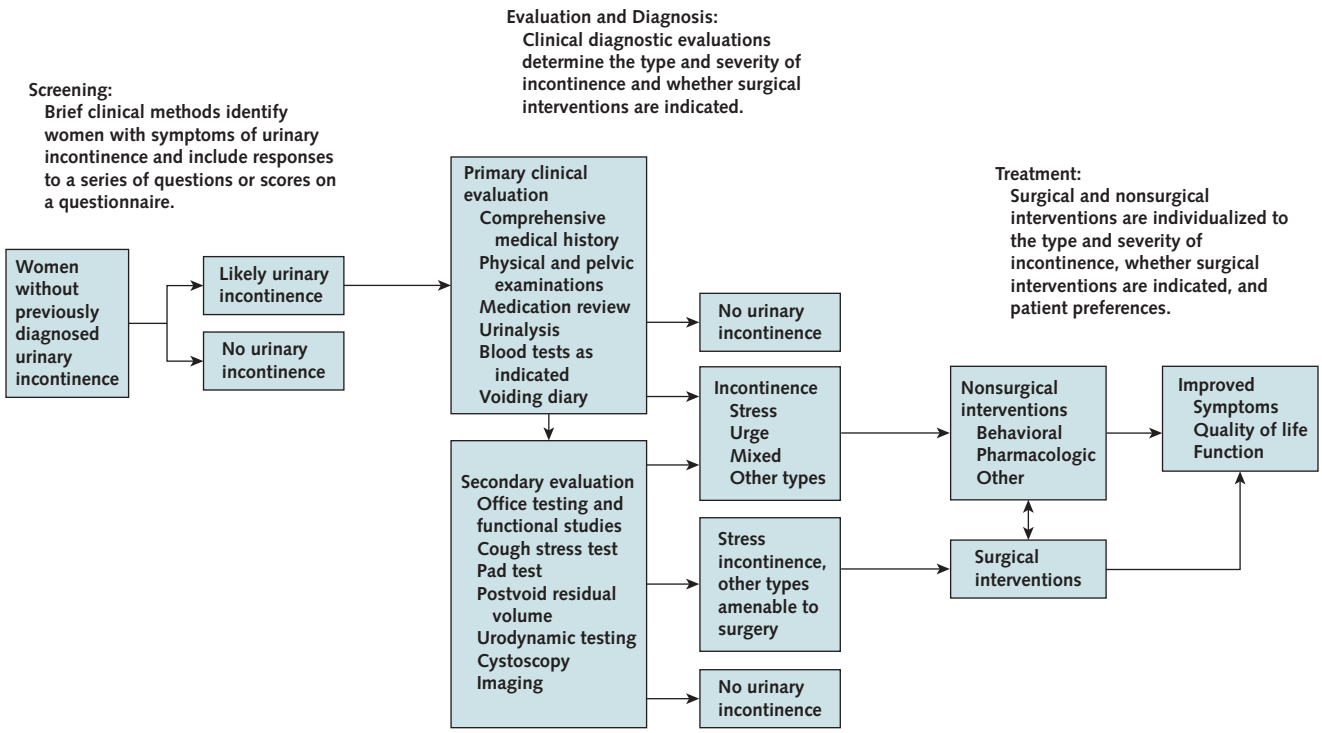
Urinary incontinence is infrequently addressed during routine health care despite its high prevalence and associated symptoms (13). Women may be reluctant to discuss incontinence because of embarrassment (14), social stigma, normalization of symptoms, lack of knowledge about treatment options (15), or concerns about surgery. In addition, most clinicians do not routinely inquire about incontinence, which may reach their attention only if a woman seeks help (16). Of women who ultimately seek medical attention, 30% are not evaluated for symptoms and 80% are not treated (13, 17).

See also:

Related article 320
Editorial comment 345

Web-Only
Supplement
CME/MOC activity

Figure. Screening for urinary incontinence clinical pathway.



No clinical practice guidelines include screening for urinary incontinence; existing recommendations involve the diagnostic evaluation and treatment (18–21). This systematic review focuses on this gap by evaluating evidence on the overall effectiveness of urinary incontinence screening in improving symptoms, quality of life, and function, as well as evidence on the accuracy of screening methods and potential harms of screening.

METHODS

The Women's Preventive Services Initiative Advisory Panel determined the scope and key questions for this review to inform the development of new screening recommendations (22). The protocol (**Supplement**, available at Annals.org) was developed using established methods (23) with input from experts and the public. The target population for screening is women who are not currently pregnant and have not previously been diagnosed with urinary incontinence. This review addressed the following key questions:

1. Does routine screening for urinary incontinence in women not previously diagnosed lead to improved symptoms, function, or quality of life?

2a. Among women not previously diagnosed with urinary incontinence, how accurate are screening methods? How does accuracy vary with age, sociodemographic characteristics, cultural group, comorbid conditions, or use of additional medications?

2b. What are the potential adverse effects of screening for urinary incontinence in women not previously diagnosed?

Screening begins with a series of questions or scores on a questionnaire designed to identify women with symptoms of urinary incontinence (**Figure**). Once identified, women with symptoms may require individualized clinical diagnostic evaluations to determine the type and severity of incontinence and appropriate options for treatment and management (24–26). The work-up includes a clinical evaluation that may be followed by office testing and functional studies, such as a cough stress test, pad test, postvoid residual volume measurement, urodynamic test, cystoscopy, or imaging. Treatment includes behavioral, pharmacologic (24, 25), surgical, and other interventions (24) specific to the type and severity of incontinence and patient preferences. Studies addressing key question 1 evaluate the overall effectiveness of screening, including subsequent diagnosis and treatment, on improving health outcomes. Studies addressing key questions 2a and 2b, as indicated in the first step in the **Figure**, evaluate the accuracy of screening methods and potential harms of screening. Other components of the clinical pathway (effectiveness and harms of diagnostic methods and treatments) have been reviewed elsewhere (24–26) and are outside the scope of this review.

Data Sources and Searches

A research librarian electronically searched Ovid MEDLINE, Cochrane Central Register of Controlled Trials, and Cochrane Database of Systematic Reviews from 1 January 1996 to 30 March 2018 for English-language articles on screening for urinary incontinence (**Appendix**, available at Annals.org). Investigators reviewed ClinicalTrials.gov for ongoing studies that may be relevant, although the only screening trial began recruitment in April 2018. Reference lists of systematic reviews and articles were also manually reviewed for relevance.

Study Selection

Using a 2-step process, 2 investigators independently evaluated abstracts and then full-text articles to identify studies meeting prespecified eligibility criteria (**Supplement**). Disagreements were resolved by consensus. We tracked results in an EndNote database (Clarivate Analytics).

Eligible studies were randomized controlled trials, large prospective cohort studies (>100 participants), diagnostic accuracy studies, or systematic reviews of studies that enrolled nonpregnant women without previously diagnosed urinary incontinence and provided data relevant to any of the key questions. We selected studies that used screening methods applicable to primary care settings in the United States (such as self-reported questionnaires, interviews by professionals, instrumental tests, and combinations thereof). Studies were relevant for screening if a test or measure was used to identify women with symptoms of urinary incontinence before diagnostic evaluations began. In most studies, the intention of the test (screening or diagnosis) was not clearly stated.

For questions about effectiveness, we considered studies comparing screening with usual care or alternative methods (such as telephone- or Web-based data collection vs. clinic-based assessments) that reported clinical outcomes, such as symptoms, quality of life, or function (days of disability, limitations in activity, absences, or other). For the diagnostic accuracy question, we included studies that reported measures of test performance, such as areas under the receiver-operating characteristic curve (AUROC) (also known as the c-statistic), sensitivity and specificity, or likelihood ratios. Potential adverse effects of screening that we evaluated included false-positive or false-negative results, anxiety, distress, and other adverse events affecting quality of life.

Data Extraction and Quality Assessment

A single investigator extracted data on characteristics of study populations, interventions, comparators, outcomes (including findings related to population subgroups), study designs, settings, and methods. A second investigator verified the completeness and accuracy of extracted data. Two investigators independently rated the quality of individual studies as good, fair, or poor using predefined criteria adapted from the U.S. Preventive Services Task Force (27). Disagreements were resolved by consensus with a third reviewer.

Critical appraisal criteria for the diagnostic accuracy of screening tests were based on U.S. Preventive Services Task Force methods (27), which are similar to other established methods, including QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies 2) (28, 29). Criteria included number (>50) and spectrum of patients, use of random or consecutive samples, description of participant eligibility criteria, low attrition, description and reproducibility of the test, use of a credible and replicable reference standard, blinding of outcome assessors to the reference standard, and application of the reference standard to all participants or a random subsample. If a study enrolled only women with symptoms of incontinence, its patient spectrum was considered inadequate for population screening. (Because prevention involves identification of incontinence before diagnostic evaluation, studies that included symptomatic women who were not yet diagnosed were included in our review but downgraded in quality ratings.) Urinary incontinence has no established gold standard test; studies that used combinations of common diagnostic tests as the reference standard were considered adequate. For each critical appraisal item, study quality was downgraded if the study did not meet the criterion or the report was unclear.

Studies were also evaluated for clinical applicability, which was rated as high if participants were selected from the community or primary care or nonspecialty clinics; the test and reference standards were relevant to U.S. clinical practice; and the test was feasible for screening by nonspecialists in clinical settings. Applicability was rated as low if these conditions were not met or described.

Data Synthesis and Analysis

We developed evidence tables of study characteristics, results, and quality ratings and summarized findings qualitatively. We organized diagnostic accuracy studies into 2 groups based on whether participants were recruited 1) from the community or primary care or nonspecialty clinics or 2) from referral clinics. We emphasized findings of fair- and good-quality studies and did not perform statistical meta-analyses because of heterogeneity in screening methods.

Role of the Funding Source

This research was funded by the Health Resources and Services Administration of the U.S. Department of Health and Human Services, which had no role in defining questions, developing the protocol, carrying out the review, interpreting data, developing conclusions, or submitting the review for publication. The investigators are solely responsible for the content and the decision to submit the manuscript for publication.

RESULTS

The **Appendix Figure** (available at Annals.org) shows results of searches and study selection. No studies evaluating the overall effectiveness of screening for urinary incontinence in women to reduce adverse outcomes or harms of screening met inclusion criteria.

Table 1. Clinical Screening Instruments to Assess Urinary Incontinence

Instrument	Description	Study, Year (Reference)
3 Incontinence Questions	3 questions about urine leakage to identify stress incontinence, urge incontinence, other causes, or mixed incontinence.	Brown et al, 2006 (34)
Actionable Bladder Symptom Screening Tool	8 items using a 4-point Likert scale and a 7-d recall period. Questions focus on frequency, leakage, urgency, and nighttime voiding and the effect on social relations, work interference, and embarrassment. A score ≥3 (range, 0–8) indicates the need for further evaluation and/or treatment.	Cardozo et al, 2014 (35)
Bladder Control Self-Assessment Questionnaire	Self-completed questionnaire with a scale ranging from 0 (not at all) to 3 (a great deal) for 4 symptom questions and 4 corresponding bother questions. Scores are totaled for each set of questions from 0 to 12, with higher scores indicating symptoms of urinary incontinence and a higher degree of bother.	Basra et al, 2012 (30)
Bristol Female Lower Urinary Tract Symptoms questionnaire	19-item questionnaire with 3 main domains: incontinence (5 items related to urge, frequency, stress, unpredictable, and nocturnal incontinence), voiding (3 items related to hesitancy, straining to start, and intermittency), and filling (4 items related to nocturia, urgency, bladder pain, and frequency) with additional subscales for sexual function (2 items related to sex life being spoiled and leakage during intercourse) and quality of life (5 items related to changing outer clothes, decreasing fluid intake, daily tasks, avoidance of situations, and overall quality of life).	Khan et al, 2004 (39)
Detrusor instability score	10 questions about the patient's urogynecologic dysfunction. Each question is scored 0, 1, or 2: 0 indicates stress urinary incontinence, and 1 or 2 indicates slight or marked detrusor instability. Scores range from 0 to 20, with a score from 0 to 7 indicating slight detrusor instability and a score ≥8 indicating marked detrusor instability.	Klovning et al, 1996 (41)
Gaudenz Incontinence Questionnaire	26 questions related to stress incontinence and detrusor instability.	Haeusler et al, 1995 (38)
Incontinence screening questionnaire	Self-administered, 6-item questionnaire designed to distinguish between stress and urge incontinence.	Gunthorpe et al, 2000 (37)
Michigan Incontinence Symptom Index	10 items using a 4-point Likert scale; subdomains include stress incontinence, urge incontinence, and pad use, along with a bother domain. Total scores range from 0 to 32, bother domain scores range from 0 to 8, stress and urge incontinence scores range from 0 to 12, and the pad use score ranges from 0 to 8. Higher scores indicate greater symptoms/bother.	Suskind et al, 2015 (45)
Overactive Bladder Awareness Tool	8 items describing symptoms, each scored on a 6-point Likert scale ranging from 0 (not at all) to 5 (a very great deal). Scores are summed, and patients with a score ≥8 are instructed to speak to their physicians about their urinary symptoms.	Basra et al, 2012 (30)
Questionnaire for Urinary Incontinence Diagnosis	Stress and urge incontinence subscales with 3 items each. For each item, scores range from 0 (none of the time) to 5 (all of the time), with total scores for each subscale ranging from 0 to 15. Stress incontinence is diagnosed with a stress score ≥4 and urge incontinence with an urge score ≥6.	Bradley et al, 2005 (33)
Urogenital Distress Inventory	Short form with 6 questions of urogenital distress rated on a scale of 0 (does not experience symptom) to 4 (bothered by symptom quite a bit). Higher scores indicate higher disability.	Lemack and Zimmern, 1999 (43)

Accuracy of Screening Methods

Seventeen diagnostic accuracy studies of 18 screening methods for urinary incontinence in women met inclusion criteria (**Appendix Table 1**, available at Annals.org) (30–46). Studies ranged in size from 69 to 1911 participants and enrolled women from the community or primary care, gynecology, or urogynecology clinics in the United States (31, 33–36, 40, 43–46), United Kingdom (30, 39), Denmark (32), Austria (38), Norway (41), Finland (42), and Australia (37). Although participant age varied, articles did not provide age-specific results. Race, body mass index, parity, and menopausal status were not uniformly reported. Most studies enrolled participants who had incontinence symptoms, although 5 studies of 6 methods did not and are most relevant to screening (30, 35–37, 45).

Most screening methods were clinician- or self-administered questionnaires about symptoms of urinary incontinence that were designed for use in clinical practice (**Table 1**) (30, 31, 33–35, 37–39, 41–43, 45). Some methods involved only 1 or 2 questions (32, 36, 40, 44, 46). Questions about the presence and severity of incontinence symptoms were similar across instruments. Responses were typically scored using a Likert scale or other point system. Diagnostic cut points were determined by comparing scores against reference standards that differed across studies, including clinical diagnosis based on

Table 2. Results of Studies of Screening Methods

Method	Accuracy Measures for Urinary Incontinence			Quality Rating	Reference
	Stress	Urge	Any		
Participants selected from the community or primary care or nonspecialty clinics					
Michigan Incontinence Symptom Index	Sensitivity: 77% Specificity: 76% PPV: 43% NPV: 86% AUROC: 0.79	Sensitivity: 86% Specificity: 73% PPV: 73% NPV: 92% AUROC: 0.88	Sensitivity: 84% Specificity: 75% PPV: 75% NPV: 84% AUROC: 0.88	Good	45
Bladder Control Self-Assessment Questionnaire	AUROC: 0.85	AUROC: 0.83	AUROC: 0.87	Fair*	30
Overactive Bladder Awareness Tool	AUROC: 0.68	AUROC: 0.82	AUROC: 0.75	Fair*	30
Incontinence screening questionnaire	–	–	Sensitivity: 66% Specificity: 80%†	Fair‡	37
Self-report	Sensitivity: 95.5% Specificity: 44.6%	Sensitivity: 40.9% Specificity: 67.6%	Sensitivity: 95.5% Specificity: 32.4%	Poor§	32
3 Incontinence Questions	Sensitivity: 86% (95% CI, 79%–90%) Specificity: 60% (CI, 51%–68%) PLR: 2.13 (CI, 1.71–2.66) NLR: 0.24 (CI, 0.16–0.35)	Sensitivity: 75% (CI, 68%–81%) Specificity: 77% (CI, 69%–84%) PLR: 3.29 (CI, 2.39–4.51) NLR: 0.32 (CI, 0.24–0.43)	–	Poor§	34
Actionable Bladder Symptom Screening Tool	–	–	Sensitivity: 79.1% Specificity: 98.2% PPV: 97.1% NPV: 86.2% AUROC: 0.96	Poor*‡‖	35
Self-report of any incontinence	Sensitivity: 57.2% Specificity: 84.1%	Sensitivity: 52.0% Specificity: 73.2%	–	Poor‡¶	36
Participants selected from referral clinics					
Gaudenz Incontinence Questionnaire	Sensitivity: 55.9% Specificity: 44.7% PPV: 88.2% NPV: 18.1%	Sensitivity: 61.5% Specificity: 56.1% PPV: 2.8% NPV: 98.5%	–	Poor‡§¶	38
Bristol Female Lower Urinary Tract Symptoms questionnaire	Incontinence Sensitivity: 14% Specificity: 98% Symptoms Sensitivity: 88% Specificity: 29%	Incontinence Sensitivity: 8% Specificity: 84% Symptoms Sensitivity: 81% Specificity: 12%	–	Poor‡§¶**	39
Detrusor instability score	–	Sensitivity: 60% Specificity: 77% PPV: 82% NPV: 52%	–	Poor§	41
Urgency score	Sensitivity: 19% Specificity: 32% PPV: 20% NPV: 31%	Sensitivity: 93% Specificity: 32% PPV: 26% NPV: 98%	Sensitivity: 64% Specificity: 62% PPV: 37% NPV: 84%	Poor‡§	42
Urogenital Distress Inventory	Sensitivity: 84.8% Specificity: 63.4%	Question 1 score ≥2 Sensitivity: 75.0% Specificity: 32.6% Question 2 score ≥2 Sensitivity: 83.3% Specificity: 50.0% Question 1 and 2 score ≥2 Sensitivity: 68.6% Sensitivity: 63.8%	–	Poor‡§‖¶	43
Questionnaire	Urine loss with cough and sneeze Sensitivity: 90% Specificity: 24% PPV: 79% Urine loss with straining Sensitivity: 95% Specificity: 43% PPV: 83%	Uncomfortable before emptying bladder Sensitivity: 90% Specificity: 41% PPV: 37% Have to hurry to toilet Sensitivity: 92% Specificity: 59% PPV: 20%	–	Poor‡§‖	31
Self-report	Sensitivity: 52% Specificity: 88% PPV: 71%	Sensitivity: 52% Specificity: 88% PPV: 71%	Sensitivity: 72% Specificity: 49% PPV: 42%	Poor‡§	40

Continued on following page

Table 2–Continued

Method	Accuracy Measures for Urinary Incontinence			Quality Rating	Reference
	Stress	Urge	Any		
Reported symptoms	Sensitivity: 100% Specificity: 65.2% PPV: 86.9% NPV: 100%	Sensitivity: 77.9% Specificity: 38.7% PPV: 36.6% NPV: 79.5%	–	Poor‡§‖¶††	44
1 question	Sensitivity: 93% Specificity: 19% PPV: 59% NPV: 41%	–	–	Poor‡§	46
Questionnaire for Urinary Incontinence Diagnosis	Sensitivity: 85% (CI, 75%–91%) Specificity: 71% (CI, 51%–87%) PPV: 90% (CI, 81%–96%) NPV: 61% (CI, 42%–77%) AUROC: 0.83 (CI, 0.74–0.92)	Sensitivity: 79% (CI, 69%–86%) Specificity: 79% (CI, 54%–94%) PPV: 95% (CI, 87%–99%) NPV: 43% (CI, 26%–60%) AUROC: 0.83 (CI, 0.75–0.92)	–	Poor§	33

AUROC = area under the receiver-operating characteristic curve; NLR = negative likelihood ratio; NPV = negative predictive value; PLR = positive likelihood ratio; PPV = positive predictive value.
* Reference standard is not credible or replicable or is unclear in publication.
† Calculated from data in publication.
‡ No blinding of outcome assessors to the reference standard or unclear in publication.
§ Patient spectrum is narrow (symptomatic referral population) or unclear in publication.
‖ Population tested is not consecutive or random or is unclear in publication.
¶ Test is not adequately described, not reproducible, or unclear in publication.
** Attrition is not reported, there was large loss to follow-up, or unclear in publication.
†† Eligibility criteria were not clearly described.

physical examinations and tests (30, 32–35, 41, 45), urodynamic testing (31, 36, 38–40, 42–44, 46), and the pad test (37). Results were expressed as AUROCs, sensitivity and specificity estimates, positive and negative predictive values, or likelihood ratios.

Table 2 presents results of the studies. Most were designed to differentiate types of incontinence and reported results specifically for stress and urge incontinence (or overactive bladder) (30–34, 36, 38–40, 42–45). Some studies reported either 1 type or general or mixed incontinence only. Four studies of 5 methods reported AUROCs for stress, urge, and general or mixed incontinence (30, 33, 35, 45). Thirteen studies reported sensitivity and specificity outcomes for stress incontinence, 13 for urge incontinence, and 6 for general or mixed incontinence. Some instruments performed particularly well for 1 type of urinary incontinence but not others.

Of included studies, 1 met criteria for good quality (45), 2 for fair (30, 37), and 14 for poor (31–36, 38–44, 46) (**Appendix Table 2**, available at Annals.org). Most were downgraded for enrolling only symptomatic women (31–34, 38–44, 46), resulting in estimates of test performance that may not apply to a screening population but could be relevant to others. Reference standards often were not credible, replicable, or blinded or had not been applied to all participants or a random subset, or these items were unclear in the report (30, 31, 35–40, 42–44, 46).

Studies of Participants From the Community or Primary Care or Nonspecialty Clinics

Seven studies enrolled participants from the community or primary care or nonspecialty clinics (30, 32, 34–37, 45). Of these, 5 studies that did not recruit participants on the basis of symptoms of incontinence are particularly applicable to population screening (30, 35–37, 45).

A good-quality study of 214 community-dwelling women in Michigan evaluated responses on the Michigan Incontinence Symptom Index (MISI) against a physician's diagnosis based on an extended clinical evaluation provided for all participants (45). Mean age was 50.5 years (range, 35 to 64 years), mean body mass index 33.1 kg/m^2, and mean parity 2.2. Of the participants, 57% were postmenopausal, 32% were white, 68% were black, and 54% self-reported symptoms of incontinence at baseline. The MISI is a 10-item questionnaire that uses a 4-point Likert scale (maximum score, 32 points) in domains specific to stress and urge incontinence. Screening thresholds for MISI include a total score of 7 or more for mixed incontinence, a subdomain score of 3 or more for stress incontinence, and a subdomain score of 5 or more for urge incontinence. All participants had the clinical evaluation, which included a pelvic examination based on the Pelvic Organ Prolapse Quantification system, vaginal examination, Q-tip angle test, measurement of bladder postvoid residual volume, urodynamic test with urethral pressure profile, leak point pressure and urine flow rate test, and paper towel test. Comparing the MISI scores against results of the clinical evaluation yielded AUROCs of 0.79 for stress, 0.88 for urge, and 0.88 for mixed incontinence.

In a fair-quality study of 223 women recruited from general gynecology, urogynecology, and primary care clinics in London, United Kingdom, results of the Bladder Control Self-Assessment Questionnaire (B-SAQ) and Overactive Bladder Awareness Tool (OAB-V8) were compared against a clinical diagnosis of incontinence (30). In this study, 46% of women had symptoms of incontinence at baseline and others reported bothersome symptoms. The B-SAQ is an 8-item questionnaire that evaluates urinary symptoms, incontinence, and bother on a 4-point Likert scale. The OAB-V8 is an 8-item questionnaire that evaluates symptoms of overactive bladder, including urinary frequency, nocturia, urgency, and urge incontinence on a 6-point Likert scale. Both instruments provide results

for stress incontinence, overactive bladder, and mixed incontinence. A clinical diagnosis by a specialist based on an evaluation form was the reference standard for this study, but the report did not provide details. Results indicated AUROCs for the B-SAQ of 0.85 for stress incontinence, 0.83 for overactive bladder, and 0.87 for mixed incontinence. For the OAB-V8, values were 0.68, 0.82, and 0.75, respectively.

A fair-quality study of 89 women recruited from general practice clinics in Australia tested an incontinence screening questionnaire against the 48-hour pad test and self-reported incontinence at the time of pad testing (37). This self-administered questionnaire includes 6 items related to symptoms of stress and urge incontinence, whereas the pad test is an objective measure of urinary leakage that compares weights of pads before and after use. Participants were generally younger (mean age, 42 years) and less obese (mean body mass index, 24 kg/m^2) than those in other studies. Whether assessors were blinded to the outcome was unclear. The test had 65.5% sensitivity and 80.0% specificity for identifying any type of urinary incontinence.

In a study of the 6-item Actionable Bladder Symptom Screening Tool in general gynecology clinics in the United States, the reference standard was a clinician assessment of whether the woman should be referred to specialists (35). Although referral is a pragmatic outcome relevant to screening, determining whether it is a credible and replicable reference standard is difficult, which contributed to the study's poor quality rating. Sensitivity, specificity, and AUROC of the test for identifying any type of incontinence were 79%, 98%, and 0.96, respectively.

A study that evaluated the accuracy of self-reported incontinence was limited by the nonreproducibility of this method (36). The other 2 studies enrolled only symptomatic women, resulting in poor quality ratings (32, 34). These studies evaluated the accuracy of using 1 to 3 screening questions, which would be highly feasible in primary care settings.

Studies of Participants From Referral Clinics

Ten studies were based in referral clinics and enrolled women with incontinence symptoms. These studies were generally designed to determine the accuracy of patient reports before urogynecologic evaluations by specialists (31, 33, 38–44, 46). Studies differentiated stress from urge incontinence or targeted 1 specific type. Methods included the Gaudenz Incontinence Questionnaire (38), Bristol Female Lower Urinary Tract Symptoms questionnaire (39), detrusor instability score (41), Urogenital Distress Inventory (43), and Questionnaire for Urinary Incontinence Diagnosis (33), as well as single questions (46), self-reports (40, 44), and unnamed questionnaires (31, 42). These studies all met criteria for poor quality because of narrow spectra of patients, although reference standards involving urodynamic evaluations by specialists were more consistent across studies. Evaluations generally included urogynecologic physical examinations, postvoid residual volume measurements, stress tests, pressure flow studies, cystometry, and cystourethrography. Accuracy measures for the instruments varied widely.

DISCUSSION

Key findings of this systematic review of screening for urinary incontinence indicate that no studies have evaluated the overall effectiveness or harms of screening women in the general population. Seventeen studies evaluated the diagnostic accuracy of 18 potential screening methods compared with a clinical diagnosis of incontinence or results of diagnostic tests. Screening methods were similar across studies and included predominantly brief questionnaires that were administered by the clinician or patient and described symptoms that were easily scored and interpreted. However, few studies had findings applicable to population screening because they evaluated women from the community or primary care or nonspecialty clinics who were not recruited on the basis of symptoms of incontinence. One good-quality and 2 fair-quality studies enrolled such women (MISI: AUROC for stress, 0.79; for urge, 0.88; for any type, 0.88; B-SAQ: AUROC for stress, 0.85; for urge, 0.83; for any type, 0.87; OAB-V8: AUROC for stress, 0.68; for urge, 0.82; for any type, 0.75; Incontinence Screening Questionnaire: sensitivity, 66%; specificity, 80%). Two poor-quality studies evaluated the Actionable Bladder Symptom Screening Tool for any type of incontinence (AUROC, 0.96) and self-reported incontinence on a household survey (sensitivity, 52% to 72%; specificity, 73% to 84%).

We found no other systematic reviews of screening for urinary incontinence or methods of screening in our extensive literature searches, and only 1 recently launched screening trial was listed on ClinicalTrials.gov. Previously published reviews concern diagnostic testing and treatment (13, 15, 17, 24–26, 47). Our systematic review is limited by its restriction to English-language articles relevant to U.S. practice settings, although this focus increases its applicability to guidelines for women in the United States. Studies to determine the overall effectiveness and harms of screening are central in supporting population screening guidelines—and are lacking in this topic. Although several studies of the diagnostic accuracy of clinical screening methods have been published, they represent single studies in patient samples that are generally small and symptomatic. These limitations compromise estimates of accuracy and their applicability to population screening. Inclusion criteria and reference standards vary across studies or are not well described, and CIs around estimates of performance measures were infrequently reported. No studies reported results by age or other patient characteristic; whether certain methods perform better among specific subgroups of women is unclear.

The lack of effectiveness trials of screening for urinary incontinence is an important deficiency in women's health research, considering the high prevalence, health burden, and stigma of this condition. Results of the URINO (Urinary Incontinence in Older Women) trial provide insights into the potential benefits of screening. In this cluster randomized trial in the Netherlands, 31% of women aged 55 years or older in family medicine practices self-reported incontinence symptoms on 2 questions in a postal questionnaire (48). Symptomatic women (n = 350) were then randomly assigned to active encouragement of diagnostic evaluations and treatment or a usual care group. Results indicated that all women in the intervention group

had diagnostic testing and 80% accepted treatment advice, whereas only 2% received treatment in the control group. At 12 months of follow-up, more women in the intervention group had improvement in symptom severity (odds ratio, 1.9 [95% CI, 1.1 to 3.3]) and fewer episodes of incontinence (odds ratio, 2.5 [CI, 1.5 to 4.1]) than in the control group (49). Although this trial does not evaluate the effectiveness of screening in itself, it is unique in showing the effectiveness of actively engaging women with previously unidentified incontinence in diagnostic evaluations and treatment.

Future research is needed to address crucial evidence gaps and improve limitations of existing research. Studies should evaluate the overall effectiveness and harms of screening for urinary incontinence in primary care settings so that women can be identified early enough to avoid more severe symptoms and their repercussions and benefit from less invasive treatment. Research on the feasibility, accuracy, and effectiveness of screening in larger, more diverse populations would help establish a standardized screening method that could be widely implemented in routine practice. Validation of existing methods, particularly those with higher accuracy in current studies (such as MISI, B-SAQ, and OAB-V8), is needed to demonstrate their performance in screening populations. Additional research to understand changes in the incidence and prevalence of urinary incontinence over time and the influence of specific risk factors, including racial and ethnic differences, could focus screening on optimal times, intervals, and population subgroups. Questions also remain regarding other steps in the clinical pathway of urinary incontinence, including the most effective approach to diagnostic evaluations after identification of urinary incontinence symptoms. The value of urodynamic studies in the management of urinary incontinence remains an important question (17, 47, 50).

In conclusion, evidence is insufficient on the overall effectiveness and harms of screening for urinary incontinence in women. Limited evidence about the accuracy of screening instruments in general populations is based on studies of brief screening questionnaires that suggest fairly high accuracy. Decisions about screening for urinary incontinence need to consider the level of existing supporting evidence in the context of the high prevalence and burden of this condition in women.

From Oregon Health & Science University, Portland, Oregon (H.D.N., A.C., M.P., L.M.).

Disclaimer: Advisory Panel or Multidisciplinary Steering Committee participation in the Women's Preventive Services Initiative or authorship of this document does not constitute organizational or individual endorsement of the recommendations or conclusions. The findings and conclusions in this document are those of the authors, who are responsible for its content, and do not necessarily represent the views of the Health Resources and Services Administration. No statement in this report should be construed as an official position of the U.S. Department of Health and Human Services.

Grant Support: By cooperative agreement UH0MC29440 from the Health Resources and Services Administration of the U.S. Department of Health and Human Services.

Disclosures: Authors have disclosed no conflicts of interest. Forms can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M18-0225.

Reproducible Research Statement: *Study protocol:* See **Supplement** (available at Annals.org). *Statistical code:* Not applicable. *Data set:* See **Appendix Tables** and **Figure** (available at Annals.org); additional information is available from the investigators (e-mail, nelsonh@ohsu.edu).

Corresponding Author: Heidi D. Nelson, MD, MPH, Oregon Health and Science University, 3181 SW Sam Jackson Park Road, Mailcode BICC, Portland, OR 97239; e-mail, nelsonh@ohsu.edu.

Current author addresses and author contributions are available at Annals.org.

References

1. **Rosenberg MT, Dmochowski RR.** Overactive bladder: evaluation and management in primary care. Cleve Clin J Med. 2005;72:149-56. [PMID: 15757170]
2. **Morrison A, Levy R.** Fraction of nursing home admissions attributable to urinary incontinence. Value Health. 2006;9:272-4. [PMID: 16903997]
3. **Abrams P, Andersson KE, Birder L, et al; Members of Committees.** 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. [PMID: 20025020] doi: 10.1002/nau.20870
4. **Haylen BT, de Ridder D, Freeman RM, et al; International Urogynecological Association.** An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. Neurourol Urodyn. 2010;29:4-20. [PMID: 19941278] doi:10.1002/nau.20798
5. **Carls C.** The prevalence of stress urinary incontinence in high school and college-age female athletes in the Midwest: implications for education and prevention. Urol Nurs. 2007;27:21-4. [PMID: 17390923]
6. **Kinchen KS, Lee J, Fireman B, Hunkeler E, Nehemiah JL, Curtice TG.** The prevalence, burden, and treatment of urinary incontinence among women in a managed care plan. J Womens Health (Larchmt). 2007;16:415-22. [PMID: 17439386]
7. **Boyington JE, Howard DL, Carter-Edwards L, et al.** Differences in resident characteristics and prevalence of urinary incontinence in nursing homes in the southeastern United States. Nurs Res. 2007;56:97-107. [PMID: 17356440]
8. **Shamliyan T, Wyman J, Bliss DZ, Kane RL, Wilt TJ.** Prevention of urinary and fecal incontinence in adults. Evid Rep Technol Assess (Full Rep). 2007:1-379. [PMID: 18457475]
9. **Nygaard I, Barber MD, Burgio KL, et al; Pelvic Floor Disorders Network.** Prevalence of symptomatic pelvic floor disorders in US women.JAMA.2008;300:1311-6.[PMID:18799443]doi:10.1001/jama.300.11.1311
10. **Al-Mukhtar Othman J, Åkervall S, Milsom I, Gyhagen M.** Urinary incontinence in nulliparous women aged 25-64 years: a national survey. Am J Obstet Gynecol. 2017;216:149. [PMID: 27720862] doi:10.1016/j.ajog.2016.09.104
11. **Lukacz ES, Lawrence JM, Contreras R, Nager CW, Luber KM.** Parity, mode of delivery, and pelvic floor disorders. Obstet Gynecol. 2006;107:1253-60. [PMID: 16738149]
12. **Lawrence JM, Lukacz ES, Nager CW, Hsu JW, Luber KM.** Prevalence and co-occurrence of pelvic floor disorders in community-dwelling women. Obstet Gynecol. 2008;111:678-85. [PMID: 18310371] doi:10.1097/AOG.0b013e3181660c1b

13. **Abrams P, Kelleher CJ, Kerr LA, Rogers RG.** Overactive bladder significantly affects quality of life. Am J Manag Care. 2000;6:S580-90. [PMID: 11183901]
14. **MacKay K, Hemmett L.** Needs assessment of women with urinary incontinence in a district health authority. Br J Gen Pract. 2001;51: 801-4. [PMID: 11677702]
15. **Davila GW, Guerette N.** Current treatment options for female urinary incontinence–a review. Int J Fertil Womens Med. 2004;49: 102-12. [PMID: 15303311]
16. **Melville JL, Newton K, Fan MY, Katon W.** Health care discussions and treatment for urinary incontinence in U.S. women. Am J Obstet Gynecol. 2006;194:729-37. [PMID: 16522405]
17. **Clement KD, Lapitan MC, Omar MI, Glazener CM.** Urodynamic studies for management of urinary incontinence in children and adults. Cochrane Database Syst Rev. 2013:CD003195. [PMID: 24166676] doi:10.1002/14651858.CD003195.pub3
18. **Appell RA, Dmochowski RR, Blaivas JM, et al.** Guidelines for the Surgical Management of Female Stress Urinary Incontinence: Update (2009). Linthicum, MD: American Urological Assoc; 2009.
19. Urinary incontinence in women. Female Pelvic Med Reconstr Surg. 2015;21:304-14. [PMID: 26506157] doi:10.1097/SPV.0000000000000231
20. **Lucas MG, Bedretdinova D, Bosch JLHR, et al.** Guidelines on Urinary Incontinence. Arnhem, the Netherlands: European Assoc Urology; 2014.
21. **Bettez M, Tu le M, Carlson K, et al.** 2012 update: guidelines for adult urinary incontinence collaborative consensus document for the Canadian Urological Association. Can Urol Assoc J. 2012;6:354-63. [PMID: 23093627] doi:10.5489/cuaj.12248
22. **Women's Preventive Services Initiative.** Screening for urinary incontinence. 2017. Accessed at www.womenspreventivehealth.org/recommendations/screening-for-urinary-incontinence on 12 June 2018.
23. **Agency for Healthcare Research and Quality.** Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ publication no. 10(14)-EHC063-EF. Rockville: Agency for Healthcare Research and Quality; January 2014.
24. **Shamliyan T, Wyman J, Kane RL.** Nonsurgical Treatments for Urinary Incontinence in Adult Women: Diagnosis and Comparative Effectiveness. Comparative Effectiveness Review no. 36. (Prepared by the University of Minnesota Evidence-based Practice Center under contract no. HHSA 290-2007-10064-I.) Rockville: Agency for Healthcare Research and Quality; 2012.
25. **McDonagh MS, Selover D, Santa J, Lopez S.** Drug Class Reviews. Drug Class Review on Agents for Overactive Bladder: Final Report. Portland: Oregon Health & Science Univ; 2005.
26. **Lukacz ES, Santiago-Lastra Y, Albo ME, Brubaker L.** Urinary incontinence in women: a review. JAMA. 2017;318:1592-604. [PMID: 29067433] doi:10.1001/jama.2017.12137
27. **U.S. Preventive Services Task Force.** Methods and Processes. Accessed at www.uspreventiveservicestaskforce.org/Page/Name/methods-and-processes on 18 July 2018.
28. **Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group.** QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155:529-36. [PMID: 22007046] doi:10.7326/0003-4819-155-8-201110180-00009
29. **Chang SM, Matchar DB, Smetana GW, Umscheid CA, eds.** Methods Guide for Medical Test Reviews. AHRQ publication no. 12-EHC017. Rockville: Agency for Healthcare Research and Quality; June 2012.
30. **Basra RK, Cortes E, Khullar V, Kelleher C.** A comparison study of two lower urinary tract symptoms screening tools in clinical practice: the B-SAQ and OAB-V8 questionnaires. J Obstet Gynaecol. 2012;32: 666-71. [PMID: 22943714] doi:10.3109/01443615.2012.696158
31. **Bergman A, Bader K.** Reliability of the patient's history in the diagnosis of urinary incontinence. Int J Gynaecol Obstet. 1990;32: 255-9. [PMID: 1972118]
32. **Borup K, Hvidman L, Nielsen JB, Mommsen S.** Validity of a self-administered questionnaire, with reference to a clinical stress urinary incontinence test. Scand J Urol Nephrol. 2008;42:148-53. [PMID: 17853006]
33. **Bradley CS, Rovner ES, Morgan MA, et al.** A new questionnaire for urinary incontinence diagnosis in women: development and testing. Am J Obstet Gynecol. 2005;192:66-73. [PMID: 15672005]
34. **Brown JS, Bradley CS, Subak LL, et al; Diagnostic Aspects of Incontinence Study (DAISy) Research Group.** The sensitivity and specificity of a simple test to distinguish between urge and stress urinary incontinence. Ann Intern Med. 2006;144:715-23. [PMID: 16702587]
35. **Cardozo L, Staskin D, Currie B, et al.** Validation of a bladder symptom screening tool in women with incontinence due to overactive bladder. Int Urogynecol J. 2014;25:1655-63. [PMID: 24859795] doi:10.1007/s00192-014-2417-7
36. **Diokno AC, Normolle DP, Brown MB, Herzog AR.** Urodynamic tests for female geriatric urinary incontinence. Urology. 1990;36: 431-9. [PMID: 2238302]
37. **Gunthorpe W, Brown W, Redman S.** The development and evaluation of an incontinence screening questionnaire for female primary care. Neurourol Urodyn. 2000;19:595-607. [PMID: 11002302]
38. **Haeusler G, Hanzal E, Joura E, Sam C, Koelbl H.** Differential diagnosis of detrusor instability and stress-incontinence by patient history: the Gaudenz-Incontinence-Questionnaire revisited. Acta Obstet Gynecol Scand. 1995;74:635-7. [PMID: 7660771]
39. **Khan MS, Chaliha C, Leskova L, Khullar V.** The relationship between urinary symptom questionnaires and urodynamic diagnoses: an analysis of two methods of questionnaire administration. BJOG. 2004;111:468-74. [PMID: 15104612]
40. **Klingele CJ, Carley ME, Hill RF.** Patient characteristics that are associated with urodynamically diagnosed detrusor instability and genuine stress incontinence. Am J Obstet Gynecol. 2002;186:866-8. [PMID: 12015497]
41. **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]
42. **Kujansuu E, Kauppila A.** Scored urological history and urethrocystometry in the differential diagnosis of female urinary incontinence. Ann Chir Gynaecol. 1982;71:197-202. [PMID: 6889831]
43. **Lemack GE, Zimmern PE.** Predictability of urodynamic findings based on the Urogenital Distress Inventory-6 questionnaire. Urology. 1999;54:461-6. [PMID: 10475355]
44. **Sand PK, Hill RC, Ostergard DR.** Incontinence history as a predictor of detrusor stability. Obstet Gynecol. 1988;71:257-60. [PMID: 3336562]
45. **Suskind AM, Dunn RL, Morgan DM, DeLancey JO, Rew KT, Wei JT.** A screening tool for clinically relevant urinary incontinence. Neurourol Urodyn. 2015;34:332-5. [PMID: 24464849] doi:10.1002/nau.22564
46. **Walters MD, Shields LE.** The diagnostic value of history, physical examination, and the Q-tip cotton swab test in women with urinary incontinence. Am J Obstet Gynecol. 1988;159:145-9. [PMID: 3394734]
47. **Clement KD, Lapitan MC, Omar MI, Glazener CM.** Urodynamic studies for management of urinary incontinence in children and adults: a short version Cochrane systematic review and meta-analysis. Neurourol Urodyn. 2015;34:407-12. [PMID: 24853652] doi:10.1002/nau.22584
48. **Visser E, de Bock GH, Messelink EJ, et al.** Active encouragement of older women with urinary incontinence in primary care to undergo diagnosis and treatment: a matched-pair cluster randomized controlled trial. Maturitas. 2015;80:212-9. [PMID: 25534171] doi:10.1016/j.maturitas.2014.11.015
49. **Visser E, Dekker JH, Vermeulen KM, et al.** The effect of systematic screening of older women for urinary incontinence on treatment uptake: the URINO trial. Maturitas. 2013;74:334-40. [PMID: 23384977] doi:10.1016/j.maturitas.2012.12.015
50. **Nager CW, Brubaker L, Litman HJ, et al; Urinary Incontinence Treatment Network.** A randomized trial of urodynamic testing before stress-incontinence surgery. N Engl J Med. 2012;366:1987-97. [PMID: 22551104] doi:10.1056/NEJMoa1113595

Current Author Addresses: Drs. Nelson and Cantor, Ms. Pappas, and Liev Miller: Oregon Health and Science University, 3181 SW Sam Jackson Park Road, Mailcode BICC, Portland, OR 97239.

Author Contributions: Conception and design: H.D. Nelson, A. Cantor.
Analysis and interpretation of the data: H.D. Nelson, A. Cantor, M. Pappas.
Drafting of the article: H.D. Nelson, A. Cantor, M. Pappas, L. Miller.
Critical revision of the article for important intellectual content: H.D. Nelson, A. Cantor, M. Pappas.
Final approval of the article: H.D. Nelson, A. Cantor, M. Pappas, L. Miller.
Provision of study materials or patients: H.D. Nelson.
Statistical expertise: H.D. Nelson.
Obtaining of funding: H.D. Nelson, A. Cantor.
Administrative, technical, or logistic support: H.D. Nelson, A. Cantor, M. Pappas, L. Miller.
Collection and assembly of data: H.D. Nelson, A. Cantor, M. Pappas, L. Miller.

Appendix: Search Strategies

Database: Ovid MEDLINE Without Revisions

1 exp Urinary Incontinence/ (18961)
2 exp Urinary Bladder, Overactive/ (3285)
3 1 or 2 (21389)
4 exp Mass Screening/ (79980)
5 exp Women's Health/ (22896)
6 Female/ (4632475)
7 exp Women's Health Services/ (4429)
8 5 or 6 or 7 (4634117)
9 3 and 4 and 8 (59)
10 ((urin* or stress* or urge*) adj5 (incontin* or leak* or ((unabl* or inabilit*) adj3 (hold* or control* or contain* or retain*)))).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (25594)
11 4 and 8 and 10 (69)
12 9 or 11 (75)
13 ((screen* or undiagnos* or undetect* or hide or hiding or hidden or occult or asymptomatic* or unrecogniz* or unacknowledg*) adj10 ((urin* or stress* or urge*) adj5 (incontin* or leak* or ((unabl* or inabilit* or incapab* or cannot or ("not" adj able) or struggl*) adj3 (hold* or control* or contain* or retain*))))).mp. (259)
14 (((well adj wom#n) or ((routin* or annual* or yearly or regular) adj5 (visit* or appointment* or consult* or physical or physicals or exam or exams or examination* or checkup or check-up))) adj10 ((urin* or stress* or urge*) adj5 (incontin* or leak* or ((unabl* or inabilit* or incapab* or cannot or ("not" adj able) or struggl*) adj3 (hold* or control* or contain* or retain*))))).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (8)
15 8 and 13 (232)
16 8 and 14 (7)
17 15 or 16 (239)
18 12 or 17 (302)
19 ((screen* or undiagnos* or undetect* or hide or hiding or hidden or occult or asymptomatic* or unrecogniz* or unacknowledg*) adj10 (overactiv* adj5 bladder*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (36)
20 (((well adj wom#n) or ((routin* or annual* or yearly or regular) adj5 (visit* or appointment* or consult* or physical or physicals or exam or exams or examination* or checkup or check-up))) adj10 (overactiv* adj5 bladder*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (0)
21 8 and 19 (31)
22 8 and 20 (0)
23 21 or 22 (31)
24 18 or 23 (323)

Database: Evidence-Based Medicine Reviews (Cochrane Central Register of Controlled Trials)

1 ((screen* or undiagnos* or undetect* or hide or hiding or hidden or occult or asymptomatic* or unrecogniz* or unacknowledg*) adj10 ((urin* or stress* or urge*) adj5 (incontin* or leak* or ((unabl* or inabilit* or incapab* or cannot or ("not" adj able) or struggl*) adj3 (hold* or control* or contain* or retain*))))).mp. (75)
2 (((well adj wom#n) or ((routin* or annual* or yearly or regular) adj5 (visit* or appointment* or consult* or physical or physicals or exam or exams or examination* or checkup or check-up))) adj10 ((urin* or stress* or urge*) adj5 (incontin* or leak* or ((unabl* or inabilit* or incapab* or cannot or ("not" adj able) or struggl*) adj3 (hold* or control* or contain* or retain*))))).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (3)
3 ((screen* or undiagnos* or undetect* or hide or hiding or hidden or occult or asymptomatic* or unrecogniz* or unacknowledg*) adj10 (overactiv* adj5 bladder*)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (5)
4 (((well adj wom#n) or ((routin* or annual* or yearly or regular) adj5 (visit* or appointment* or consult* or physical or physicals or exam or exams or examination* or checkup or check-up))) adj10 (overactiv* adj5 bladder*)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (1)

5 ((urin* or stress* or urge*) adj5 (incontin* or leak* or ((unabl* or inabilit*) adj3 (hold* or control* or contain* or retain*))))).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (4660)

6 (overactiv* adj5 bladder*).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (1564)

7 5 or 6 (5520)

8 screen*.mp. (26747)

9 7 and 8 (116)

10 ((routin* or annual* or yearly or regular) adj5 (visit* or appointment* or consult* or physical or physicals or exam or exams or examination* or checkup or check-up)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (3340)

11 7 and 10 (17)

12 1 or 2 or 3 or 4 or 9 or 11 (176)

13 (woman* or women* or female*).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword] (503769)

14 12 and 13 (140)

15 limit 14 to English language (121)

16 limit 14 to abstracts (131)

17 15 or 16 (132)

Database: Evidence-Based Medicine Reviews (Cochrane Database of Systematic Reviews)

1 ((screen* or undiagnos* or undetect* or hide or hiding or hidden or occult or asymptomatic* or unrecogniz* or unacknowledg*) adj25 ((urin* or stress* or urge*) adj5 (incontin* or leak* or ((unabl* or inabilit* or incapab* or cannot or ("not" adj able) or struggl*) adj3 (hold* or control* or contain* or retain*))))).mp. (11)

2 (((well adj wom#n) or ((routin* or annual* or yearly or regular) adj5 (visit* or appointment* or consult* or physical or physicals or exam or exams or examination* or checkup or check-up))) adj25 ((urin* or stress* or urge*) adj5 (incontin* or leak* or ((unabl* or inabilit* or incapab* or cannot or ("not" adj able) or struggl*) adj3 (hold* or control* or contain* or retain*))))).mp. [mp=title, abstract, full text, keywords, caption text] (0)

3 ((screen* or undiagnos* or undetect* or hide or hiding or hidden or occult or asymptomatic* or unrecogniz* or unacknowledg*) adj25 (overactiv* adj5 bladder*)).mp. [mp=title, abstract, full text, keywords, caption text] (1)

4 (((well adj wom#n) or ((routin* or annual* or yearly or regular) adj5 (visit* or appointment* or consult* or physical or physicals or exam or exams or examination* or checkup or check-up))) adj25 (overactiv* adj5 bladder*)).mp. [mp=title, abstract, full text, keywords, caption text] (0)

5 1 or 2 or 3 or 4 (12)

6 ((urin* or stress* or urge*) adj5 (incontin* or leak* or ((unabl* or inabilit*) adj3 (hold* or control* or contain* or retain*))))).mp. [mp=title, abstract, full text, keywords, caption text] (274)

7 (overactiv* adj5 bladder*).mp. [mp=title, abstract, full text, keywords, caption text] (53)

8 6 or 7 (282)

9 screen*.mp. (6017)

10 8 and 9 (182)

11 ((routin* or annual* or yearly or regular) adj5 (visit* or appointment* or consult* or physical or physicals or exam or exams or examination* or checkup or check-up)).mp. [mp=title, abstract, full text, keywords, caption text] (497)

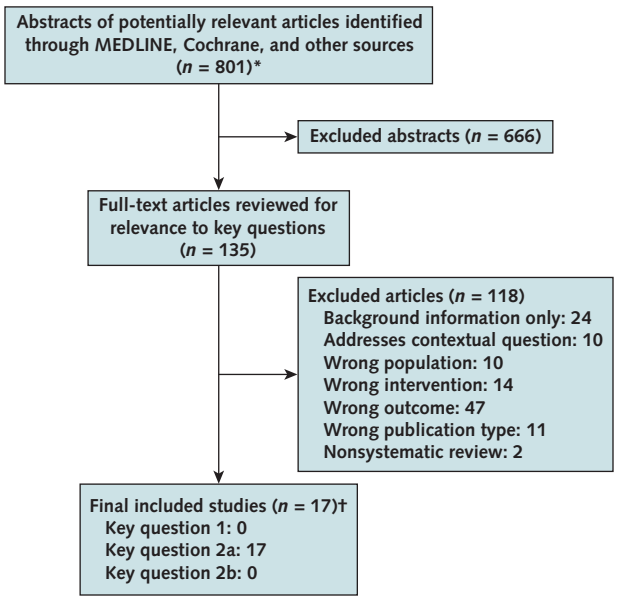
12 8 and 11 (20)

13 (woman* or women* or female*).mp. [mp=title, abstract, full text, keywords, caption text] (4972)

14 5 or 10 or 12 (187)

15 13 and 14 (145)

Appendix Figure. Evidence search and selection.



* Cochrane databases include Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews. Other sources include reference lists and hand-searching.
† Studies that provided data and contributed to the body of evidence were included.

Appendix Table 1. Evidence Table of Studies of the Accuracy of Screening Instruments

Study, Year (Reference)	Design	Participants, *n*	Age, *y*	Characteristics	Setting
Basra et al, 2012 (30)	Prospective	223	Mean 49	75% white, 17% black, 8% Asian	Recruited from general gynecology, urogynecology, and primary care clinics in London, United Kingdom
Bergman and Bader, 1990 (31)	Prospective	154	Mean 54	Mean parity 3; 47% postmenopausal	Urogynecology clinic at University of Southern California Medical Center
Borup et al, 2008 (32)	Prospective	96	20–59	Not reported	7410 randomly selected women from communities of Aarhus and Randers, Denmark; of women indicating episodes of incontinence, a subsample of 96 answered additional questions and had further evaluations
Bradley et al, 2005 (33)	Prospective	117	Median 56 (range 22–87)	73% white, 21% black, 2% Asian, 3% Hispanic; median parity 2 (range 0–8); 52% postmenopausal; median BMI 26.6 kg/m^2	Urogynecology clinic at the University of Pennsylvania
Brown et al, 2006 (34)	Prospective	301	Mean 56	69% white; 13% black; 12% Latina; 2% Asian/Pacific Islander; 4% other; parity 0 (6%); 1–2 (52%); 3–4 (32%); >4 (9.0%); 33% postmenopausal; 34% hysterectomy	Community-dwelling women recruited from 5 U.S. nonspecialty clinical sites with baseline symptoms of urinary incontinence
Cardozo et al, 2014 (35)	Prospective	100	Mean 47.9	71% white, 19% black; mean BMI 28.9 kg/m^2	6 U.S. general gynecology clinics
Diokno et al, 1990 (36)	Prospective	167	≥60 (60–86)	95% white	Of 1108 women responding to household surveys in Michigan, 167 participated in evaluations and urodynamic testing
Gunthorpe et al, 2000 (37)	Prospective	89	Mean 42.4	Mean BMI 24 kg/m^2	Recruited from general practice clinic in Australia
Haeusler et al, 1995 (38)	Prospective	1911	Mean 52.4	Mean parity 2.4; 66% postmenopausal	Referred for urodynamic evaluation in Vienna, Austria
Khan et al, 2004 (39)	Randomized crossover	69	Not reported	Not reported	Women referred to a tertiary urogynecology clinic with lower urinary tract symptoms in London, United Kingdom
Klingele et al, 2002 (40)	Retrospective	278*	Mean 53.7†	61% white, 32% black, 2% Hispanic, 5% other; mean parity 3.0†; 39% hysterectomy	Referred to urogynecology clinic for evaluation of urinary incontinence in the United States
Klovning et al, 1996 (41)	Prospective	250	Mean 49.2	Not reported	Referred for urogenital dysfunction including urinary incontinence in Norway
Kujansuu and Kauppila, 1982 (42)	Prospective	121	Mean 51.6	Not reported	Referred for urinary incontinence to university hospital in Finland
Lemack and Zimmern, 1999 (43)	Prospective	128	Mean 61	Not reported	Referred with lower urinary tract symptoms or incontinence in Texas
Sand et al, 1988 (44)	Prospective	218	Mean 51.8	Mean parity 2.5	Referred for evaluation for incontinence symptoms in California
Suskind et al, 2015 (45)	Prospective	214	Mean 50.5	32% white, 68% black; mean parity 2.2; 57% postmenopausal; mean BMI 33.1 kg/m^2	Community-dwelling women from southeastern Michigan included women with and without symptoms
Walters and Shields, 1988 (46)	Prospective	106	Mean 46.3	Mean weight 165.3 lb; mean parity 4.4	Referral clinic in Texas

ABSST = Actionable Bladder Symptom Screening Tool; AUROC = area under the receiver-operating characteristic curve; BMI = body mass index; B-SAQ = Bladder Control Self-Assessment Questionnaire; MISI = Michigan Incontinence Symptom Index; NLR = negative likelihood ratio; NPV = negative predictive value; OAB-V8 = 8-item Overactive Bladder Awareness Tool; PLR = positive likelihood ratio; PPV = positive predictive value; PVR = postvoid residual; QUID = Questionnaire for Urinary Incontinence Diagnosis.
* Number of participants based on final analysis.
† Calculated.

Appendix Table 1–Continued

Baseline Symptoms	Screening Test and Threshold	Reference Standard	Accuracy Measures
Not recruited on the basis of symptoms; 46% lower urinary tract symptoms; of these 60% had bothersome symptoms	B-SAQ: aggregate symptom score ≥3 or bother score ≥1 OAB-V8: total score ≥8	Clinical diagnosis using an evaluation form by clinicians in the United Kingdom	B-SAQ; OAB-V8 (AUROC): overactive bladder (0.83; 0.82); mixed incontinence (0.87; 0.75); stress incontinence (0.85; 0.68).
122 with urinary complaints; 32 without	64-item questionnaire with 12 items for stress incontinence and 24 items for detrusor instability: positive test result based on response to a single selected question, or to mean of all questions	Clinical and urodynamic evaluations by specialists (urine culture, Q-tip test, pessary test, dynamic water urethrocystometry, urethral pressure profiles at rest and during stress)	Stress incontinence (sensitivity; specificity; PPV of top 2 single questions and mean of 12 items): urine loss with cough, sneeze (90%; 24%; 79%); urine loss with straining (95%; 43%; 83%); mean (56%; 70%; 77%). Detrusor instability (sensitivity; specificity; PPV of top 2 single questions and mean of 24 items): uncomfortable before emptying bladder (90%; 41%; 37%); have to hurry to toilet (92%; 59%; 20%); mean (38%; 80%; 25%).
74% with any urinary incontinence; of these, 65% with stress and 34% urge	Self-reported symptoms: any urinary incontinence during previous 6 mo; stress incontinence while coughing, sneezing, laughing, lifting, or straining; urge incontinence when strong desire to void was associated with involuntary loss of urine	Gynecologic examination and urinary incontinence stress test by specialists	Any type of urinary incontinence: sensitivity 95.5%; specificity 32.4%. Stress: sensitivity 95.5%; specificity 44.6%. Urge: sensitivity 40.9%; specificity 67.6%.
All had symptoms; duration: 15% <1 y; 50% 1–5 y; 30% >5 y	QUID: stress incontinence with score ≥4; urge with score ≥6	Clinical diagnosis by specialists (physical and pelvic examinations; cough stress test, PVR volume, urinalysis)	Stress incontinence: sensitivity 85% (95% CI 75% to 91%); specificity 71% (95% CI 51% to 87%); PPV 90% (95% CI 81% to 96%); NPV 61% (95% CI 42% to 77%); AUROC 0.83 (95% CI 0.74 to 0.92). Urge incontinence: sensitivity 79% (95% CI 69% to 86%); specificity 79% (95% CI 54% to 94%); PPV 95% (95% CI 87% to 99%); NPV 43% (95% CI 26% to 60%); AUROC 0.83 (95% CI 0.75 to 0.92).
≥3 episodes/wk for ≥3 mo; 7 y mean duration of incontinence; 30 mean total episodes/wk	3 Incontinence Questions: response to third question	Clinical diagnosis by specialists (history, physical and pelvic examination, cough stress test, PVR, 3-d voiding diary)	Stress incontinence: sensitivity 86% (95% CI 79% to 90%); specificity 60% (95% CI 51% to 68%); PLR: 2.13 (95% CI 1.71 to 2.66); NLR: 0.24 (95% CI 0.16 to 0.35). Urge incontinence: sensitivity 75% (95% CI 68% to 81%); specificity 77% (95% CI 69% to 84%); PLR 3.29 (95% CI 2.39 to 4.51); NLR 0.32 (95% CI 0.24 to 0.43).
53% with symptoms of urgency or overactive bladder	ABSST: total score ≥3	Clinician assessment of whether women should be referred to specialists	Incontinence: sensitivity 79.1%; specificity 98.2%; PPV 97.1%; NPV 86.2%; AUROC 0.958.
Inclusion not based on symptoms; 100 with symptoms, 67 without	Self-report of any incontinence	Urodynamic tests by specialists (uroflowmetry, PVR, cystometry, stress test)	Stress incontinence: sensitivity 57.2%; specificity 84.1%. Detrusor instability: sensitivity 52.0%; specificity 73.2%.
Not reported at baseline; after evaluation, 29 diagnosed with incontinence	Incontinence Screening Questionnaire: score ≥3	48-h pad test and self-reported incontinence at the time of pad-testing	Incontinence: sensitivity 65.5%; specificity: 80.0%; PPV 61.3%; NPV 82.8%; PLR 3.3; NLR 0.4.
Not reported	Gaudenz Incontinence Questionnaire: scores classify patients with types of incontinence (stress, urge, genuine stress, detrusor instability)	Urodynamic evaluation by specialists (PVR, filling cystometry, urethral pressure, clinical stress test, urethrocystoscopy for those with urge symptoms)	Stress incontinence: sensitivity 55.9%; specificity: 44.7%; PPV 88.2%; NPV 18.1%. Detrusor instability: sensitivity 61.5%; specificity 56.1%; PPV 2.8%; NPV 98.5%.
Not reported	Bristol Female Lower Urinary Tract Symptoms Questionnaire: several combinations of responses evaluated	Urodynamic evaluation by specialists (uroflowmetry, video cystourethrography, pressure flow studies)	Stress incontinence: sensitivity 14%; specificity 98%; any symptoms of stress incontinence: sensitivity 88%; specificity 29%. Detrusor instability incontinence; sensitivity 8%; specificity 84%; any symptoms of detrusor instability: sensitivity 81%; specificity 12%.
21% with urge urinary incontinence; 26% with stress; 53% mixed	Self-reported symptoms: stress incontinence occurs with physical activity; urge associated with a strong desire to urinate	Urodynamic evaluation by specialists (urethrocystometry, stress maneuvers)	Stress: sensitivity 52%; specificity 88%; PPV 71%. Urge: sensitivity 37%; specificity 87%; PPV 59%. Mixed: sensitivity 72%; specificity 49%; PPV 42%.
Not reported	Detrusor Instability Score: score ≥5	Clinical diagnosis by specialists (urogynecologic examination, PVR, stress test, urodynamic pressure measurements urethrocystoscopy)	Detrusor instability: sensitivity 60%; specificity 77%; PPV 82%; NPV 52%.
12% with urge urinary incontinence; 47% stress; 26% mixed; 15% no findings	Urgency score: score ≥6	Urodynamic tests by specialists (pelvic examination, urine analysis, PVR, urethrocystometry)	Stress: sensitivity 19%; specificity 32%; PPV 20%; NPV 31%. Urge: sensitivity 93%; specificity 32%; PPV 26%; NPV 98%. Mixed: sensitivity 64%; specificity 62%; PPV 37%; NPV 84%.
14% with urge incontinence; 20% stress; 27% mixed; 10% symptomatic prolapse; 2% total incontinence; 2% urinary retention; 2% pelvic pain; 6% other	Urogenital Distress Inventory, 6 items: stress incontinence if question 3 score ≥2; bladder outlet obstruction if question 5 score ≥2 or score ≥ all others; detrusor over activity if question 1 score ≥2 and/or question 2 score ≥2	Urodynamic tests by specialists (measurement of bladder volumes and leak pressures)	Stress: sensitivity 84.8%; specificity 63.4%. Bladder outlet obstruction (sensitivity; specificity): question 5 score ≥2 (3.9%; 70.1%); question 5 score ≥all others (39.0%; 85.1%). Detrusor instability (sensitivity; specificity): question 1 score ≥2 (75.0%; 32.6%); question 2 score ≥2 (83.3%; 50.0%); question 1 and 2 score ≥2 (68.6%; 63.8%).
83% with symptoms; 14% urgency, frequency, and dysuria without urine loss	Stress and urge incontinence: self-reported symptoms	Urodynamic evaluations by specialists (urine culture, physical examinations, 24-h voiding diary, uroflowmetry, Q-tip test, urethral calibration, cystourethroscopy)	Stress: sensitivity 100%; specificity: 65.2%; PPV 86.9%; NPV 100%. Urge: sensitivity 77.9%; specificity 38.7%; PPV 36.6%; NPV 79.5%.
54% self-reported incontinence; 53% using pads	MISI total: total score ≥7; stress incontinence: subdomain score ≥3; urge incontinence: subdomain score ≥5	Clinical evaluation by specialists (pelvic examination, Q-tip test, PVR, urodynamics, paper towel test)	Total: sensitivity 84%; specificity 75%; PPV 75%; NPV 84%; AUROC 0.88. Stress: sensitivity 77%; specificity 76%; PPV 43%; NPV 86%; AUROC 0.79. Urge: sensitivity 86%; specificity 73%; PPV 73%; NPV 92%; AUROC 0.88.
56% with stress urinary incontinence	Stress incontinence question "Do you lose urine by spurts during coughing, sneezing, or lifting?": positive response	Clinical evaluation by specialists (pelvic examination, Q-tip test, urodynamic tests, cystometry)	Stress incontinence: sensitivity 93%; specificity 19%; PPV 59%; NPV 41%.

Appendix Table 2. Quality and Applicability Ratings of Studies Meeting Inclusion Criteria

Study, Year (Reference)	Patient Selection*						Reference Standard*				Sensitivity; Specificity; AUROC	Rating	
	1. Spectrum	1. Sample Size >50	2. Sample Selection	3. Eligibility Criteria	4. Minimal Attrition	5. Test Described	6. Credible	6. Replicable	7. Blinding	8. Applied to All		Quality†	Applicability‡
Basra et al, 2012 (30)	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Fair	High
Bergman and Bader, 1990 (31)	No	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Poor	Low
Borup et al, 2008 (32)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Poor	High
Bradley et al, 2005 (33)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Poor	Low
Brown et al, 2006 (34)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Poor	High
Cardozo et al, 2014 (35)	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear	Unclear	Unclear	Yes	Yes	Poor	High
Diokno et al, 1990 (36)	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Unclear	Yes	Yes	Poor	High
Gunthorpe et al, 2000 (37)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Fair	High
Haeusler et al, 1995 (38)	No	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Unclear	Yes	Yes	Poor	Low
Khan et al, 2004 (39)	No	Yes	Yes	Yes	No	Unclear	Yes	Yes	Unclear	Yes	Yes	Poor	Low
Klingele et al, 2002 (40)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Poor	Low
Klovning et al, 1996 (41)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Poor	Low
Kujansuuand Kauppila, 1982 (42)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Poor	Low
Lemack and Zimmern, 1999 (43)	No	Yes	Unclear	Yes	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Poor	Low
Sand et al, 1988 (44)	No	Yes	Unclear	No	Yes	No	Yes	Yes	Unclear	Yes	Yes	Poor	Low
Suskind et al, 2015 (45)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Good	High
Walters and Shields, 1988 (46)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Poor	Low

AUROC = area under the receiver-operating characteristic curve.

* Quality criteria definitions: 1: Test was applied to an appropriate number and spectrum of patients (>50 participants with and without urinary incontinence). 2: Population tested was consecutive or random. 3: Clear eligibility criteria were described. 4: Attrition was reported and loss to follow-up was minimal. 5: Test was adequately described and reproducible. 6: Reference standard was credible and replicable. 7: Outcome assessors were blinded to the reference standard. 8: Reference standard was applied to all patients or a random subset.

† Definition of ratings based on criteria in footnote 1: Good: Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (>500) of broad-spectrum patients with and without disease; attempts to enroll a random or consecutive sample of patients who meet inclusion criteria; screening cutoffs were stated before enrollment. Fair: Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; moderate sample size (100–500 participants) and a "medium" spectrum of patients (i.e., applicable to many settings where the diagnostic test would be applied). Poor: Has important limitation, such as use of inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; small sample size (<100) of very narrow selected spectrum of patients; or these components of study not well described.

‡ High: Participants were selected from the community or primary care or nonspecialty clinics; test and reference standards are relevant to clinical practice in the United States; test is feasible for screening by nonspecialists in clinical settings. Low: Participants were selected exclusively from referral clinics; test and reference standards are not relevant to clinical practice in the United States; test is may not be feasible for screening by nonspecialists in clinical settings; or these components of the study were not described.