Aetna considers transvaginal ultrasonography (TV-US) medically necessary for a number of indications:

- Assessment of a pelvic mass (e.g., adenomyosis, cancer, cyst, and fibroid)
- Diagnosis of bowel endometriosis
- Diagnosis of vasa previa
- Evaluation of abnormal uterine bleeding
- Evaluation of congenital uterine anomalies
- Evaluation of infertility (see [CPB 0327 - Infertility](../300_399/0327.html))
- Evaluation of sequelae of pelvic infection (e.g., abscess, and hydrosalpinx)
- Evaluation of women with new symptoms (bloating, difficulty eating or feeling full quickly, pelvic or abdominal pain, or urinary frequency and urgency) that have persisted for 3 or more weeks, and the clinician has performed a pelvic and rectal examination and suspects ovarian cancer
- Evaluation of women with post-menopausal bleeding
- Guidance during embryo transfer
- Monitoring natural or stimulated follicular development
during infertility therapy (see CPB 0327 - Infertility (../300_399/0327.html)

- Monitoring of women with Lynch II syndrome (BRCA2 mutation) for ovarian cancer. (Note: The Doppler ultrasound mode is considered not medically necessary for TV-US monitoring of women with Lynch II syndrome)
- Verifying position of intrauterine device if IUD string is not visible or if there is a suspicion that IUD is not in the correct position

Aetna considers TV-US experimental and investigational for screening for ovarian cancer, endometrial cancer, or other gynecologic cancers because of insufficient evidence of effectiveness for screening.

**Background**

Pelvic ultrasound is considered to be clinically integral to the transvaginal examination and does not warrant separate reimbursement. A transvaginal ultrasound (TV-US) provides superior detail in images of pelvic structures. When TV-US is performed on a patient whose pelvic structures are within the bony pelvis, pelvic echography using an abdominal approach is duplicative of the TV-US.

Persadie (2002) stated that measurement of endometrial thickness with ultrasonography is a modality commonly used today. Its clinical importance and applications extend throughout the phases of the reproductive lives of women. In pre-menopausal women, endometrial thickness is used to monitor infertility treatment, while in post-menopausal women with abnormal uterine bleeding it is useful as an initial investigation for endometrial hyperplasia or cancer. Moreover, endometrial thickness can vary with the menstrual cycle and with the use of hormone replacement therapy or selective estrogen receptor modulators.

Champaneria et al (2010) stated that adenomyosis is a common condition that causes substantial morbidity. Until recently, the reference standard for a definitive diagnosis was histology of
hysterectomy specimens. Ultrasound and magnetic resonance imaging (MRI) may allow accurate non-invasive diagnosis. In a systematic review with meta-analysis, these investigators compared the diagnostic accuracy of these techniques. Subjects were women who had ultrasound and/or MRI, and whose results were compared with a reference standard. Electronic searches were conducted in literature databases from database inception to 2010. The reference lists of known relevant articles were searched for further articles. Selected studies reported data on ultrasound and/or MRI with histological confirmation of diagnosis. Two reviewers independently selected articles without language restrictions, and extracted data in the form of 2 × 2 tables. They computed sensitivity and specificity for individual studies and pooled these results in a meta-analysis. They also performed meta-regression to examine how the index tests compared on diagnostic accuracy. A total of 23 articles (involving 2,312 women) satisfied the inclusion criteria. Transvaginal ultrasound had a pooled sensitivity of 72 % (95 % confidence intervals [CI]: 65 to 79 %), specificity of 81 % (95 % CI: 77 to 85 %), positive likelihood ratio of 3.7 (95 % CI: 2.1 to 6.4) and negative likelihood ratio of 0.3 (95 % CI: 0.1 to 0.5); MRI had a pooled sensitivity of 77 % (95 % CI: 67 to 85 %), specificity of 89 % (95 % CI: 84 to 92 %), positive likelihood ratio of 6.5 (95 % CI: 4.5 to 9.3), and negative likelihood ratio of 0.2 (95 % CI: 0.1 to 0.4). The results showed that a correct diagnosis was obtained more often with MRI. The authors concluded that transvaginal ultrasound and MRI showed high levels of accuracy for the non-invasive diagnosis of adenomyosis.

The ACOG practice bulletin’s on “Diagnosis of abnormal uterine bleeding in reproductive-aged women” (2012) stated that some experts recommend transvaginal ultrasonography as the initial screening test for abnormal uterine bleeding and MRI as a second-line test to be used when the diagnosis is inconclusive, when further delineation would affect patient management, or when co-existing uterine myomas are suspected.

An UpToDate review on “Ultrasound examination in obstetrics
“and gynecology” (Shipp, 2013) states that gynecologic ultrasound examination has multiple uses, including but not limited to:

- Evaluation of the menstrual cycle (endometrial thickness, follicular development)
- Monitoring natural or stimulated follicular development during infertility therapy
- Localization of an intrauterine device
- Evaluation of abnormal uterine bleeding
- Assessment of a pelvic mass (e.g., adenomyosis, fibroid, cancer, cysts)
- Evaluation for sequelae of pelvic infection (e.g., abscess, hydrosalpinx)
- Evaluation of congenital uterine anomalies
- Screening for malignancy

**Screening for Ovarian Cancer:**

Ovarian cancer is among the deadliest types of cancer because diagnosis usually comes very late, after the cancer has spread. If the cancer is found and surgically removed before it spreads outside the ovary, the 5-year survival rate is 93%. However, only 19% of cases are detected early enough for that kind of successful intervention. It is estimated that 22,430 new cases and 15,280 deaths will be reported in 2007 (ACS, 2007).

A Committee Opinion by the American College of Obstetricians and Gynecologists has concluded that TV-US has not been proven as a screening test for ovarian cancer (ACOG, 2002). The National Cancer Institute (NCI, 2004) has stated that there is insufficient evidence to establish that screening for ovarian cancer with TV-US would result in a decrease in mortality from ovarian cancer. The NCI notes that a serious potential harm is the false-positive test result, which may lead to anxiety and invasive diagnostic procedures. The NCI states that there is good evidence that screening for ovarian cancer with TV-US would result in more diagnostic laparoscopies and laparotomies than new ovarian cancers found. Unnecessary oophorectomies
Transvaginal ultrasound may be medically necessary for monitoring women with Lynch II syndrome (BRCA2 mutation) for ovarian cancer. Use of the Doppler mode, however, is not medically necessary for this indication. Although transvaginal Doppler ultrasonography (TV-DUS) may improve upon the ability of other imaging methods (i.e., TV-US) in distinguishing benign from malignant ovarian neoplasms, it is not clear whether this improvement will have any impact on the management of patients with adnexal lesions. Specifically, it is unknown whether the diagnostic abilities of TV-DUS are sufficient to confidently identify those patients who can forego surgery and be followed conservatively. Furthermore, the diagnostic abilities of TV-DUS are probably different among pre- and post-menopausal patients due to the differing prevalence of malignancy between these 2 groups. Unfortunately, studies of TV-DUS have included a mixture of pre- and post-menopausal patients. It is unclear how TV-DUS will alter patient management in those patients with ovarian masses.

Fields and Chevlen (2006) stated that currently available tests (CA-125, TV-US, or a combination of both) lack the sensitivity and specificity to be useful for screening ovarian cancer in the general population.

Lacey and colleagues (2006) examined the positive predictive values of CA-125 or TV-US screening for ovarian cancer according to family history of breast or ovarian cancer. In the screening arm of a randomized controlled trial of screening compared with usual care, a total of 28,460 women with family history data received baseline and annual CA-125 and TV-US examinations. These investigators analyzed CA-125 and TV-US results from the first 4 rounds of screening. They classified women as average (n = 22,687), moderate (n = 2,572), or high (n = 2,163) risk based on family history, or high risk due to a personal history of breast cancer (n = 1,038). Cancers were identified by active follow-up of women with abnormal screening results and annual questionnaires. These researchers
calculated positive predictive values for screening combinations. Similar proportions (4.8 to 5.0 %) of women in each group had abnormal screening results. Higher-risk women were more likely than lower-risk women to undergo biopsy after a positive screen. Screening identified 43 invasive ovarian cancers. The positive predictive values for abnormal screening results were 0.7 % in average-risk, 1.3 % in moderate-risk, and 1.6 % in high-risk groups; 1 ovarian cancer occurred among the breast cancer survivors. The positive predictive values for post-baseline abnormal screening results were also higher in the higher-risk groups. The positive-predictive values did not significantly differ across risk groups. The authors concluded that the probabilities of abnormal annual CA-125 and TV-US screens were similar across groups based on family history of breast or ovarian cancer. However, ovarian cancer was more likely to be diagnosed after an abnormal screening result among women at higher family history-based risk than among women at lower risk.

The authors noted that ongoing studies, including the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, will ascertain the efficacy of ovarian cancer screening. Until the results of these studies are available, the findings of this analysis demonstrated that stratifying women into risk groups based on family history slightly enhanced the positive predictive values of a combined CA-125 and TV-US-based screening approach. These researchers stated that whether these differences prove to be efficacious, cost-effective, or clinically useful in screened populations awaits the results of the PLCO and other cancer screening studies.

The Gynecologic Cancer Foundation, the Society of Gynecologic Oncologists, and the American Cancer Society have issued a consensus statement to promote early detection of ovarian cancer, which recommends that women who have symptoms -- specifically bloating, pelvic or abdominal pain, difficulty eating or feeling full quickly, and urinary frequency and urgency -- are urged to see a gynecologist if symptoms are new and persist for more than 3 weeks (ACS, 2007; SGO, 2007). The consensus
Statement recommendations are based on studies that show the above symptoms appeared in women with ovarian cancer more than in other women (Goff et al, 2004; Daly and Ozols, 2004). The recommendations acknowledge that there is not consensus on what physicians should do when patients present with these symptoms. According to a consensus statement issued by the Gynecologic Cancer Foundation, patients should be evaluated with pelvic and rectal examinations. If there is a suspicion of cancer, the clinician may choose to perform a TV-US to check the ovaries for signs of cancer. Testing for CA-125 levels should also be considered.

Jokubkiene et al (2007) examined if tumor vascularity as assessed by 3-dimensional (3D) power Doppler ultrasound can be used to discriminate between benign and malignant ovarian tumors, if adding 3D power Doppler ultrasound to gray-scale imaging improves differentiation between benignity and malignancy, and if 3D power Doppler ultrasound adds more to gray-scale ultrasound than does 2-dimensional (2D) power Doppler ultrasound. A total of 106 women scheduled for surgery because of an ovarian mass were examined with transvaginal gray-scale ultrasound as well as 2D and 3D power Doppler ultrasound. The color content of the tumor scan was rated subjectively by the ultrasound examiner on a visual analog scale. Vascularization index (VI), flow index (FI) and vascularization flow index (VFI) were calculated in the whole tumor and in a 5-cm(3) sample taken from the most vascularized area of the tumor. Logistic regression analysis was used to build models to predict malignancy. There were 79 benign tumors, 6 borderline tumors and 21 invasive malignancies. A logistic regression model including only gray-scale ultrasound variables (the size of the largest solid component, wall irregularity, and lesion size) was built to predict malignancy. It had an area under the receiver-operating characteristics (ROC) curve of 0.98, sensitivity of 100 %, false-positive rate of 10 %, and positive likelihood ratio (LR) of 10 when using the mathematically best cut-off value for risk of malignancy (0.12). The diagnostic performance of the 3D flow index with the best diagnostic performance, i.e., VI in a 5-cm(3)
sample, was superior to that of the color content of the tumor scan (area under ROC curve 0.92 versus 0.80, sensitivity 93 % versus 78 %, false-positive rate 16 % versus 27 % using the mathematically best cut-off value). Adding the color content of the tumor scan or F1 in a 5-cm(3) sample to the logistic regression model including the 3 gray-scale variables described above improved diagnostic performance only marginally, an additional 2 tumors being correctly classified. The authors concluded that even though 2D and 3D power Doppler ultrasound can be used to discriminate between benign and malignant ovarian tumors, their use adds little to a correct diagnosis of malignancy in an ordinary population of ovarian tumors. Objective quantitation of the color content of the tumor scan using 3D power Doppler ultrasound does not seem to add more to gray-scale imaging than does subjective quantitation by the ultrasound examiner using 2D power Doppler ultrasound.

Partridge and associates (2009) examined if annual screening with TV-US and CA-125 reduces ovarian cancer mortality. Data from the first 4 annual screens, denoted T0-T3, were reported. A CA-125 value at or above 35 units/ml or an abnormality on TV-US was considered a positive screen. Diagnostic follow-up of positive screens was performed at the discretion of participants' physicians. Diagnostic procedures and cancers were tracked and verified through medical records. Among 34,261 screening arm women without prior oophorectomy, compliance with screening ranged from 83.1 % (T0) to 77.6 % (T3). Screen positivity rates declined slightly with TV-US, from 4.6 at T0 to 2.9 to 3.4 at T1-T3; CA-125 positivity rates (range of 1.4 % to 1.8 %) showed no time trend. A total of 89 invasive ovarian or peritoneal cancers were diagnosed; 60 were screen-detected. The positive-predictive value (PPV) and cancer yield per 10,000 women screened on the combination of tests were similar across screening rounds (range of 1.0 % to 1.3 % for PPV and 4.7 to 6.2 for yield); however, the biopsy (surgery) rate among screen positives decreased from 34 % at T0 to 15 % to 20 % at T1-T3. The overall ratio of surgeries to screen-detected cancers was 19.5:1. A total of 72 % of screen-detected cases
were late stage (III/IV). The authors concluded that through 4 screening rounds, the ratio of surgeries to screen-detected cancers was high, and most cases were late stage. However, the effect of screening on mortality is as yet unknown.

Menon and colleagues (2009) noted that the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) is a randomized controlled trial designed to assess the effect of screening on mortality. These investigators summarized the outcome of the prevalence (initial) screen in UKCTOCS. Between 2001 and 2005, a total of 202,638 post-menopausal women aged 50 to 74 years were randomly assigned to no treatment (control; n = 101,359); annual CA-125 screening (interpreted using a risk of ovarian cancer algorithm) with TV-US scan as a second-line test (multi-modal screening [MMS]; n = 50,640); or annual screening with TV-US (USS; n = 50,639) alone in a 2:1:1 ratio using a computer-generated random number algorithm. All women provided a blood sample at recruitment. Women randomized to the MMS group had their blood tested for CA-125 and those randomized to the USS group were sent an appointment to attend for a TV-US. Women with abnormal screens had repeat tests. Women with persistent abnormality on repeat screens underwent clinical evaluation and, where appropriate, surgery. In the prevalence screen, 50,078 (98.9 %) women underwent MMS, and 48,230 (95.2 %) underwent USS. The main reasons for withdrawal were death (2 MMS, 28 USS), non-ovarian cancer or other disease (none MMS, 66 USS), removal of ovaries (5 MMS, 29 USS), relocation (none MMS, 39 USS), failure to attend 3 appointments for the screen (72 MMS, 757 USS), and participant changing their mind (483 MMS, 1,490 USS). Overall, 4,355 of 50,078 (8.7 %) women in the MMS group and 5,779 of 48,230 (12.0 %) women in the USS group required a repeat test, and 167 (0.3 %) women in the MMS group and 1,894 (3.9 %) women in the USS group required clinical evaluation. A total of 97 of 50,078 (0.2 %) women from the MMS group and 845 of 48,230 (1.8 %) from the USS group underwent surgery; 42 (MMS) and 45 (USS) primary ovarian and tubal cancers were detected, including 28 borderline tumors (8 MMS, 20 USS).
(16 MMS, 12 USS) of 58 (48.3 %; 95 % confidence interval [CI]: 35.0 to 61.8) of the invasive cancers were stage I/II, with no difference (p = 0.396) in stage distribution between the groups. A further 13 (5 MMS, 8 USS) women developed primary ovarian cancer during the year after the screen. The sensitivity, specificity, and PPV for all primary ovarian and tubal cancers were 89.4 %, 99.8 %, and 43.3 % for MMS, and 84.9 %, 98.2 %, and 5.3 % for USS, respectively. For primary invasive epithelial ovarian and tubal cancers, the sensitivity, specificity, and PPV were 89.5 %, 99.8 %, and 35.1 % for MMS, and 75.0 %, 98.2 %, and 2.8 % for USS, respectively. There was a significant difference in specificity (p < 0.0001) but not sensitivity between the 2 screening groups for both primary ovarian and tubal cancers as well as primary epithelial invasive ovarian and tubal cancers. The authors concluded that the sensitivity of the MMS and USS screening strategies is encouraging. Specificity was higher in the MMS than in the USS group, resulting in lower rates of repeat testing and surgery. This in part reflects the high prevalence of benign adnexal abnormalities and the more frequent detection of borderline tumors in the USS group. The prevalence screen has established that the screening strategies are feasible. The results of ongoing screening are awaited so that the effect of screening on mortality can be determined.

Nelson and colleagues (2009) stated that the National Breast and Ovarian Cancer Center’s position statement on population screening and early detection of ovarian cancer in asymptomatic women was developed and agreed following a Forum in February 2009 attended by key Australian stakeholders. The final position statement and supporting background information have been endorsed by key Australian colleges and agencies. Position statement on population screening and early detection of ovarian cancer in asymptomatic women noted that (i) currently there is no evidence that any test, including pelvic examination, CA-125 or other biomarkers, ultrasound (including TV-US), or combination of tests, results in reduced mortality from ovarian cancer, and (ii) there is no evidence to support the use of any test, including pelvic examination, CA-125 or other biomarkers, ultrasound
(including TV-US), or combination of tests, for routine population-based screening for ovarian cancer.

The Royal Australian College of General Practitioners' guidelines on “Preventive activities in general practice” (2012) stated that “There is no evidence to support the use of any test -- including pelvic examination, CA125, or other biomarkers, ultrasound (including transvaginal ultrasound), or combination of tests -- for routine population-based screening for ovarian cancer .... The specificity of transvaginal ultrasound is low. The low prevalence of ovarian cancer means that even screening tests that have very high sensitivity and specificity have a low positive predictive value for disease detection”.

Screening for Endometrial Cancer:

Montgomery et al (2004) noted that endometrial hyperplasia is a precursor to the most common gynecological cancer diagnosed in women: endometrial cancer of endometrioid histology. It is most often diagnosed in post-menopausal women, but women at any age with unopposed estrogen from any source are at an increased risk for developing endometrial hyperplasia. Hyperplasia with cytological atypia represents the greatest risk for progression to endometrial carcinoma and the presence of concomitant carcinoma in women with endometrial hyperplasia. Abnormal uterine bleeding is the most common presenting symptom of endometrial hyperplasia. Specific Pap smear findings and endometrial thickness per ultrasound could also suggest the diagnosis. Epstein and Valentin (2004) stated that a measurement of endometrial thickness is a simple and accurate method for estimating the risk of endometrial cancer. However, the reliability of ultrasound evaluation of endometrial morphology and/or vascularization for risk estimation of endometrial malignancy remains to be determined.

In addressing whether TV-US should be performed at annual examination in asymptomatic women, Cohen (2003) stated that there is little evidence that death rates from either endometrial
or ovarian cancer would improve with this approach. If TV-US is to be used in screening asymptomatic women, it should be as part of a controlled study and at no cost to the patient. This is in agreement with Robertson (2003) who stated that routine screening for endometrial carcinoma is currently not justified. Post-menopausal women need to be educated about the importance of seeking attention if any vaginal bleeding occurs. All post-menopausal bleeding requires review and appropriate investigation. Women taking tamoxifen have a higher risk of endometrial cancer and should report any bleeding or spotting; however, ultrasound screening is not recommended for asymptomatic women taking tamoxifen. Families with hereditary non-polyposis colon cancer have a higher risk of endometrial cancer and require counseling about this risk. A Pap test is not a screening test for endometrial cancer, but the incidental finding of endometrial cells on a Pap smear in a post-menopausal woman requires investigation.

The National Cancer Institute (NCI, 2004) has stated that there is insufficient evidence to establish whether a decrease in mortality from endometrial cancer occurs with screening by TV-US. The NCI notes that risks associated with false-positive test results include anxiety and additional diagnostic testing and surgery. In addition, endometrial cancers may be missed by ultrasound.

Meyer et al (2009) stated that about 2 % to 5 % of endometrial cancers may be due to an inherited susceptibility. Lynch syndrome (also known as hereditary non-polyposis colorectal cancer syndrome), an autosomal-dominant inherited cancer susceptibility syndrome caused by a germline mutation in one of the DNA mismatch repair genes, accounts for the majority of inherited cases. Lynch syndrome is associated with early onset of cancer and the development of multiple cancer types, especially colon and endometrial cancer. These researchers reviewed the current status of knowledge regarding Lynch syndrome-associated endometrial cancer and methods for diagnosis, screening, and prevention of cancers. The lifetime cumulative risk of endometrial cancer for women with
Lynch syndrome is 40% to 60%, which equals or exceeds their risk of colorectal cancer. No current evidence suggests either a survival advantage or disadvantage to endometrial cancer that is associated with Lynch syndrome when these cases are compared with sporadic cases. A combination of family and personal medical history and tumor testing provides an efficient basis for diagnosing Lynch syndrome in women with endometrial cancer. Current gynecologic cancer screening guidelines for women with Lynch syndrome include annual endometrial sampling and TV-US beginning at age 30 to 35 years. The authors concluded that diagnosing endometrial cancer patients with Lynch syndrome has important clinical implications for the individual and family members. Screening and prevention practices can decrease the likelihood of developing additional cancers.

**Screening for Other Gynecological Cancers:**

Sharma and Menon (2006) noted that the role of screening in gynecological cancers is being studied. With mass screening proven effective in significantly lowering cervical cancer mortality, there is an intense interest in developing other screening methods to detect gynecological cancers early. These researchers reviewed advances in cervical cancer screening, strategies being investigated in ovarian cancer screening and the lack of justification in screening for endometrial, vaginal and vulvar cancers. A Medline-based literature search was performed for studies relating to screening for different gynecological malignancies. Additional papers cited in those identified by the initial search were also reviewed. Advances in cervical cancer screening include liquid-based cytology and human papillomavirus testing. Results of ongoing trials are awaited before these can be fully implemented. The results of the 2 large, multi-center, randomized controlled trials being conducted in the United Kingdom and United States (UKCTOCS and PLCO study, respectively) to evaluate the impact of screening on ovarian cancer mortality will shed some light on the need to implement screening for ovarian cancer in the general population. Novel markers, serum proteomic profiles
and Doppler ultrasonography are some of the other technologies being examined. Presently, screening for endometrial cancer is not advocated since most women present with symptoms in early disease exhibit good survival outcomes. Vaginal and vulval cancers are too rare to justify mass screening. The authors concluded that methods to screen for various gynecological malignancies need further evaluation before implementation in the general population; results of large multi-center trials are awaited. They stated that currently, screening for endometrial, vaginal and vulval cancers is not justified.

**Evaluation of Post-Menopausal Bleeding:**

The ACOG's Committee Opinion on the role of TV-US in the evaluation of women with post-menopausal bleeding (2009) stated that the clinical approach to post-menopausal bleeding requires prompt and efficient evaluation to exclude or diagnose carcinoma. Women with post-menopausal bleeding may be evaluated initially with either endometrial biopsy or TV-US; this initial evaluation does not require performance of both tests. Transvaginal ultrasonography can be useful in the triage of patients in whom endometrial sampling was performed but tissue was insufficient for diagnosis. When TV-US is performed for patients with post-menopausal bleeding and an endometrial thickness of less than or equal to 4 mm is found, endometrial sampling is not required. Meaningful assessment of the endometrium by ultrasonography is not possible in all patients. In such cases, alternative assessment should be completed. When bleeding persists despite negative initial evaluations, additional assessment usually is indicated.

**Confirmation of Placement of an Intra-Uterine Device Following Insertion:**

In a prospective comparative study, de Kroon and colleagues (2003) evaluated the accuracy of clinical evaluation of intra-uterine device (IUD) position. The clinical evaluation was compared with the TV-US measurement of IUD position both
The authors concluded that clinical evaluation is an excellent test for the evaluation of the position of an IUD and routine TV-US is not indicated for this purpose.

Diagnosis of Bowel Endometriosis:

Bazot et al (2007) compared the accuracy of TV-US and rectal endoscopic sonography (RES) for the diagnosis of deep infiltrating endometriosis (DIE), with respect to surgical and histological findings. This was a longitudinal study of 81 consecutive patients referred for surgical management of DIE, who underwent both TV-US and RES pre-operatively. The diagnostic criteria were identical for TV-US and RES, and were based on visualization of hypoechoic areas in specific locations (utero-sacral ligaments, vagina, recto-vaginal septum and intestine). These investigators calculated the sensitivity, specificity, predictive values and accuracy of TV-US and RES for the diagnosis of DIE. Endometriosis was confirmed histologically in 80/81 (98.7 %) patients. Endometriomas and DIE were present in 43.2 % and 97.5 % of the women, respectively. For the diagnosis of DIE overall, TV-US and RES, respectively, had a sensitivity of 87.3 % and 74.7 %, a PPV of 98.6 % and 98.3 %, and an accuracy of 86.4 % and 74 %. For the diagnosis of utero-sacral endometriosis, they had a sensitivity of 80.8 % and 46.6 %, a specificity of 75 % and 50.0 %, a PPV of 96.7 % and 89.5 % and a NPV of 30 % and 9.3 %. For the diagnosis of intestinal endometriosis, they had a
sensitivity of 92.6 % and 88.9 %, a specificity of 100 % and 92.6 %, a PPV of 100 % and 96 % and a NPV of 87 % and 80.6 %. The authors concluded that TV-US is apparently more accurate than RES for predicting DIE in specific locations, and should thus be the first-line imaging technique in this setting.

Hudelist and Keckstein (2009) noted that over the past years, additional diagnostic tools such as TV-US and/or magnetic resonance imaging have been recommended as an appropriate investigation to diagnose ovarian endometriomas or adenomyosis. Several lines of recent evidence strongly suggests that the use of TV-US also has an important role in detecting DIE of the pelvis not only involving the ovaries but also structures such as the vagina, the recto-vaginal space, the utero-sacral ligaments, the bladder or the rectal wall.

Hudelist et al (2011) analyzed the diagnostic value of TV-US for non-invasive, pre-surgical detection of bowel endometriosis. MEDLINE (1966 to 2010) and EMBASE (1980 to 2010) databases were searched for relevant studies investigating the diagnostic accuracy of TV-US for diagnosing deep infiltrating endometriosis involving the bowel. Diagnosis was established by laparoscopy and/or histopathological analysis. Likelihood ratios (LRs) were re-calculated in addition to traditional measures of effectiveness. Out of 188 papers, a total of 10 studies fulfilled pre-defined inclusion criteria involving 1,106 patients with suspected endometriosis. The prevalence of bowel endometriosis varied from 24 to 73.3 %. LR+ ranged from 4.8 to 48.56 and LR- ranged from 0.02 to 0.36, with wide confidence intervals. Pooled estimates of sensitivities and specificities were 91 and 98 %; LR+ and LR- were 30.36 and 0.09; and PPV and NPV were 98 and 95 %, respectively. Three of the studies used bowel preparations to enhance the visibility of the rectal wall; 1 study directly compared the use of water contrast versus no prior bowel enema, for which the LR- was 1.4 and 0.47, respectively. The authors concluded that TV-US with or without the use of prior bowel preparation is an accurate test for non-invasive, pre-surgical detection of DIE of the rectosigmoid.
Egekvist and colleagues (2012) evaluated the inter-observer variation of transvaginal ultrasonographic measurements of the size of DIE lesions in the recto-sigmoid wall performed by an experienced and a less experienced sonographer. Fifteen outpatient women were seen for a gynecologic examination and 24 women were seen with recto-sigmoid endometriosis. Transvaginal ultrasonography was performed independently by 2 observers with a focus on the presence and size of recto-sigmoid endometriosis. The senior observer had several years of experience, whereas the junior observer was a medical student with no prior experience in TV-US. The findings of the 2 observers were identical concerning the identification of recto-sigmoid endometriosis. The probability of differences in size within 30% of the mean was 0.81, 0.63 and 0.61 for length, width and depth, respectively. The authors concluded that these findings suggested that fair skills in this technique can be acquired by inexperienced examiners within a short period of time.

Cervical Assessment for Prevention of Preterm Birth:

The Society of Obstetricians and Gynaecologists of Canada stated that routine transvaginal cervical length assessment was not indicated in women at low-risk (Lim et al, 2011). The Institute for Clinical Systems Improvement’s clinical practice guideline on “Management of labor” (Creedon et al, 2013) recommended the use of transvaginal sonogram for cervical length for monitoring of patients with sign/symptoms of preterm labor and early cervical change. However, this recommendation is based on low-quality evidence.

In a Cochrane review, Berghella et al (2013) evaluated the effectiveness of antenatal management based on transvaginal ultrasound of cervical length (TVU CL) screening for preventing preterm birth (PTB). These investigators searched the Cochrane Pregnancy and Childbirth Group’s Trials Register (August 31, 2012), reviewed the reference lists of all articles and contacted experts in the field for additional and ongoing trials. Published and unpublished randomized controlled trials (RCTs) including
pregnant women between the gestational ages of 14 to 32 weeks screened with TVU CL for risk of PTB were selected for analysis. This review focused exclusively on studies based on knowledge versus no knowledge of TVU CL results. All potential studies identified from the search were independently assessed for inclusion by 3 review authors. They also analyzed studies for quality measures and extracted data. Of the 13 trials identified, 5 were eligible for inclusion (n = 507). Three included singleton gestations with preterm labor (PTL); 1 included singleton gestations with preterm premature rupture of membranes (PPROM); and 1 included twin gestations with or without PTL. In the 3 trials of singleton gestations with PTL, 290 women were randomized; 147 to knowledge and 143 to no knowledge of TVU CL. Knowledge of TVU CL results was associated with a non-significant decrease in PTB at less than 37 weeks (22.3 % versus 34.7 %, respectively; average risk ratio [RR] 0.59, 95 % CI: 0.26 to 1.32; 2 trials, 242 women) and at less than 34 weeks (6.9 % versus 12.6 %; RR 0.55, 95 % CI: 0.25 to 1.20; 3 trials, 256 women). Delivery occurred at a later gestational age in the knowledge versus no knowledge groups (mean difference (MD) 0.64 weeks, 95 % CI: 0.03 to 1.25; 3 trials, 290 women). For all other outcomes for which there were available data (PTB at less than 34 or 28 weeks; birthweight less than 2,500 grams; perinatal death; maternal hospitalization; tocolysis; and steroids for fetal lung maturity), there was no evidence of a difference between groups. The trial of singleton gestations with PPROM (n = 92) evaluated as its primary outcome safety of TVU CL in this population, and not its effect on management. There was no evidence of a difference in incidence of maternal and neonatal infections between the TVU CL and no TVU CL groups. In the trial of twin gestations with or without PTL (n = 125), there was no evidence of a difference in PTB at less than 36, 34, or 30 weeks, gestational age at delivery, and other perinatal and maternal outcomes between the TVU CL and the no TVU CL groups. Life-table analysis revealed significantly less PTB at less than 35 weeks in the TVU CL group compared with the no TVU CL group (p = 0.02). The authors concluded that currently, there is insufficient evidence to recommend routine screening of
asymptomatic or symptomatic pregnant women with TVU CL. Since there is a non-significant association between knowledge of TVU CL results and a lower incidence of PTB at less than 37 weeks in symptomatic women, the authors encouraged further research. Future studies should look at specific populations separately (e.g., singleton versus twins; symptoms of PTL or no such symptoms), report on all pertinent maternal and perinatal outcomes, and include cost-effectiveness analyses. Most importantly, they stated that future studies should include a clear protocol for management of women based on TVU CL results, so that it can be easily evaluated and replicated.

In an observational, prospective study, Kuusela and colleagues (2015) evaluated cervical length in asymptomatic women with singleton pregnancies in the 2nd trimester by means of TV-US, and examined the relation between cervical length and spontaneous preterm delivery. A total of 2,122 asymptomatic women with live singleton pregnancies without fetal anomalies were included in this study. Cervical length was measured at between 16 and 23 weeks of gestation by means of TV-US; data were analyzed using logistic regression analysis. Main outcome measure were cervical length in relation to spontaneous preterm delivery less than 34 weeks (primary outcome) and less than 37 weeks of gestation (secondary outcome). Eleven women had a cervical length of less than or equal to 25 mm (0.5 %) and 73 women had a cervical length of less than or equal to 30 mm (3.4 %). Spontaneous preterm delivery at less than 34 weeks occurred in 22/2,061 women (1.1 %) and at less than 37 weeks in 87/2061 women (4.2 %). There was a significant association between cervical length and spontaneous preterm delivery at less than 34 weeks (odds ratio [OR] 1.78; 95 % CI: 1.19 to 2.65 for a decrease of cervical length by 5 mm) but no significant association at less than 37 weeks. The authors concluded that the rate of short cervical length of less than or equal to 25 mm was lower than expected. The study confirmed the increased risk of spontaneous preterm delivery in women with a short cervix, although the analysis was based on only a few cases. They stated that in Sweden, a larger study is needed to evaluate the prevalence of short
In a systematic review and meta-analysis, Conde-Agudelo and Romero (2015) examined the accuracy of changes in transvaginal sonographic cervical length over time in predicting preterm birth in women with singleton and twin gestations. Data sources included PubMed, Embase, Cinahl, Lilacs, and Medion (all from inception to June 30, 2015), bibliographies, Google scholar, and conference proceedings. Cohort or cross-sectional studies reporting on the predictive accuracy for preterm birth of changes in cervical length over time were selected for analysis. Two reviewers independently selected studies, assessed the risk of bias, and extracted the data. Summary receiver-operating characteristic curves, pooled sensitivities and specificities, and summary likelihood ratios were generated. A total of 14 studies met the inclusion criteria, of which 7 provided data on singleton gestations (3,374 women) and 8 on twin gestations (1,024 women). Among women with singleton gestations, the shortening of cervical length over time had a low predictive accuracy for preterm birth at less than 37 and less than 35 weeks of gestation with pooled sensitivities and specificities, and summary positive and negative likelihood ratios ranging from 49 % to 74 %, 44 % to 85 %, 1.3 to 4.1, and 0.3 to 0.7, respectively. In women with twin gestations, the shortening of cervical length over time had a low-to-moderate predictive accuracy for preterm birth at less than 34, less than 32, less than 30, and less than 28 weeks of gestation with pooled sensitivities and specificities, and summary positive and negative likelihood ratios ranging from 47 % to 73 %, 84 % to 89 %, 3.8 to 5.3, and 0.3 to 0.6, respectively. There were no statistically significant differences between the predictive accuracies for preterm birth of cervical length shortening over time and the single initial and/or final cervical length measurement in 8 of 11 studies that provided data for making these comparisons. In the largest and highest-quality study, a single measurement of cervical length obtained at 24 or 28 weeks of gestation was significantly more predictive of preterm birth than any decrease in cervical length between
these gestational ages. The authors concluded that change in transvaginal sonographic cervical length over time is not a clinically useful test to predict preterm birth in women with singleton or twin gestations. A single cervical length measurement obtained between 18 and 24 weeks of gestation appeared to be a better test to predict preterm birth than changes in cervical length over time.

An ACOG Committee Opinion (2012) reached the following conclusions:

- ACOG and the American Institute of Ultrasound in Medicine recommend a cervical length measurement at around 18 to 22 gestational weeks, at the same time as the ultrasound for fetal anatomic survey, because this is a useful screening test to predict spontaneous preterm birth.
- Women with cervical length less than 25 mm at 14 to 28 weeks should undergo a subsequent transvaginal ultrasound to confirm this finding.
- Women in whom short cervical length is confirmed should have a review of risk factors for preterm birth, as well as of management options.
- TVU measurement of cervical length should be performed only when interventions to reduce risk for preterm birth are available.
- The utility of universal cervical length screening to prevent preterm birth is still controversial and under debate.

Cervical length measurement should be done according to strict quality criteria, which are available to practitioners in the United States via CLEAR (Cervical Length Education and Review) and the Perinatal Quality Foundation and in Europe through the FetalMedicine Foundation. The CLEAR program provides educational lectures, optional examinations, and scored image reviews. Those who complete the lectures and who pass the examination and image review receive documents verifying that they have completed the CLEAR program. They also qualify for CME provided by both Society of Diagnostic Medical Sonographers (SDMS) and ACOG. Names of those who
complete the program are listed on the CLEAR website.

Transvaginal Ultrasonography for Guidance during Embryo Transfer:

Teixeira et al (2015) summarized the current evidence on the effect of using US guidance during embryo transfer (ET). In this systematic review, these investigators included RCTs examining the effect of the use of US guidance during ET; data from studies using the same catheter type in study arms were not pooled with the results from studies that used different catheter types. A total of 21 studies were included in the quantitative analysis: 18 compared “US guidance” with “clinical touch”, of which 1 was subsequently excluded from the quantitative meta-analysis owing to a lack of available data; 3 studies compared TVU guidance with trans-abdominal US guidance; and 1 study compared “hysterosonometry before ET” with US guidance. Comparison of the use of US guidance with clinical touch, in studies that used the same catheter type in the study arms, indicated a benefit of using US guidance during ET on the rates of live-birth (RR, 1.48; 95 % CI: 1.16 to 1.87), based on 2 studies involving 888 women with moderate-quality evidence, and on the rates of clinical pregnancy (RR, 1.32; 95 % CI: 1.18 to 1.46), based on 13 studies involving 3,641 women with high-quality evidence. However, when comparing the use of US guidance with clinical touch in studies that used different catheter types, the results suggest that using US guidance during ET has no effect on the rates of reproductive outcome: live-birth (RR, 0.99; 95 % CI: 0.83 to 1.19), based on 1 study involving 1,649 women with moderate-quality evidence; clinical pregnancy (RR, 1.04; 95 % CI: 0.89 to 1.21), based on 5 studies involving 2,949 women with moderate-quality evidence. The estimates for the rate of miscarriage and for the other identified comparisons were imprecise. The authors concluded that the available evidence suggested that there is a benefit of using US guidance during ET. However, both US-guided transfer and clinical touch should be considered acceptable, as the benefit of US is not large and should be balanced against the increased cost and need to change the catheter type. They
stated that more studies are needed before conclusions can be drawn regarding the effect of other techniques on reproductive outcome.

Diagnosis of Vasa Previa:

Ruiter and colleagues (2015) stated that vasa previa is an obstetric complication in which the fetal blood vessels lie outside the chorionic plate in close proximity to the internal cervical os. In women with vasa previa, the risk of rupture of these vessels is increased, thus potentially causing fetal death or serious morbidity. These investigators evaluated the accuracy of ultrasound in the prenatal diagnosis of vasa previa. They searched Medline, Embase, the Cochrane Library and PubMed for studies on vasa previa. Two reviewers independently selected studies on the accuracy of ultrasound in the diagnosis of vasa previa. The studies were scored on methodological quality using the Quality Assessment of Diagnostic Accuracy Studies tool (QUADAS-2). Data on sensitivity and specificity were subsequently extracted. The literature search revealed 583 articles, of which 2 prospective and 6 retrospective cohort studies were eligible for inclusion in the qualitative analysis. All studies documented methods suitable for the prenatal diagnosis of vasa previa; 4 out of the 8 studies used TV-US for primary evaluation, while the remaining four studies used trans-abdominal US and performed a subsequent TV-US when vasa previa was suspected. The QUADAS-2 tool reflected poor methodology in 6 of the 8 included studies, and prenatal detection rates varied from 53 % (10/19) to 100 % (total of 442,633 patients, including 138 cases of vasa previa). In the 2 prospective studies (n = 33,795, including 11 cases of vasa previa), trans-vaginal color Doppler performed during the 2nd trimester detected all cases of vasa previa (sensitivity, 100 %) with a specificity of 99.0 to 99.8 %. The authors concluded that the accuracy of US in the diagnosis of vasa previa is high when performed trans-vaginally in combination with color Doppler.

An UpToDate review on “Velamentous umbilical cord insertion
Prenatal diagnosis of vasa previa is based on identification of membranous fetal vessels passing across or in close proximity to the internal cervical os by real-time transvaginal ultrasound examination with color Doppler. Close proximity has been defined as within 2 cm of the internal os; however, only limited data are available to support this specific measurement. In prospective studies in which the investigators were specifically looking for vasa previa, sonography plus color Doppler had high diagnostic sensitivity: 10/10 cases after 26 weeks and 1/1 case at 18 to 20 weeks; in each study, 1 additional case diagnosed prenatally could not be confirmed at delivery. In a systematic review including both prospective and retrospective studies, sensitivity ranged from 53 to 100%.

<table>
<thead>
<tr>
<th>CPT Codes / HCPCS Codes / ICD-10 Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information in the [brackets] below has been added for clarification purposes. Codes requiring a 7th character are represented by &quot;+&quot;.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>ICD-10 codes will become effective as of October 1, 2015:</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>CPT codes covered if selection criteria are met:</td>
</tr>
<tr>
<td>76830</td>
</tr>
<tr>
<td>ICD-10 codes covered if selection criteria are met (not all-inclusive):</td>
</tr>
<tr>
<td>C53.0 - C55</td>
</tr>
<tr>
<td>C56.1 - C56.9</td>
</tr>
<tr>
<td>C57.4</td>
</tr>
<tr>
<td>C79.60 - C79.62</td>
</tr>
<tr>
<td>C79.82</td>
</tr>
<tr>
<td>D07.0</td>
</tr>
<tr>
<td>Code</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>D25.0 - D25.9</td>
</tr>
<tr>
<td>D26.1</td>
</tr>
<tr>
<td>D27.0 - D27.9</td>
</tr>
<tr>
<td>D39.0</td>
</tr>
<tr>
<td>D39.10 - D39.12</td>
</tr>
<tr>
<td>N39.3</td>
</tr>
<tr>
<td>N70.01 - N77.1</td>
</tr>
<tr>
<td>N80.0</td>
</tr>
<tr>
<td>N80.5</td>
</tr>
<tr>
<td>N83.20 - N83.29</td>
</tr>
<tr>
<td>N91.0 - N94.9</td>
</tr>
<tr>
<td>N94.0</td>
</tr>
<tr>
<td>N94.89</td>
</tr>
<tr>
<td>N95.0</td>
</tr>
<tr>
<td>N97.0 - N97.9</td>
</tr>
<tr>
<td>Q51.0 - Q51.5, Q51.810 - Q51.818, Q51.9</td>
</tr>
<tr>
<td>R10.0 - R10.829, R10.84 - R10.9</td>
</tr>
<tr>
<td>Code</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>R14.0</td>
</tr>
<tr>
<td>R19.00</td>
</tr>
<tr>
<td>R19.09</td>
</tr>
<tr>
<td>R35.0</td>
</tr>
<tr>
<td>R39.15</td>
</tr>
<tr>
<td>R68.81</td>
</tr>
<tr>
<td>T83.31x</td>
</tr>
<tr>
<td>T83.39x</td>
</tr>
<tr>
<td>T83.81x</td>
</tr>
<tr>
<td>T83.9xx</td>
</tr>
<tr>
<td>Z15.02</td>
</tr>
<tr>
<td>Z30.431</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Z01.411</td>
</tr>
<tr>
<td>Z01.419</td>
</tr>
<tr>
<td>Z12.73</td>
</tr>
<tr>
<td>Z12.79</td>
</tr>
<tr>
<td>Z30.430</td>
</tr>
<tr>
<td>Z97.5</td>
</tr>
</tbody>
</table>

**The above policy is based on the following references:**


15. NHS Centre for Reviews and Dissemination. Screening for ovarian cancer: A systematic review. York, UK: Centre for Reviews and Dissemination (CRD); 1998.
22. Villanueva EVS. The diagnostic characteristics of transvaginal ultrasound in the detection of endometrial cancer in postmenopausal women taking hormone replacement therapy. Clayton, Victoria, Australia: Centre for Clinical Effectiveness (CCE); 2000.


44. Valentin L. Use of morphology to characterize and


65. American College of Obstetricians and Gynecologists. ACOG Committee Opinion No. 440: The role of transvaginal ultrasonography in the evaluation of


82. Shipp TD. Ultrasound examination in obstetrics and gynecology. UpToDate [online serial]. Waltham, MA: UpToDate; reviewed April 2013.


85. Berghella V, Baxter JK, Hendrix NW. Cervical assessment


AETNA BETTER HEALTH® OF PENNSYLVANIA

Amendment to
Aetna Clinical Policy Bulletin Number: 0530
Transvaginal Ultrasonography

There are no amendments for Medicaid.

www.aetnabetterhealth.com/pennsylvania
Revised 04/2017