JAMA | US Preventive Services Task Force | EVIDENCE REPORT

Periodic Screening Pelvic Examination
Evidence Report and Systematic Review for the US Preventive Services Task Force

Janelle M. Guirguis-Blake, MD; Jillian T. Henderson, PhD; Leslie A. Perdue, MPH

IMPORTANCE Recent changes in the periodicity of cervical cancer screening have led to questions about the role of screening pelvic examinations among asymptomatic women.

OBJECTIVE To systematically review literature on health benefits, accuracy, and harms of the screening pelvic examination for gynecologic conditions for the US Preventive Services Task Force (USPSTF).

DATA SOURCES MEDLINE, PubMed, and Cochrane Central Register of Controlled Trials for relevant English-language studies published through January 13, 2016, with surveillance through August 3, 2016.

STUDY SELECTION Two reviewers independently screened abstracts and studies. The search yielded 8678 unique citations; 316 full-text articles were reviewed, and 9 studies including 27 630 patients met inclusion criteria.

DATA EXTRACTION AND SYNTHESIS Two reviewers rated study quality using USPSTF criteria.

MAIN OUTCOMES AND MEASURES Morbidity; mortality; diagnostic accuracy for any gynecologic cancer or condition except cervical cancer, gonorrhea, and chlamydia, which are covered by other USPSTF screening recommendations; harms (false-positive rates, false-negative rates, surgery rates).

RESULTS No trials examined the effectiveness of the pelvic examination in reducing all-cause mortality, reducing cancer- and disease-specific morbidity and mortality, or improving quality of life. Eight studies reported accuracy for the screening pelvic examination: ovarian cancer (4 studies; n = 26 432), bacterial vaginosis (2 studies; n = 930), trichomoniasis (1 study; n = 779), and genital herpes (1 study; n = 779). In the 4 ovarian cancer screening studies, low prevalence of ovarian cancer consistently resulted in low positive predictive values (PPVs) and false-positive rates, with a lack of precision in accuracy estimates (sensitivity range, 0%-100%; specificity range, 91%-99%; PPV range, 0%-3.6%; negative predictive value [NPV] range, ≥99%). Each diagnostic accuracy study for bacterial vaginosis, trichomoniasis, and genital herpes was performed in a high-prevalence population with substantial proportions of symptomatic patients and reported accuracy characteristics for individual physical examination findings (bacterial vaginosis, homogeneous discharge: sensitivity range, 69%-79%; specificity range, 54%-97%; PPV range, 52%-95%; NPV range, 79%-80%; herpes simplex virus, vulvar ulcerations: sensitivity, 20%; specificity, 98%; PPV, 88%; NPV, 57%; trichomoniasis, colpitis macularis: sensitivity, 2%; specificity, 100%; PPV, 100%; NPV, 85%). Surgery rates resulting from an abnormal screening pelvic examination for ovarian cancer ranged from 5% to 36% at 1 year, with the largest study reporting an 11% surgery rate and 1% complication rate within 1 year of a screening pelvic examination with abnormal findings.

CONCLUSIONS AND RELEVANCE No direct evidence was identified for overall benefits and harms of the pelvic examination as a 1-time or periodic screening test. Limited evidence was identified regarding the diagnostic accuracy and harms of routine screening pelvic examinations in asymptomatic primary care populations.

JAMA. 2017;317(9):954-966. doi:10.1001/jama.2016.12819

Editorial page 910

Related article page 947 and JAMA Patient Page page 984

Supplemental content

Related article at jamainternalmedicine.com

Author Affiliations: Department of Family Medicine, University of Washington, Tacoma (Guirguis-Blake); Kaiser Permanente Research Affiliates Evidence-based Practice Center, Center for Health Research, Kaiser Permanente, Portland, Oregon (Guirguis-Blake, Henderson, Perdue).

Corresponding Author: Janelle M. Guirguis-Blake, MD, Kaiser Permanente Research Affiliates EPC, Department of Family Medicine, University of Washington, 521 Martin Luther King Jr Way, Tacoma, WA 98405 (jguirgui@u.washington.edu).

In 2012, 44.2 million pelvic examinations were performed in US outpatient visits,[1] and in a 2008-2009 survey of physicians, 69.1% agreed that the pelvic examination is an effective screening test for ovarian cancer.[2] Multiple malignant and benign gynecologic conditions can be detected with pelvic examination, and the value of early detection among asymptomatic women depends on natural history, prevalence, disease morbidity, and early treatment effectiveness.[3,4] For some gynecologic conditions, alternative effective screening approaches are recommended (eg, nucleic acid amplification test for gonorrhea and chlamydia,[5] Papanicolaou smear and human papillomavirus testing for cervical cancer[6]). As recommended intervals for cervical cancer screening have been lengthened,[6-8] the health benefits of annual pelvic examinations among asymptomatic women have increasingly been questioned.[9-12] Current professional organizations' guidelines for the administration of the screening pelvic examination are based on limited evidence and expert opinion, and these guidelines vary widely.[13-15] The current evidence review was undertaken to inform US Preventive Services Task Force (USPSTF) deliberations on whether nonpregnant women without gynecologic symptoms would obtain net health benefits from periodic screening pelvic examinations.

Methods

Scope of Review

In this evidence review, the pelvic examination was defined as visual inspection of the external genitalia; a speculum examination of the vagina and cervix; bimanual examination of the adnexa, uterus, and cervix; and may include the rectovaginal examination.[13] This review addresses the benefits and harms of screening with the pelvic examination for benign and malignant gynecologic conditions as well as the diagnostic accuracy of the pelvic examination in detecting individual benign and malignant gynecologic conditions. The analytic framework and key questions that guided the review are shown in **Figure 1**. Detailed methods are available in the full evidence report at https://www.uspreventiveservicestaskforce.org/Page/Document/final-evidence-review140/gynecological-conditions-screening-with-the-pelvic-examination.

Although the pelvic examination is commonly performed in adolescent and pregnant women, these populations were specifically excluded from the scope of this review. The purpose of conducting the pelvic examination in unselected nonpregnant adult women often differs from that in special populations of adolescents (eg, Tanner staging, congenital abnormality case-finding) or pregnant women (eg, assessment of pregnancy dating, pelvic outlet adequacy, cervical dilation). Likewise, screening for congenital gynecological conditions was excluded because this review focuses on routine periodic screening, and many of the congenital conditions would be detected at the symptomatic stage, during pregnancy, at infertility workup, or incidentally during cervical cancer screening.

In addition, the USPSTF previously determined that there is good evidence for primary screening approaches for cervical cancer (such as Papanicolaou and human papillomavirus cotesting) and for gonorrhea and chlamydia (such as nucleic acid amplification testing). Since the pelvic examination alone is less accurate than these screening approaches, these conditions were not included in this review.

Data Sources and Searches

MEDLINE, PubMed (publisher-supplied references only), and the Cochrane Central Register of Controlled Trials were searched to locate primary studies informing the key questions (eMethods in the Supplement) and published from the earliest date indexed (1946 for MEDLINE) through January 13, 2016. The database searches were supplemented with experts' suggestions and reference lists from all other recent systematic reviews.[17,18] ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform were searched for ongoing trials. The National Cancer Institute provided previously unpublished data on the subset of randomized women from the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial who had undergone bimanual palpation of the ovaries and adnexa as well as rectovaginal examination (Paul Pinsky, PhD, National Cancer Institute, written communication, May 2, 2016); the 5-year follow-up results were subsequently published.[19] Since January 2016, ongoing surveillance was conducted through article alerts and targeted searches of high-impact journals to identify major studies published in the interim that may affect the conclusions or understanding of the evidence and therefore the related USPSTF recommendation. The last surveillance was conducted on August 3, 2016, and identified no new studies.

Study Selection

Two investigators independently reviewed 8678 titles and abstracts by using an online platform (abstrackr[20]) and 316 articles (**Figure 2**) with specified inclusion criteria (eTable 1 in the Supplement). Discrepancies were resolved through consensus and consultation with a third investigator. Articles that did not meet inclusion criteria or those rated as poor quality were excluded; criteria for establishing study quality are noted in eTable 2 in the Supplement. To avoid missing studies using the pelvic examination as a secondary screening test (eg, ovarian cancer screening studies using cancer antigen 125 [CA-125] measurement and ultrasound technology that also included a pelvic examination component), reviewers were more inclusive during the review of abstracts and titles. As a result, many studies were excluded at the full-text review.

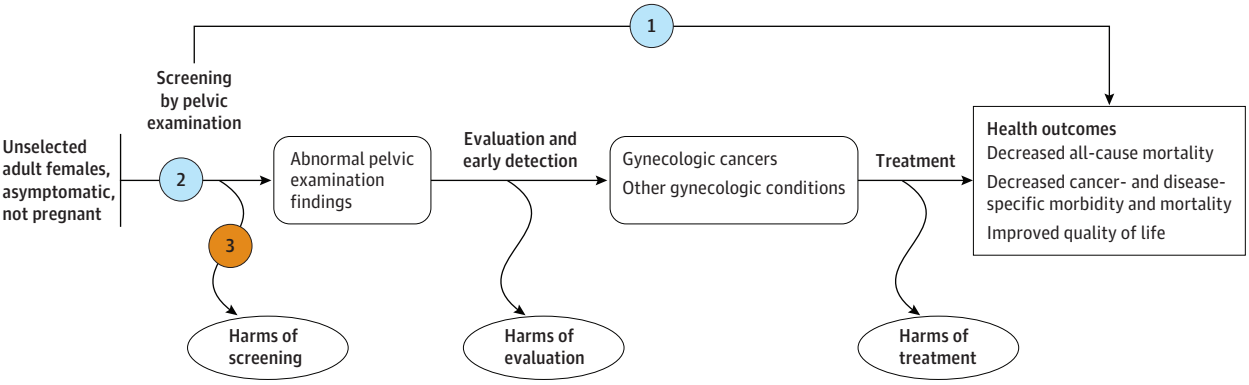
Eligible studies included unselected adult females who were not symptomatic or pregnant and were conducted in developed countries, as defined by "very high" development according to the 2014 United Nations Human Development Index.[21] Studies conducted solely in symptomatic populations were excluded.

Any study that examined the relationship between pelvic examination and all-cause mortality, cancer- or disease-specific morbidity or mortality, or quality of life was eligible for inclusion. In addition, studies examining the screening accuracy of the pelvic examination in a single encounter or as a periodic program of screening were eligible.

Data Extraction and Quality Assessment

One reviewer extracted study-level data into standardized evidence tables, and a second checked for accuracy. At least 2 reviewers critically appraised included studies using the

Figure 1. Analytic Framework



Key questions

1 What is the direct evidence for the effectiveness of the pelvic examination in (a) reducing all-cause mortality, (b) reducing cancer- and disease-specific morbidity and mortality, and (c) improving quality of life?

2 What are the test performance characteristics of the pelvic examination (sensitivity, specificity, and positive and negative predictive values) in screening for gynecologic cancers and other gynecologic conditions?

3 What are the adverse effects of screening by pelvic examination?

Refer to USPSTF Procedure Manual for interpretation of the analytic framework.[16]

Newcastle Ottawa Scales[22] for cohort and case-control studies and Quality Assessment of Diagnostic Accuracy Studies I and II[23,24] for studies of diagnostic accuracy adapted to align with the USPSTF's design-specific quality criteria[16] (eTable 2 in the Supplement). Disagreements in quality rating were resolved by consensus or consultation with a third independent reviewer. Included studies were limited to those published in English that were rated as good or fair quality using USPSTF quality-rating standards.[16]

Data Synthesis and Analysis

Results were qualitatively synthesized by key question (Figure 1). For all of the studies of diagnostic accuracy, sensitivity and specificity were calculated in Stata version 13.1 (StataCorp), using Jeffrey confidence intervals from 2 × 2 tables constructed from data reported in the primary studies. In many cases the data presented differ slightly from the data in the published article because of these calculations. For diagnostic accuracy studies, in addition to the standard test performance characteristics (sensitivity, specificity, positive predictive value [PPV], and negative predictive value [NPV]), the following outcomes were calculated: condition prevalence in the study population, percentage of patients screening positive, false-positive rate, and false-negative rate. Since there were a limited number of studies for each condition, no pooled analyses were conducted.

Results

Nine studies met the inclusion criteria for this systematic review (Figure 2). The results are organized by key question. Eight of the 9 included studies reported outcomes for both key question 2 (accuracy) and key question 3 (harms). A single additional study on key question 3 (harms) was also included.

Benefits of Screening

Key Question 1. What is the direct evidence for the effectiveness of the pelvic examination in (a) reducing all-cause mortality, (b) reducing cancer- and disease-specific morbidity and mortality, and (c) improving quality of life?

There was no direct evidence comparing the effectiveness of pelvic examination screening and no screening on patient health outcomes.

Accuracy of Screening Pelvic Examination

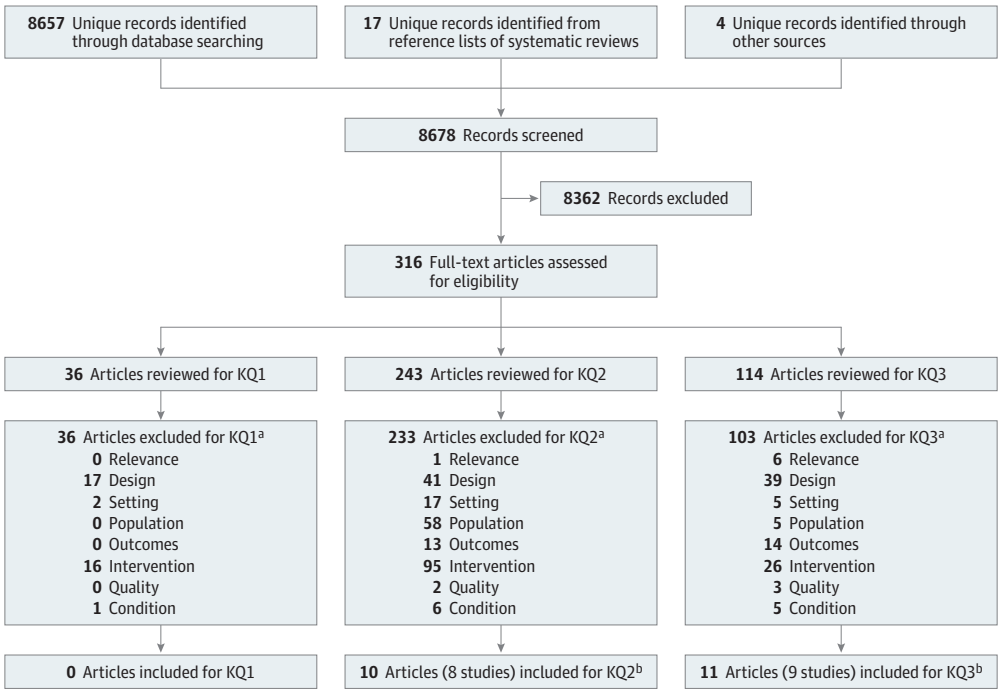
Key Question 2. What are the test performance characteristics of the pelvic examination (sensitivity, specificity, PPV, and NPV) in screening for gynecologic cancers and other gynecologic conditions?

Studies examining the diagnostic screening accuracy of the pelvic examination were identified for 4 gynecologic conditions: ovarian cancer, bacterial vaginosis, trichomoniasis, and genital herpes. In total, 4 studies examined the accuracy of the pelvic examination to detect ovarian cancer in asymptomatic populations, while 1 to 2 studies in high-prevalence populations with substantial proportions of symptomatic patients were identified for each of the infectious diseases (bacterial vaginosis, trichomoniasis, and genital herpes).

Ovarian Cancer

Data from 1 good-quality study and 3 fair-quality studies (n = 26 432) were identified to estimate the screening accuracy of pelvic examination for ovarian cancer detection.[25-27] The PLCO

Figure 2. Literature Search Flow Diagram



KQ indicates key question.

[a] Reasons for exclusion: Relevance: Study aim not relevant; Design: Study did not use an included design; Setting: Study was not conducted in a country relevant to US practice; Population: Study was not conducted in an unselected population; Outcomes: Study did not have relevant outcomes or had incomplete outcomes; Intervention: Study used an excluded intervention/screening approach; Poor quality: Study did not meet criteria for fair or good quality; Condition: Study examined an excluded condition (cervical cancer, gonorrhea, or chlamydia).

[b] Eight studies (in 10 articles) were included for both KQ2 and KQ3.

trial (n = 20 872), a large multicenter US randomized clinical trial, recruited average-risk women aged 55 to 74 years from the community to test the benefits and harms of ovarian cancer screening using a combination of 3 modalities: measurement of CA-125 levels, ultrasound, and ovarian palpation/rectovaginal examination.[28] Ovarian palpation was stopped after the first 5 years of the trial because no ovarian cancers were detected solely with this test; accuracy information and clinical outcomes resulting from ovarian palpation were not included in published results from the trial. The unpublished results were obtained from National Cancer Institute PLCO investigators. The other 3 were prospective diagnostic accuracy studies (n = 5560) conducted in Greece, Australia, and the United Kingdom; they primarily recruited average-risk women aged 40 to 45 years and older from the community (**Table 1**).

Mean or median age ranged from 51 to 63 years. In the 2 studies reporting menopausal status, 43%[27] and 65%[26] of participants were postmenopausal. In the other 2 studies, all or nearly all were postmenopausal.[25,28] None of the studies excluded women with a family history of ovarian cancer, and 1 actively recruited younger women with family history.[27] One study defined the test as bimanual ovarian palpation plus rectovaginal examination,[28] another clearly defined the index pelvic examination test as bimanual and speculum examination,[27] while the other 2 defined the index test only as a "pelvic exam"[26] or "vaginal exam."[25] Two studies specified that experienced gynecologists or examiners performed the examination,[26,28] and in the other studies a single examiner[27] or 1 of 2 examiners[25] performed the examinations.

A positive index test finding was a palpable pelvic or adnexal mass of any size. The PLCO trial used a reference standard for ovarian cancer diagnoses captured in medical records and participant questionnaires; investigators reported cancer incidence at 1 year and up to 5 years for the purposes of this analysis.[19,28,29] The other 3 studies used a reference standard of incident ovarian cancer reported in a 1-year patient questionnaire and measurement of CA-125 levels with or without ultrasound.[25-27] Abnormal ultrasound findings or CA-125 levels were variably defined in the different studies.

Ovarian cancer prevalence was reported as 0.1% in 3 studies[25,26] and 0.04% in 1 study[27]; the longer follow-up from the PLCO trial (5 years) yielded 0.3% prevalence of ovarian cancer (**Table 2**). The comparable 1-year follow-up results for the proportion of screened patients with a positive pelvic examination finding ranged from 1.2% to 8.7%. Sensitivity was reported as 100.0% in 2 of the studies[25,26] (1-2 ovarian cancer cases were palpable on pelvic examination) and as 0% (95% CI, 0%-85.3%) in the study in which the single case of ovarian cancer was not detected with the pelvic examination.[27] The PLCO trial reported a sensitivity of 4.3% (95% CI, 0.5%-18.6%) and 2.8% (95% CI, 0.6%-8.6%) from the first screening examination at 1-year follow-up (23 cases) and up to 5 years' follow-up (72 cases). Specificity ranged from 91.4% (95% CI, 90.1%-92.6%) to 98.8%

Table 1. Study Characteristics, Ovarian Cancer Screening

Source	Country	Recruitment Setting	Study Aim	Inclusion/Exclusion Criteria	Screening Test Description	Reference Standard Description
PLCO,[19,28] 2016 (Paul Pinsky, PhD, National Cancer Institute, written communication, May 2, 2016)	United States	Community	Determine the effect of specific cancer screening tests on cause-specific mortality	Inclusion: Age 55-74 y Exclusion: Undergoing treatment for cancer (excluding basal cell and squamous cell skin cancer); known prior cancer of the lung, colon, rectum, or ovary; previous surgical removal of 1 lung or entire colon; had colonoscopy, sigmoidoscopy, or barium enema in past 3 y; unable or unwilling to sign consent form	Palpable ovarian mass or cul-de-sac nodularity (for obese patients with nonpalpable ovaries, the examination was considered negative)	Diagnosis of ovarian cancer within 1 to 5 y of examination based on medical records and patient questionnaires for ovarian cancer diagnoses. (All women also received TVU and CA-125 measurement. Although these results were available to the provider, they are not being used as a reference standard.)
Adonakis et al,[26] 1996	Greece	Community	Investigate effectiveness of pelvic examination and CA-125 measurement followed by ultrasonography as a screening method	Inclusion: Aged ≥45 y without any evidence of adnexal pathology Exclusion: History of ovarian cancer (familial or not) or any other malignancy; bilateral oophorectomy; ascites	Detection of palpable adnexal mass on pelvic examination	TVU for participants with abnormal pelvic examination result or serum CA-125 ≥35 U/mL 1-y follow-up visit to measure CA-125 levels for those with normal pelvic examination results and CA-125 levels
Grover et al,[27] 1995	Australia	Community	Assess effectiveness of serum CA-125 measurement plus vaginal examination as a screening test	Inclusion: Apparently healthy and aged ≥40 y (younger included if they had a family history of ovarian cancer) Exclusion: NR	Adnexal mass palpable during bimanual examination in postmenopausal women, or if a larger than normal-size ovary was palpable in premenopausal women	Abdominal and/or vaginal ultrasonography for participants with abnormal pelvic examination result or serum CA-125 >35 U/mL 1-y postal questionnaire for all patients
Jacobs et al,[25] 1988	United Kingdom	Community	Examine screening capabilities of vaginal examination, CA-125 measurement, and ultrasonography in various combinations	Inclusion: Age >45 y; amenorrheic for >12 mo Exclusion: History of ovarian cancer or bilateral oophorectomy; being treated for any malignancy	Palpable pelvic mass of any size that could be clinically distinguished as being separate from the uterus and gastrointestinal tract during vaginal examination	Abdominal ultrasonography for those with abnormal pelvic examination result or serum CA-125 >30 U/mL 1-y postal questionnaire for those with normal pelvic examination results and CA-125 levels

Abbreviations: CA-125, cancer antigen 125; NR, not reported; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; TVU, transvaginal ultrasonography.

(95% CI, 98.7%-99.0%). Calculated PPVs ranged from 0% (95% CI, 0%-6.0%) to 3.6% (95% CI, 0.4%-15.5%), and NPV was 99% or greater for all studies.[19,25-27] In all studies except the PLCO trial, accuracy estimates had wide confidence intervals because of the very low event rate.

Bacterial Vaginosis

No screening studies were conducted solely in asymptomatic primary care populations. Two fair-quality US studies (n = 930) with large proportions of symptomatic patients assessed the accuracy of different approaches to diagnosing bacterial vaginosis, including findings visible on the pelvic examination.[30,31] Gutman et al (n = 269)[30] recruited any woman undergoing a "routine pelvic examination" from a hospital-based primary care clinic, colposcopy clinic, or research clinic, whereas Eschenbach et al (n = 661)[31] recruited nonpregnant women aged 16 to 50 years from a sexually transmitted infection (STI) clinic undergoing a "standardized pelvic examination"(**Table 3**). Neither study had a primary aim of estimating the accuracy of the pelvic examination; instead, they explored different clinical signs and diagnostic criteria for bacterial vaginosis measured against a gold standard (Nugent criteria[30] or criteria based on pH; Gram stain microscopy[31,34]).

The patients in the study by Gutman et al[30] had a mean age of 24.1 years, with 38% of patients being white, 30% black, and 27% Hispanic; the study by Eschenbach et al[31] did not report patient characteristics. In the study by Gutman et al,[30] 33% of patients were symptomatic, while the study by Eschenbach et al[31] reported that 59% of participants presented with some pelvic or abdominal symptom as a chief presenting symptom. Both studies reported a high prevalence of bacterial vaginosis (39% and 47%).[30,31] Gutman et al[30] reported the sensitivity and specificity of thin, homogeneous discharge as 78.8% (95% CI, 70.3%-85.8%) and 53.9% (95% CI, 46.3%-61.4%), respectively (PPV, 51.9% [95% CI, 44.1%-59.6%] and NPV, 80.2% [95% CI, 72.0%-86.8%]) (**Table 4**). Eschenbach et al[31] reported the sensitivity and specificity of homogeneous discharge as 69.2% (95% CI, 63.4%-74.5%) and 97.2% (95% CI, 94.9%-98.6%), respectively (PPV, 95.3% [95% CI, 91.7%-97.7%]; NPV, 79.0% [95% CI, 74.8%-82.8%]).

Genital Herpes (Herpes Simplex Virus Type 1 or Type 2)

No screening studies were conducted solely in asymptomatic primary care populations. One fair-quality trial by Koutsky et al (n = 779)[32] provided data on the accuracy of specific pelvic examination findings in detecting this condition. The study

Table 2. Summary of Diagnostic Accuracy of Pelvic Examination for Ovarian Cancer (Key Questions 2 and 3)

Source	Quality[a]	No. (Age, y)	Follow-up, y[b]	Yield, No./Total (%)	True Positive	False Negative	False Positive	True Negative	Measure (95% CI)[c]					
									Sensitivity	Specificity	PPV	NPV	False-Positive Rate	False-Negative Rate
PLCO,[19,28] 2016 (Paul Pinsky, PhD, National Cancer Institute, written communication, May 2, 2016)	Good	20 872 (63 [mean])	1	23/20 872 (0.1)	1	22	242	20 607	4.3 (0.5-18.6)	98.8 (98.7-99.0)	0.4 (0.04-1.9)	99.9 (99.8-99.9)	1.2 (1.0-1.3)	95.7 (81.4-99.5)
			5	72/20 872 (0.3)[d]	2	70	241	20 559	2.8 (0.6-8.6)	98.8 (98.7-99.0)	0.8 (0.2-2.6)	99.7 (99.6-99.7)	1.2 (1.0-1.3)	97.2 (91.4-99.4)
Adonakis et al,[26] 1996	Fair	2000 (58 [mean])	1	2/2000 (0.1)	2	0	172	1826	100.0 (33.3-100.0)[e]	91.4 (90.1-92.6)[e]	1.2 (0.2-3.6)[e]	100.0 (99.9-100.0)[e]	8.6 (7.4-9.9)[e]	0 (0-66.7)[e]
Grover et al,[27] 1995	Fair	2550 (51 [median])	1	1/2550 (0.04)	0	1	40	2509	0 (0-85.3)	98.4 (97.9-98.9)	0 (0-6.0)	100.0 (99.8-100.0)	1.6 (1.1-2.1)	100.0 (14.7-100.0)
Jacobs et al,[25] 1988	Fair	1010 (54 [median])	1	1/1010 (0.1)	1	0	27	982	100.0 (14.7-100.0)	97.3 (96.2-98.2)	3.6 (0.4-15.5)	100.0 (99.7-100.0)	2.7 (1.8-3.8)	0 (0-85.3)

Abbreviations: NPV, negative predictive value; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; PPV, positive predictive value.

[a] Quality assessed using criteria from Quality Assessment of Diagnostic Accuracy Studies (QUADAS) I[23] and II[24] instrument.

[b] Length of follow-up for ovarian cancer diagnosis.

[c] Calculated (not reported in study).

[d] Includes all participants also diagnosed within the first year (n = 23).

[e] Adonakis et al[26] treated ambiguous pelvic examinations as negative screens, whereas we treated the ambiguous examinations as positive screens. Author-reported sensitivity was 97.2%; specificity, 66.7%; and PPV, 3.4%.

recruited nonpregnant women aged 16 to 50 years from the same population seen in the STI clinic for the bacterial vaginosis study by Eschenbach et al[31] and for the trichomoniasis study by Wølner-Hansson et al[33] (Table 3). All pelvic examinations were performed by 1 "women's health care specialist." Mean age was 24 years, and 70% of participants were white. Almost all patients were sexually active (98%). Condoms were used as the primary method of contraception by 7%, and 33% did not use any contraception. Of the participants, 10% were symptomatic. The index test was a "genital examination" (colposcopy findings were not considered for this review). A positive pelvic examination result was defined as clinician-detected lesions, and only vulvar ulcerations and tender inguinal nodes were used in this review for accuracy calculations. All patients received the reference test, which included cultures from urine, cervical swabs, anal swabs, and any lesion swabs (all herpes simplex virus [HSV] isolates were confirmed and typed by direct immunofluorescence with use of mouse monoclonal antibodies), as well as serum testing for HSV-1 or HSV-2 antibodies using Western blotting.

Nearly half (48%) of all study participants were diagnosed with genital herpes at some stage of the disease: 6% were diagnosed at the first episode, 5% had symptomatic recurrence, 2% had asymptomatic shedding, and 35% had latent HSV-2 infection (defined as HSV-2 antibodies present without signs or symptoms). Of those with any stage of genital herpes, 78% had latent disease. The presence of vulvar ulcerations had a sensitivity of 19.6% (95% CI, 15.8%-23.9%) and specificity of 97.5% (95% CI, 95.7%-98.7%) in detecting genital HSV at any stage (PPV, 88.0% [95% CI, 79.7%-93.6%]; NPV, 57.0% [95% CI, 53.3%-60.7%]) (Table 4). Similarly, the presence of tender inguinal lymphadenopathy had a sensitivity of 14.2% (95% CI, 11.0%-18.1%) and specificity of 97.1% (95% CI, 95.1%-98.4%) (PPV, 81.5% [95% CI, 70.8%-89.5%]; NPV, 55.3% [95% CI, 51.7%-58.9%]).[32]

Trichomoniasis

There were no screening studies conducted solely in asymptomatic primary care populations. One fair-quality trial by Wølner-Hanssen et al (n = 779)[33] was set in the same STI clinic as the trials by Eschenbach et al[31] and Koutsky et al,[32] with a high proportion of symptomatic patients aimed to analyze the clinical manifestations of trichomoniasis and determined the accuracy of specific clinical findings on pelvic examination in detecting trichomoniasis (Table 3). An abnormal pelvic examination finding was defined as clinician-reported moderate to markedly increased vaginal discharge compared with that seen in patients without genital infections. The reference test for identifying trichomoniasis was culture. At least half of the patients reported symptoms of yellow discharge (23%), abnormal vaginal odor (36%), or vulvar itching (51%). The prevalence of culture-confirmed trichomoniasis was 15%. For the most specific clinical sign, colpitis macularis or "strawberry cervix" (detected grossly, without a colposcope), sensitivity was 1.7% (95% CI, 0.4%-5.3%) and specificity was 100.0% (95% CI, 99.6%-100.0%) (PPV, 100.0% [95% CI, 33.3%-100.0%]; NPV, 85.1% [95% CI, 82.4%-87.4%]) (Table 4). For other individual clinical findings, sensitivity ranged from 8.0% (95% CI, 3.8%-14.5%) to 59.2% (95% CI, 49.3%-68.5%) and specificity from 76.1% (72.4%-79.5%) to 99.1% (98.1%-99.7%); PPVs ranged from 30.2% (95% CI, 24.0%-37.0%) to 61.5%

Table 3. Study Characteristics, Infectious Diseases

Source	Country	Recruitment Setting	Study Aim	Inclusion/Exclusion Criteria	Race/Ethnicity, No. (%)	Prevalence of Symptoms, No. (%)	Screening Test	Reference Standard Description
Bacterial Vaginosis								
Gutman et al,[30] 2005	United States	Hospital	Determine whether current clinical criteria for diagnosing bacterial vaginosis can be simplified by using 2 clinical criteria rather than the standard (3 of 4 Amsel criteria)	Inclusion: any woman undergoing a speculum examination Exclusion: large amount of vaginal bleeding on examination	White: 103 (38)	Any symptoms: 88 (32.7) Vaginal discharge: 64 (23.8) Foul odor: 38 (14.1) Vaginal itching: 17 (6.3) Vaginal burning: 7 (2.6)	Thin, homogeneous vaginal discharge	Nugent criteria: score of ≥7 defined a diagnosis of bacterial vaginosis[a]
Eschenbach et al,[31] 1988[b]	United States	STI clinic	Compare accuracy of Gram stain criteria for bacterial vaginosis with composite clinical criteria for diagnosing bacterial vaginosis	Inclusion: age 16-50 y, English speaking Exclusion: pregnant, used oral antibiotics or vaginal medication in previous 14 d, hysterectomized. severely mentally or physically incapacitated, *Trichomonas vaginalis* (by culture), no evaluable Gram stain	NR	Any pelvic or abdominal symptom as chief presenting symptom: 390 (59)	Standardized pelvic examination, with attention to appearance of vulva, vagina, and cervix (eg, erythema, friability of cervix, color of cervical mucus), characteristics of vaginal discharge (amount, color, other characteristics), and tenderness (cervical, uterine, adnexal)	pH of vaginal contents, clue and epithelial cells present on microscopy; fishy, amine-like odor
Genital Herpes								
Koutsky et al,[32] 1992[b]	United States	STI clinic	Assess relative merits of different approaches to detecting genital HSV infection, including the approach of clinical examination and viral isolation	Inclusion: age 16-50 y; English-speaking Exclusion: pregnant, used oral antibiotics or vaginal medication in previous 14 d, hysterectomized, severely mentally or physically incapacitated	White: 545 (70)	22% of women with evidence of herpes presented symptomatically.	Genital examination, looking for vulva ulcerations and tender inguinal nodes	For HSV isolation: collection of urine, specimens from cervix and anal canal, swabs from external genital lesions; serum Western blot for antibodies
Trichomoniasis								
Wølner-Hansson et al,[33] 1989[b]	United States	STI clinic	Identify relationships of specific genital microbial pathogens to clinical manifestations	Inclusion: age 16-50 y, English-speaking Exclusion: pregnant, used oral antibiotics or vaginal medication in previous 14 d, hysterectomized, severely mentally or physically incapacitated	White: NR (70)	Yellow discharge: 179 (23) Abnormal vaginal odor: 278 (36) Vulvar itching: 397 (51)	Standardized pelvic examination with specific attention to appearance of vulva, vagina, and cervix; abnormal results included colitis macularis ("strawberry cervix"), purulent discharge, frothy discharge, vulvar or vaginal erythema	Cultures from 2 vaginal specimens examined for growth of *Trichomonas vaginalis* (identified by characteristic morphology and motility in unstained wet mounts)

Abbreviations: HSV, herpes simplex virus; NR, not reported; STI, sexually transmitted infection.

[a] 0- to 10-point score describing numbers of *Lactobacilli*, *Gardenerella*, and curved gram-negative bacilli in ×100 microscopy field of Gram stain sample.

[b] These 3 studies include the same sample of women.

(35.0%-83.5%) and NPVs from 86.3% (83.5%-88.7%) to 91.4% (95% CI, 88.6%-93.7%) (Table 4).

Harms of Screening

Key Question 3. What are the adverse effects of screening by pelvic examination?

All accuracy studies (8 studies) were included for harms (false-positive rates and resulting diagnostic workup), and 1 additional study was included for other harms.

Ovarian Cancer

Additional imaging and unnecessary surgical intervention are potential harms of pelvic examination screening for ovarian cancer. The prevalence of laparoscopy or laparotomy for those with abnormal findings on pelvic examination ranged from 5% to 36%.[19,25-27] In the Greek study,[26] 17% of women with abnormal pelvic examination results underwent surgery because of the results. Pathology findings revealed 2 cases of ovarian cancer (1 metastatic, 1 stage Ia serous cystadenocarcinoma), 4 serous cystadenomas, 3 mucinous

Table 4. Summary of Diagnostic Accuracy of Pelvic Examination for Infectious Disease (Key Question 2 and Key Question 3)

Source	Quality[a]	No. (Mean Age, y)	Yield of Disease, No./Total (%)	Reference Standard	Screening Test	Measure (95% CI)[b]					
						Sensitivity	Specificity	PPV	NPV	False-Positive Rate	False-Negative Rate
Bacterial Vaginosis											
Gutman et al,[30] 2005	Fair	269 (24.1)	104/269 (38.7)	Nugent criteria	Thin, homogeneous discharge on pelvic examination	78.8 (70.3-85.8)	53.9 (46.3-61.4)	51.9 (44.1-59.6)	80.2 (72.0-86.8)	46.1 (38.6-53.7)	21.2 (14.2-29.7)
Eschenbach et al,[31] 1988[c]	Fair	661 (NR)	311/661 (47.0)	pH, microscopy	Homogeneous discharge	69.2 (63.4-74.5)	97.2 (94.9-98.6)	95.3 (91.7-97.7)	79.0 (74.8-82.8)	2.8 (1.4-5.1)	30.8 (25.5-36.6)
					Frothy discharge	2.3 (0.9-45.9)	100.0 (99.2-100.0)	100 (67.0-100.0)	55.0 (50.9-59.0)	0 (0-7.9)	97.7 (95.4-99.0)
					Increased discharge	9.4 (6.3-13.3)	95.6 (92.9-97.5)	64.1 (48.5-77.7)	55.8 (51.6-59.9)	4.4 (2.5-7.1)	90.6 (86.7-93.7)
					Yellow discharge	31.7 (26.3-37.5)	82.3 (77.6-86.3)	61.0 (52.7-68.9)	57.9 (53.1-62.5)	17.7 (13.7-22.4)	68.3 (62.5-73.7)
Genital Herpes											
Koutsky et al,[32] 1992[c]	Fair	779 (24)	372/779 (47.8)	Culture or serologic evidence of HSV	Vulvar ulcerations	19.6 (15.8-23.9)	97.5 (95.7-98.7)	88.0 (79.7-93.6)	57.0 (53.3-60.7)	2.5 (1.3-4.3)	80.4 (76.1-84.2)
					Tender inguinal nodes	14.2 (11.0-18.1)	97.1 (95.1-98.4)	81.5 (70.8-89.5)	55.3 (51.7-58.9)	2.9 (1.6-4.9)	85.8 (81.9-89.0)
Trichomoniasis											
Wølner-Hansson et al,[33] 1989[c]	Fair	779 (24)	118/778 (15.2)	Culture	Colpitis macularis	1.7 (0.4-5.3)	100.0 (99.6-100.0)	100.0 (33.3-100.0)	85.1 (82.4-87.4)	0 (0-0.4)	98.3 (94.7-99.6)
					Purulent discharge	59.2 (49.3-68.5)	76.1 (72.4-79.5)	30.2 (24.0-37.0)	91.4 (88.6-93.7)	23.9 (20.5-27.6)	40.8 (31.5-50.7)
					Frothy discharge	8.0 (3.8-14.5)	99.1 (98.1-99.7)	61.5 (35.0-83.5)	86.3 (83.5-88.7)	0.9 (0.3-1.9)	92.0 (85.5-96.1)
					Vaginal erythema	19.5 (13.1-27.3)	93.0 (90.9-94.8)	33.3 (23.1-45.0)	86.6 (84.0-89.0)	7.0 (5.2-9.1)	80.5 (72.7-86.9)

Abbreviations: HSV, herpes simplex virus; NPV, negative predictive value; NR, not reported; PPV, positive predictive value.

[a] Quality assessed using criteria from Quality Assessment of Diagnostic Accuracy Studies (QUADAS) I[23] and II[24] instrument.

[b] Calculated (not reported in study).

[c] These 3 studies include the same sample of women recruited from a sexually transmitted diseases clinic; population at higher risk and more symptomatic than average.

Table 5. Diagnostic Procedures Within 1 Year of a Positive Palpation Examination in Women Without an Ovarian Cancer Diagnosis; PLCO Only (Key Question 3)[a]

Diagnostic Procedure	No. (%) of Women Receiving Diagnostic Procedure Within 1 y of Positive Palpation Examination	
	Positive Palpation Examination in First Screening[b] (n = 242)	Positive Palpation Examination Anytime During 4 Rounds of Annual Palpation Screening Examinations[c] (n = 475)
Additional CA-125[d]	10 (4.1)	26 (5.5)
Additional TVU[d]	47 (19.4)	87 (18.3)
Abdominal CT	7 (2.9)	11 (2.3)
Surgery	31 (12.8)	53 (11.2)
Any complication[e]	4 (1.7)	5 (1.0)
Surgical complication	1 (0.4)	1 (0.2)

Abbreviations: CA-125, cancer antigen 125; CT, computed tomography; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; TVU, transvaginal ultrasound.

[a] Because of rolling recruitment and early termination of the palpation component in the screening intervention group, the number of palpation screening visits women completed was variable. On average, women received 2.4 palpation screening visits (28.0% received 1; 24.8% received 2; 24.4% received 3; and 22.8% received 4).

[b] Data source: Paul Pinsky, PhD, National Cancer Institute, written communication, May 2, 2016.

[c] Data source: Doroudi et al.[19]

[d] All women received CA-125 measurement and TVU as part of the PLCO trial protocol for ovarian cancer screening, and these results were available to each woman's primary care clinician. These 2 diagnostic procedures in the table indicate that additional CA-125 measurement and TVU were conducted with a woman's clinician within 1 y.

[e] Includes surgical, infection, cardiovascular, pulmonary, or other complication.

cystadenomas, 5 endometroid cysts of the ovary, 12 benign cysts, and 3 normal pathology results.[26] In the Australian study,[27] 2 women (5%) with abnormal findings on pelvic examination had surgery. The operations revealed that 1 patient had a fibroid uterus and 1 a normal (negative) result; the single case of ovarian cancer was not detected on pelvic examination.[27] In the UK study that recruited solely postmenopausal women,[25] 36% of women with an abnormal pelvic examination result underwent surgery based on the pelvic examination result: 1 woman had ovarian cancer, and 9 had benign conditions.[25] In the PLCO trial, the surgery rate occurring subsequent to an ovarian palpation examination with abnormal findings was 11% over up to 4 rounds of screening (mean, 2.4 ovarian pelvic palpation examinations per woman), with a complication rate of 1%; additional ultrasounds, CA-125 measurements, pelvic examinations, and computed tomography scans occurring subsequent to an ovarian palpation examination with abnormal findings are reported in **Table 5**.

Bacterial Vaginosis

False-positive and false-negative rates for thin homogeneous discharge were 46.1% (95% CI, 38.6%-53.7%) and 21.2% (95% CI, 14.2%-29.7%), respectively.[30] Eschenbach et al[31] reported that false-positive and false-negative rates for homogeneous discharge were 2.8% (95% CI, 1.4%-5.1%) and 30.8% (95% CI, 25.5%-36.6%), respectively (Table 4).

Genital Herpes (HSV-1 or HSV-2)

False-positive and false-negative rates for vulvar ulcerations were 2.5% (95% CI, 1.3%-4.3%) and 80.4% (95% CI, 76.1%-84.2%), respectively, for any stage of genital herpes. False-positive and false-negative rates for tender inguinal lymphadenopathy were 2.9% (95% CI, 1.6%-4.9%) and 85.8% (95% CI, 81.9%-89.0%), respectively, for any stage of genital herpes (Table 4).[32]

Trichomoniasis

False-positive and false-negative rates for colpitis macularis were 0% (95% CI, 0%-0.4%) and 98.3% (95% CI, 94.7%-99.6%), respectively (Table 4).[33] False-positive and false-negative rates for purulent discharge were 23.9% (95% CI, 20.5%-27.6%) and 40.8% (95% CI, 31.5%-50.7%), respectively.

Other Harms

Beyond the downstream diagnostic workup and surgical procedures resulting from the pelvic examination for ovarian cancer screening and the false-positive rates and missed cases from accuracy studies, 1 additional small, poor-to-fair–quality 4-week prospective cohort study with high attrition (49%) reported an association between the pelvic examination and subsequent development of urinary symptoms (dysuria [11/63 vs 6/87; $P < .01$] and urinary frequency [17/63 vs 12/87; $P < .01$]).[35] Further research is needed, in larger studies with urine culture–confirmed urinary tract infection as the outcome, to confirm or disprove this potential harm.

Discussion

An overall summary of the evidence is presented in **Table 6**. There is no direct evidence examining the overall effectiveness of the pelvic examination in improving health outcomes for any of the gynecologic conditions included in this review. Cervical cancer, gonorrhea, and chlamydia were not included because other effective screening tests are already recommended for these conditions.[5,6] Despite the many diseases that are plausibly detectable or that physicians cite as important to detect with routine screening pelvic examinations,[36,37] this systematic review identified diagnostic accuracy studies for only 1 cancerous condition (ovarian cancer) and 3 infectious conditions (bacterial vaginosis, HSV, and trichomoniasis). The limited evidence identified on the accuracy of the screening pelvic examination for detecting any single gynecologic condition and the overall conclusion of its relatively low PPV for ovarian cancer are consistent with other systematic reviews[17,18] and are not surprising, given that ovarian cancer is rare (approximately 12 cases per 100 000 women).[38] Prior cross-sectional studies have shown that even under circumstances optimal for accurate disease detection (ie, patients examined preoperatively under anesthesia, all with some pelvic abnormality as indication for surgery, examination performed by attending physicians), the accuracy of the pelvic examination for detecting pelvic masses is low.[39,40] Performance of the examination is notably worse when performed by inexperienced examiners and in

Table 6. Summary of Evidence by Key Question and Condition

Condition	No. of Studies (Study Design), Sample Size	Summary of Findings (Includes Consistency, Precision)	Applicability	Limitations (Includes Reporting Bias)	Study Quality
KQ1: What Is the Direct Evidence for the Effectiveness of the Pelvic Examination in (a) Reducing All-Cause Mortality,(b) Reducing Cancer- and Disease-Specific Morbidity and Mortality, and (c) Improving Quality of Life?					
All	None (NA)	NA	NA	NA	NA
KQ2: What Are the Test Performance Characteristics of the Pelvic Examination (Sensitivity, Specificity, and Positive and Negative Predictive Values) in Screening for Gynecologic Cancers and Other Gynecologic Conditions?					
Ovarian cancer	4 (3 prospective diagnostic accuracy, 1 RCT) n = 26 432	Sensitivity range: 0%-100% Specificity range: 91%-99%	Fair Average-risk population, low prevalence of ovarian cancer; ultrasound technology from 2 decades ago	Rare ovarian cancer events; accuracy estimates had wide confidence intervals due to very low event rate.	Fair
Bacterial vaginosis	2 (prospective diagnostic accuracy) n = 930	Thin homogeneous discharge: Sensitivity range: 69%-79% Specificity range: 54%-97%	Poor High-risk population; likely overestimates test performance	No screening studies conducted solely in asymptomatic primary care populations; studies had large proportions of symptomatic patients and substantial clinical heterogeneity (populations, personnel performing index test, description of results of index tests, reference standards) and statistical heterogeneity (disparate accuracy results).	Fair
Genital herpes	1 (prospective diagnostic accuracy) n = 779	Vulvar ulcerations: Sensitivity: 20% Specificity: 98%	Poor High-risk population, likely overestimates test performance	No screening studies conducted solely in asymptomatic primary care populations; 78% of women with genital herpes had latent disease that could not be detected by pelvic examination.	Fair
Trichomoniasis	1 (prospective diagnostic accuracy) n = 779	Colpitis macularis: Sensitivity: 2% Specificity: 100% Other signs: Sensitivity range: 8%-59% Specificity range: 76%-99%	Poor High-risk population, likely overestimates test performance	No screening studies in solely asymptomatic primary care populations; data reporting insufficient to calculate accuracy for the presence of ≥1 abnormal finding on pelvic examination.	Fair
KQ3: What Are the Adverse Effects of Screening Pelvic Examination?					
Pelvic examination	1 (prospective cohort) n = 150	Dysuria and urinary frequency ($P < .01$) were more common in the pelvic examination group compared with control group during the 4-week follow-up.	Poor to fair Reflects community practice, single exploratory study	Baseline differences in study groups, high attrition, underpowered to detect urinary tract infection outcome	Poor to fair
Ovarian cancer	4 (3 rospective diagnostic accuracy, 1 RCT) n = 26 432	Surgery rates in participants with abnormal pelvic examination: 5% to 36% Single large study reported complication rate of 1% at 5 y downstream from abnormal pelvic examination. Downstream diagnostic workups include additional ultrasounds (18%), CA-125 measurements (6%), CT scans (2%).	Fair	Few studies, different ultrasound techniques and threshold positivity, and rare ovarian cancer events	Fair
Bacterial vaginosis	2 (prospective diagnostic accuracy) n = 930	Thin, homogeneous discharge: FPR range: 3%-46% FNR range: 21%-31%	Poor	High-risk, symptomatic population	Fair
Genital herpes	1 (prospective diagnostic accuracy) n = 779	Vulvar ulceration: FPR: 2% FNR: 80% Tender lymphadenopathy: FPR: 3% FNR: 86%	Poor	High-risk, symptomatic population	Fair
Trichomoniasis	1 (prospective diagnostic accuracy) n = 779	Colpitis macularis: FPR: 0% FNR: 98% Other individual signs: FPR range: 1%-24% FNR range: 41%-92%	Poor	High-risk, symptomatic population	Fair

Abbreviations: CA-125, cancer antigen 125; CT, computed tomography; FNR, false-negative rate; FPR, false-positive rate; KQ, key question; NA, not applicable; RCT, randomized clinical trial.

patients who are obese or have enlarged uteri.[39-42] To our knowledge, this review is the first systematic review to include accuracy data from the large PLCO trial confirming the very low sensitivity of ovarian palpation for the detection of ovarian cancer and estimating an 11% surgery rate resulting from abnormal examination findings.

The 4 studies reporting diagnostic accuracy for infectious diseases in high-risk settings (STI clinics or populations with high rates of symptoms) likely overestimate the accuracy obtained when pelvic examinations are administered to average-risk, asymptomatic primary care populations. Although the sensitivity of the pelvic examination was generally low for detecting a single condition (eg, ovarian cancer or bacterial vaginosis), on the basis of current evidence it is not possible to estimate the value of the screening pelvic examination to detect any of a host of conditions that might be found before symptomatic clinical presentation. In the absence of trials and test accuracy data on summary performance of screening pelvic examinations, clinical guidance could be informed by epidemiologic evidence to estimate the burden of single conditions, biological plausibility of early detection using the pelvic examination, treatability of the disease at earlier stages, and the presence of alternative, superior screening approaches.

The role of the annual or periodic screening pelvic examination is controversial.[9,43-47] There are concerns that scrutiny of the accuracy of screening pelvic examination is misdirected, as the benefit of other routinely provided physical examination components (eg, heart and lung auscultation screening) in the context of screening is similarly lacking.[44] Rather than viewing the examination as a screening test, this perspective suggests a broader clinical purpose, including to facilitate discussion of sensitive topics that would otherwise not be discussed.[47] On the other hand, the screening pelvic examination can cause anxiety and discomfort and could pose unnecessary barriers to care,[48] particularly in certain subpopulations of women (eg, those with a history of sexual abuse[49-51]).

Women's preventive care is provided by clinicians with diverse training and professional orientations, and the type of clinician can influence the comprehensiveness of primary care.[52-56] In the face of changing clinical practice with respect to cervical cancer screening frequency and new pelvic examination screening recommendations from professional organizations,[14,57] longstanding patterns of women's primary health care delivery may be altered. Research into the benefits or unintended consequences of different screening pelvic examination guidelines, ranging from recommendations against their provision[14] to recommendations for shared decision making, is needed.[15,57] Patients' expectations and preferences for pelvic examination also warrant investigation, as it is not clear how shared decision making may change current clinical screening practices. Regardless of recommendations for screening pelvic examinations, it has been noted that women may wish to have an annual gynecological visit to ensure a full spectrum of preventive care.[56]

Limitations

This review included English-language studies only, consistent with the USPSTF methods.[16] This review excluded conditions such as gonorrhea, chlamydia, and cervical cancer, given that there are more accurate, alternative screening modalities for these conditions. Otherwise, despite having broad search terms to ensure an exhaustive literature search, the review identified little evidence to address the key questions.

Conclusions

No direct evidence was identified for overall benefits and harms of the pelvic examination as a 1-time or periodic screening test. Limited evidence was identified regarding the diagnostic accuracy and harms of routine screening pelvic examinations in asymptomatic primary care populations.

ARTICLE INFORMATION

Author Contributions: Dr Guirguis-Blake had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Funding/Support: This research was funded by the Agency for Healthcare Research and Quality (AHRQ) under a contract to support the USPSTF.

Role of the Funder/Sponsor: Investigators worked with USPSTF members and AHRQ staff to develop the scope, analytic framework, and key questions for this review. AHRQ had no role in study selection, quality assessment, or synthesis. AHRQ staff provided project oversight; reviewed the report to ensure that the analysis met methodological standards; and distributed the draft for peer review. Otherwise, AHRQ had no role in the conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript findings.

The opinions expressed in this document are those of the authors and do not reflect the official position of AHRQ or the US Department of Health and Human Services.

Additional Contributions: We gratefully acknowledge the following individuals for their contributions to this project: Nadia Redmond, MSPH, Smyth Lai, MLS, and Elizabeth Hess, MA (Kaiser Permanente Center for Health Research); Barnett Kramer, MD, Paul Pinsky, PhD, and Maryam Doroudi, PhD, for providing the unpublished data from the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial participants undergoing ovarian palpation; Tina Fan, MD (Agency for Healthcare Research and Quality); and current and former members of the USPSTF who contributed to topic deliberations. USPSTF members, peer reviewers, and federal partner reviewers did not receive financial compensation for their contributions.

Additional Information: A draft version of this evidence report underwent external peer review from 3 content experts (Hanna Bloomfield, MD, Minneapolis VA Health Care, University of Minnesota School of Medicine; David Chelmow, MD, VCU Medical Center; Giang Nguyen, MD, University of Pennsylvania) and 3 federal partners (Military Health System, National Institutes of Health, and Centers for Disease Control and Prevention). Comments were presented to the USPSTF during its deliberation of the evidence and were considered in preparing the final evidence review.

Editorial Disclaimer: This evidence report is presented as a document in support of the accompanying USPSTF Recommendation Statement. It did not undergo additional peer review after submission to *JAMA*.

REFERENCES

1. National Center for Health Statistics. National Ambulatory Medical Care Survey: 2012 summary tables. Centers for Disease Control and Prevention website. http://www.cdc.gov/nchs/data/ahcd/namcs_summary/2012_namcs_web_tables.pdf. Accessed July 28, 2016.

2. Miller JW, Baldwin LM, Matthews B, et al. Physicians' beliefs about effectiveness of cancer screening tests: a national survey of family physicians, general internists, and obstetrician-gynecologists. *Prev Med*. 2014; 69:37-42.

3. Herman CR, Gill HK, Eng J, Fajardo LL. Screening for preclinical disease: test and disease characteristics. *AJR Am J Roentgenol*. 2002; 179(4):825-831.

4. Herman C. What makes a screening exam "good"? *Virtual Mentor*. 2006;8(1):34-37.

5. LeFevre ML; U.S. Preventive Services Task Force. Screening for chlamydia and gonorrhea: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2014;161(12):902-910.

6. Moyer VA; U.S. Preventive Services Task Force. Screening for cervical cancer: U.S. Preventive Services Task Force recommendation statement [published correction appears in *Ann Intern Med*. 2013;158(11):852.]. *Ann Intern Med*. 2012;156(12): 880-891.

7. ACOG Committee on Practice Bulletins—Gynecology. ACOG Practice Bulletin No. 109: cervical cytology screening. *Obstet Gynecol*. 2009; 114(6):1409-1420.

8. Committee on Practice Bulletins—Gynecology. ACOG Practice Bulletin Number 131: screening for cervical cancer. *Obstet Gynecol*. 2012;120(5):1222-1238.

9. Westhoff CL, Jones HE, Guiahi M. Do new guidelines and technology make the routine pelvic examination obsolete? *J Womens Health (Larchmt)*. 2011;20(1):5-10.

10. Sawaya GF. Pelvic examinations in asymptomatic women: tipping a sacred cow. *Arch Intern Med*. 2011;171(22):2054-2055.

11. Sawaya GF, Jacoby V. Screening pelvic examinations: right, wrong, or rite? *Ann Intern Med*. 2014;161(1):78-79.

12. Burns RB, Potter JE, Ricciotti HA, Reynolds EE. Screening pelvic examinations in adult women: grand rounds discussion from the Beth Israel Deaconess Medical Center. *Ann Intern Med*. 2015; 163(7):537-547.

13. American College of Obstetricians and Gynecologists (ACOG). Committee Opinion: well-woman visit. ACOG website. http://www.acog.org/Resources-And-Publications/Committee-Opinions/Committee-on-Gynecologic-Practice/Well-Woman-Visit. 2012. Accessed August 21, 2014.

14. Qaseem A, Humphrey LL, Harris R, Starkey M, Denberg TD; Clinical Guidelines Committee of the American College of Physicians. Screening pelvic examination in adult women: a clinical practice guideline from the American College of Physicians. *Ann Intern Med*. 2014;161(1):67-72.

15. American College of Obstetricians and Gynecologists (ACOG). ACOG practice advisory on annual pelvic examination recommendations. ACOG website. http://www.acog.org/About-ACOG/News-Room/Practice-Advisories/ACOG-Practice-Advisory-on-Annual-Pelvic-Examination-Recommendations. 2014. Accessed November 5, 2015.

16. Harris RP, Helfand M, Woolf SH, et al; Methods Work Group, Third US Preventive Services Task Force. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med*. 2001;20(3)(suppl):21-35.

17. Bloomfield HE, Olson A, Greer N, et al. Screening pelvic examinations in asymptomatic, average-risk adult women: an evidence report for a clinical practice guideline from the American College of Physicians. *Ann Intern Med*. 2014;161(1): 46-53.

18. Ebell MH, Culp M, Lastinger K, Dasigi T. A systematic review of the bimanual examination as a test for ovarian cancer. *Am J Prev Med*. 2015;48 (3):350-356.

19. Doroudi M, Kramer BS, Pinsky PF. The bimanual ovarian palpation examination in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial: performance and complications [published online November 30, 2016]. *J Med Screen*. doi:10.1177/0969141316680381

20. Wallace BC, Small K, Brodley CE, Lau J, Trikalinos TA. Deploying an interactive machine learning system in an Evidence-based Practice Center. *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*; January 28-30, 2012; Miami, FL. doi:10.1145/2110363.2110464.

21. United Nations Development Programme (UNDP). 2014 Human Development Report. UNDP website. http://www.undp.org/content/undp/en/home/presscenter/events/2014/july/HDR2014.html. 2014. Accessed April 13, 2016.

22. Wells GA, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analysis. The Ottawa Hospital website. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. 2014. Accessed November 9, 2015.

23. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003;3:25.

24. 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(8):529-536.

25. Jacobs I, Stabile I, Bridges J, et al. Multimodal approach to screening for ovarian cancer. *Lancet*. 1988;1(8580):268-271.

26. Adonakis GL, Paraskevaidis E, Tsiga S, Seferiadis K, Lolis DE. A combined approach for the early detection of ovarian cancer in asymptomatic women. *Eur J Obstet Gynecol Reprod Biol*. 1996;65 (2):221-225.

27. Grover S, Quinn MA, Weideman P, et al. Screening for ovarian cancer using serum CA125 and vaginal examination: report on 2550 females. *Int J Gynecol Cancer*. 1995;5(4):291-295.

28. Prorok PC, Andriole GL, Bresalier RS, et al; Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000;21 (6)(suppl):273S-309S.

29. Buys SS, Partridge E, Greene MH, et al; PLCO Project Team. Ovarian cancer screening in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial: findings from the initial screen of a randomized trial [published correction appears in *Am J Obstet Gynecol*. 2005;193(6):2183-2184]. *Am J Obstet Gynecol*. 2005;193(5):1630-1639.

30. Gutman RE, Peipert JF, Weitzen S, Blume J. Evaluation of clinical methods for diagnosing bacterial vaginosis. *Obstet Gynecol*. 2005;105(3): 551-556.

31. Eschenbach DA, Hillier S, Critchlow C, Stevens C, DeRouen T, Holmes KK. Diagnosis and clinical manifestations of bacterial vaginosis. *Am J Obstet Gynecol*. 1988;158(4):819-828.

32. Koutsky LA, Stevens CE, Holmes KK, et al. Underdiagnosis of genital herpes by current clinical and viral-isolation procedures. *N Engl J Med*. 1992; 326(23):1533-1539.

33. Wølner-Hanssen P, Krieger JN, Stevens CE, et al. Clinical manifestations of vaginal trichomoniasis. *JAMA*. 1989;261(4):571-576.

34. Spiegel CA, Amsel R, Holmes KK. Diagnosis of bacterial vaginosis by direct gram stain of vaginal fluid. *J Clin Microbiol*. 1983;18(1):170-177.

35. Tiemstra JD, Pela E. Urinary symptoms after a routine pelvic exam. *J Am Board Fam Med*. 2011;24 (3):290-295.

36. Stormo AR, Cooper CP, Hawkins NA, Saraiya M. Physician characteristics and beliefs associated with use of pelvic examinations in asymptomatic women. *Prev Med*. 2012;54(6):415-421.

37. Henderson JT, Harper CC, Gutin S, Saraiya M, Chapman J, Sawaya GF. Routine bimanual pelvic examinations: practices and beliefs of US obstetrician-gynecologists. *Am J Obstet Gynecol*. 2013;208(2):109.e1-109.e7.

38. Howlader N, Noone AM, Krapcho M, et al, eds. *SEER Cancer Statistics Review, 1975–2012*. Bethesda, MD: National Cancer Institute; 2015.

39. Padilla LA, Radosevich DM, Milad MP. Limitations of the pelvic examination for evaluation of the female pelvic organs. *Int J Gynaecol Obstet*. 2005;88(1):84-88.

40. Padilla LA, Radosevich DM, Milad MP. Accuracy of the pelvic examination in detecting adnexal masses. *Obstet Gynecol*. 2000;96(4):593-598.

41. Tayal VS, Crean CA, Norton HJ, Schulz CJ, Bacalis KN, Bliss S. Prospective comparative trial of endovaginal sonographic bimanual examination versus traditional digital bimanual examination in nonpregnant women with lower abdominal pain with regard to body mass index classification. *J Ultrasound Med*. 2008;27(8):1171-1177.

42. Ueland FR, Depriest PD, Desimone CP, et al. The accuracy of examination under anesthesia and transvaginal sonography in evaluating ovarian size. *Gynecol Oncol*. 2005;99(2):400-403.

43. Saraiya M. Evidence is lacking to support pelvic examinations as a screening tool for non-cervical cancers or other conditions. *Evid Based Med*. 2015; 20(1):31.

44. Bump RC. "Pelvic: deferred"—have nongynecologists been right all along? *Am J Obstet Gynecol*. 2013;208(2):91.

45. Bates C, Potter J. Screening pelvic examination in adult women. *Ann Intern Med*. 2014;161(12):925-926.

46. Metz S, Gussman D. Screening pelvic examination in adult women. *Ann Intern Med*. 2014; 161(12):925.

47. Jennings JC, Blake J. Screening pelvic examinations in asymptomatic, average-risk adult women. *Ann Intern Med*. 2014;161(12):924.

48. Harper C, Balistreri E, Boggess J, Leon K, Darney P. Provision of hormonal contraceptives without a mandatory pelvic examination: the First Stop demonstration project. *Fam Plann Perspect*. 2001;33(1):13-18.

49. Swahnberg K, Wijma B, Siwe K. Strong discomfort during vaginal examination: why consider a history of abuse? *Eur J Obstet Gynecol Reprod Biol*. 2011;157(2):200-205.

50. Weitlauf JC, Finney JW, Ruzek JI, et al. Distress and pain during pelvic examinations: effect of sexual violence. *Obstet Gynecol*. 2008;112(6):1343-1350.

51. Weitlauf JC, Frayne SM, Finney JW, et al. Sexual violence, posttraumatic stress disorder, and the pelvic examination: how do beliefs about the safety, necessity, and utility of the examination influence patient experiences? *J Womens Health (Larchmt)*. 2010;19(7):1271-1280.

52. Henderson JT, Weisman CS, Grason H. Are two doctors better than one? women's physician use and appropriate care. *Womens Health Issues*. 2002; 12(3):138-149.

53. Henderson JT, Weisman CS. Women's patterns of provider use across the lifespan and satisfaction with primary care coordination and comprehensiveness. *Med Care*. 2005;43(8):826-833.

54. Stormo AR, Saraiya M, Hing E, Henderson JT, Sawaya GF. Women's clinical preventive services in the United States: who is doing what? *JAMA Intern Med*. 2014;174(9):1512-1514.

55. Clancy CM, Massion CT. American women's health care: a patchwork quilt with gaps. *JAMA*. 1992;268(14):1918-1920.

56. Institute of Medicine. *Clinical Preventive Services for Women: Closing the Gaps*. Washington, DC: Institute of Medicine; 2011.

57. Conry JA, Brown H. Well-Woman Task Force: components of the well-woman visit. *Obstet Gynecol*. 2015;126(4):697-701.