2.04.66	Serum Biomarker Human Epididymis Protein 4		
Original Policy Date:	July 2, 2010	**Effective Date:**	February 1, 2024
Section:	2.0 Medicine	**Page:**	Page 1 of 23

Policy Statement

I. Measurement of human epididymis protein 4 (HE4) is considered **investigational** for all indications.

NOTE: Refer to Appendix A to see the policy statement changes (if any) from the previous version.

Policy Guidelines

Coding

The following CPT code is specific for this test:

- **86305**: Human epididymis protein 4 (HE4)

If HE4 is performed as part of the Risk of Ovarian Malignancy Algorithm (ROMA™) test, the following code would be reported:

- **81500**: Oncology (ovarian), biochemical assays of two proteins (CA-125 and HE4), utilizing serum, with menopausal status, algorithm reported as a risk score

Description

Human epididymis protein 4 (HE4) is a novel biomarker that has been cleared by the U.S. Food and Drug Administration for monitoring patients with epithelial ovarian cancer. HE4 is proposed as a replacement for or a complement to cancer antigen 125 (CA 125) for monitoring disease progression and recurrence. HE4 has also been proposed as a test to evaluate women with ovarian masses and to screen for ovarian cancer in asymptomatic women.

Related Policies

- N/A

Benefit Application

Benefit determinations should be based in all cases on the applicable contract language. To the extent there are any conflicts between these guidelines and the contract language, the contract language will control. Please refer to the member's contract benefits in effect at the time of service to determine coverage or non-coverage of these services as it applies to an individual member.

Some state or federal mandates (e.g., Federal Employee Program [FEP]) prohibits plans from denying Food and Drug Administration (FDA)-approved technologies as investigational. In these instances, plans may have to consider the coverage eligibility of FDA-approved technologies on the basis of medical necessity alone.

Regulatory Status

Multiple HE4 test kits have been cleared by the FDA through the 510(k) process and summarized in Table 1. The FDA determined that this device was substantially equivalent to a CA 125 assay kit for use as an aid in monitoring disease progression or recurrence in patients with epithelial ovarian cancer.

The FDA-approved indication states that serial testing for HE4 should be done in conjunction with other clinical methods used for monitoring ovarian cancer and that the HE4 test is not intended to assess the risk of disease outcomes.

Table 1. Serum Human Epididymis Protein 4 Tests Cleared by the U.S. Food and Drug Administration

Test	Manufacturer	Location	Date Cleared	510(k) No.
HE4 EIA Kit	Fujirebio Diagnostics	Malvern, PA	06/09/2008	K072939
ARCHITECT HE4 assay (CMIA)	Fujirebio Diagnostics	Malvern, PA	03/18/2010	K093957
ELECSYS HE4 (CMIA)	Roche Diagnostics	Indianapolis, IN	09/10/2012	K112624
Lumipulse G HE4 Immunoreaction Cartridges	Fujirebio Diagnostics	Malvern, PA	11/24/2015	K151378

CMIA: chemiluminescent microparticle immunoassay; HE4: human epididymis protein 4; EIA: enzymatic immunoassay.
FDA product code: OIU.

Rationale

Background
Ovarian Cancer
Ovarian cancer is the fifth most common cause of cancer mortality among U.S. women. According to Surveillance Epidemiology and End Results data, in 2023 , an estimated 19,710 women will be diagnosed with ovarian cancer and 13,270 women will die of the disease.[1] The stage at diagnosis is an important predictor of survival; however, most women are not diagnosed until the disease has spread. For the period 2012 to 2018, 57% of women with ovarian cancer were diagnosed when the disease had distant metastases (stage IV), and this was associated with a 5-year survival rate of 31%. In contrast, 17% of women diagnosed with localized cancer (stage I) had a 5-year survival rate of 93%. Epithelial ovarian tumors account for 85% to 90% of ovarian cancers.[2]

Research from the Ovarian Cancer in Women of African Ancestry (OCWAA) consortium reports that Black women with ovarian cancer have worse survival than White women.[3] Contributors to this disparity may include education level, nulliparity, smoking status, body mass index, diabetes, and postmenopausal hormone therapy duration.

Treatment
The standard treatment for epithelial ovarian cancer is surgical staging and primary cytoreductive surgery followed by chemotherapy in most cases. There is a lack of consensus about an optimal approach to the follow-up of patients with ovarian cancer after or during primary treatment. Patients undergo regular physical examinations and may have imaging studies. In addition, managing patients with serial measurements of the biomarker cancer antigen 125 (CA 125) to detect early recurrence of disease is common. A rising CA 125 level has been found to correlate with disease recurrence and has been found to detect recurrent ovarian cancer earlier than clinical detection. However, a survival advantage of initiating treatment based on early detection with CA 125 has not been demonstrated to date. For example, a 2010 randomized controlled trial in women with ovarian cancer that was in complete remission did not find a significant difference in overall survival when treatment for remission was initiated after CA 125 concentration exceeded twice the limit of normal compared to delaying treatment initiation until symptom onset.[4]

Human epididymis protein 4 (HE4) is a protein that circulates in the serum and has been found to be overexpressed in epithelial ovarian cancer, lung adenocarcinoma, breast cancer, pancreatic cancer, endometrial cancer, and bladder cancer. HE4 is made up of 2 whey acidic proteins with a 4 disulfide core domain and has been proposed as a biomarker for monitoring patients with epithelial ovarian cancer.

Evaluation of Adnexal Masses

This evidence review also addresses the use of the HE4 as a stand-alone test for evaluating women with ovarian masses who have not been diagnosed with ovarian cancer. Such patients undergo a diagnostic workup to determine whether the risk of malignancy is sufficiently high to warrant surgical removal. In patients for whom surgery is indicated, further evaluation may be warranted to determine if a surgical referral to a specialist with expertise in ovarian cancer is warranted. The Risk of Ovarian Malignancy Algorithm (ROMA) test combines HE4, CA 125, and menopausal status into a numeric score. The ROMA test has been cleared by U.S. Food and Drug Administration (FDA) for predicting the risk that an adnexal mass is malignant.

Literature Review

Evidence reviews assess whether a medical test is clinically useful. A useful test provides information to make a clinical management decision that improves the net health outcome. That is, the balance of benefits and harms is better when the test is used to manage the condition than when another test or no test is used to manage the condition.

The first step in assessing a medical test is to formulate the clinical context and purpose of the test. The test must be technically reliable, clinically valid, and clinically useful for that purpose. Evidence reviews assess the evidence on whether a test is clinically valid and clinically useful. Technical reliability is outside the scope of these reviews, and credible information on technical reliability is available from other sources.

Promotion of greater diversity and inclusion in clinical research of historically marginalized groups (e.g., People of Color [African-American, Asian, Black, Latino and Native American]; LGBTQIA (Lesbian, Gay, Bisexual, Transgender, Queer, Intersex, Asexual); Women; and People with Disabilities [Physical and Invisible]) allows policy populations to be more reflective of and findings more applicable to our diverse members. While we also strive to use inclusive language related to these groups in our policies, use of gender-specific nouns (e.g., women, men, sisters, etc.) will continue when reflective of language used in publications describing study populations.

Biomarker Human Epididymis Protein 4 Testing for Ovarian Cancer
Clinical Context and Test Purpose

The purpose of testing serum biomarker human epididymis protein 4 (HE4) levels is to provide an alternative to or an improvement on existing testing in individuals with ovarian cancer.

The following PICO was used to select literature to inform this review.

Populations

The relevant populations of interest is individuals with epithelial ovarian cancer who have had primary treatment.

Interventions

The test being considered is testing serum biomarker HE4 levels. These levels are used to monitor for surveillance of progression (response to primary treatment) or recurrence in individuals with ovarian cancer.

Comparators

Comparators of interest include measurement of the cancer antigen 125 (CA 125) test and measurement of the combination CA 125 plus HE4. Typically, individuals with ovarian cancer undergoing primary chemotherapy after cytoreductive surgery will also have monitoring for a response with a computed tomography scan. After the completion of primary treatment, patients may have other monitoring imaging studies such as positron emission tomography.

Outcomes
The general outcomes of interest are overall survival (OS), disease-specific survival, test accuracy, test validity, and other test performance measures. Change in disease status is also an outcome of interest in individuals with ovarian cancer.

The timing of follow-up after testing HE4 serum levels in an individual with ovarian cancer is based on the stage of the disease, type of prior therapy and guideline recommendations.

Study Selection Criteria
For the evaluation of clinical validity of HE4 testing, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology.
- Included a suitable reference standard.
- Patient/sample characteristics were described.
- Patient/sample selection criteria were described.

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
Systematic Review
Han et al (2021) published a systematic review and meta-analysis on the value of HE4 in predicting chemotherapy resistance in patients with ovarian cancer.[5,] An analysis of 8 studies (I^2=74%) found that preoperative HE4 had a sensitivity of 80% (95% confidence interval [CI], 65% to 90%) and specificity of 67% (95% CI, 54% to 77%) in predicting resistance to platinum chemotherapy. After the third cycle of chemotherapy (5 studies; I^2=49%), the sensitivity and specificity were 86% (95% CI, 72% to 94%) and 85% (95% CI, 70% to 93%), respectively.

Food and Drug Administration Pivotal Studies
The U.S. Food and Drug Administration (FDA) documents included information on the diagnostic performance of HE4 for monitoring the progression and recurrence of ovarian cancer. The FDA materials addressed the noninferiority rather than the superiority of HE4 tests to CA 125. A study reported in the 510(k) substantial equivalence determination decision summary for the HE4 enzyme immunoassay (EIA) evaluated whether this test is noninferior to the CA 125 test. The study included samples from 80 women with epithelial ovarian cancer (EOC) who were undergoing serial surveillance of cancer progression.[6,] Blood samples were obtained from a large cancer center in the U.S.; they were not drawn specifically for this study. A total of 354 samples were obtained for the 80 women (women had multiple visits over time). Receiver operating characteristic curve analysis was used to compare the 2 assays, and clinical evidence of progression was used as the reference standard. When a positive change in HE4 level (i.e., to indicate disease progression) was defined as a value at least 25% higher than the previous value of the test, the sensitivity of the test was 76 (60.3%) of 126, and the specificity was 171 (75%) of 228 (note that the unit of analysis was the number of samples rather than the number of women.) The area under the receiver operating characteristic curves were found to be similar (HE4=0.725 vs. CA 125=0.709), with an overlap in the CIs; according to the authors, this indicated that the HE4 assay was not inferior to the CA 125 assay for detecting cancer progression.

Another analysis estimated the cutoff values and specificities for the HE4 and CA 125 assays across a range of fixed sensitivities, where the sensitivities of the HE4 and CA 125 assays were set at the same values. The specificity values for CA 125 and HE4 did not differ statistically at the respective cutoffs and sensitivities. These data were also said to confirm that the HE4 EIA test was not inferior to the CA 125 test for detecting ovarian cancer progression.

The 510(k) substantial equivalence determination decision summary for the ARCHITECT HE4 assay reported data from a retrospective study using remnant serial samples from 76 women diagnosed with EOC being monitored after completion of chemotherapy.[7,] The eligibility criteria included the availability of at least 3 serial specimens; samples could have been drawn during and/or after treatment. Clinical determination of disease progression was used as the reference standard. A positive test was defined as an HE4 level that was 14% higher than the previous reading. Using this cutoff, the sensitivity of the assay for detecting progressive disease was 53 (53.5%) of 99 events. The specificity of the assay was 260 (78.5%) of 331. Of note, the sensitivity is lower than that previously reported for the HE4 EIA test at a similar specificity, when a cutoff of a 25% increase was used (sensitivity, 60.3%; specificity, 75%).

The FDA documents noted that there is no clinically accepted cutoff for monitoring cancer progression in EOC patients using the HE4 assays. As mentioned, a study included in the HE4 EIA assay materials defined a positive test as a level 25% higher than a previous measurement, and a study on the ARCHITECT HE4 test defined a positive test as an increase of at least 14% in the level of HE4. The FDA documents further stated that clinicians may decide whether to use the cutoffs in the studies or another cutoff that reflects personal preferences in the tradeoff between sensitivity and specificity.

Published observational studies on the diagnostic performance of HE4 for monitoring progression and/or recurrence of EOC are described next.

Nonrandomized Observational Studies

Nassir et al(2016) published an analysis of data from an earlier study by Braicu et al (OVCAD study, 2013).[8,] The OVCAD study evaluated 275 patients with advanced primary ovarian cancer who underwent cytoreductive surgery and adjuvant platinum-based chemotherapy at a specialized clinic.Ninety-two (33%) of 275 patients, who had preoperative and follow-up plasma samples for analyzing HE4 and CA 125, were included in the analysis; however, 13 preoperative HE4 samples and 10 postoperative CA 125 samples were missing. Both preoperative HE4 and CA 125 levels significantly predicted 12-month recurrence or death. Among responders, median OS was worse among patients for whom both biomarkers were elevated (hazard ratio [HR], 17.96; 95% CI, 4.00 to 80.85; p<.001) compared to patients for whom no biomarker was elevated. The CI for the OS analysis was wide, indicating an imprecise estimate. There was no significant association with median OS when only 1 biomarker was elevated; the sample size may have been inadequate for this analysis.

Vallius et al (2017) reported a study that was designed to assess fluorodeoxyglucose-positron emission tomography/computed tomography imaging and serum tumor markers in epithelial ovarian cancer staging and chemotherapy response. A substudy analysis evaluated the use of HE4 profiles to predict treatment outcomes during the first line of chemotherapy after primary cytoreductive surgery.[9,] HE4 and CA 125 were measured in patients with the Federation of International Gynecology and Obstetrics III/IV EOC who received primary debulking surgery followed by platinum-based chemotherapy or neoadjuvant chemotherapy followed by interval debulking surgery. HE4 at the time of diagnosis was not associated with progression-free survival (PFS) (p=.24), whereas lower CA 125 at the time of diagnosis predicted longer PFS (HR, 1.45; 95% CI, 1.09 to 1.94; p=.01). When patients who underwent either surgical approach were combined (n=40), those with no macroscopic residual disease after cytoreductive surgery were more likely to have lower postoperative HE4 values. Both HE4 and CA 125 nadir values were associated with a greater complete response to chemotherapy. Tables 2 through 5 below summarize findings for this study.

Potenza et al (2020) retrospectively assessed 78 patients with EOC to determine whether HE4 and CA 125 measured at diagnosis and before each platinum-based chemotherapy cycle could predict lack of response to chemotherapy and disease recurrence.[10,] The proportions of patients who were sensitive, partially resistant, and refractory to chemotherapy were 73%, 16.6%, and 6.4%, respectively. After a median follow-up of 10 months, both HE4 and CA 125 had a positive correlation to PFS when

measured after the third chemotherapy cycle (both p=.0001). At the time of diagnosis, HE4 and CA 125 levels lower than the population mean value were also positively correlated to PFS (both p<.05).

Salminen et al (2020) conducted a prospective observational study in 143 women with histologically confirmed high-grade serous carcinoma (a common and aggressive form of EOC) to assess biomarkers, including CA 125 and HE4, for treatment monitoring and prognostic stratification.[11,] Included patients received primary treatment with either primary debulking surgery followed by chemotherapy (n=58) or neoadjuvant chemotherapy with interval debulking surgery plus adjuvant chemotherapy (n=85). Chemotherapy regimens consisted of carboplatin plus a taxane (n=125), carboplatin alone (n=16) or other/unknown (n=2). Follow-up times ranged between 1.5 months to 10.2 years. At the time of progression, multivariate analysis showed that HE4 concentration elevations greater than 199.20 pmol/L were significantly associated with a reduction in OS(HR, 5.85; 95% CI, 2.07 to 16.51; p=.0001); elevations in CA 125 greater than 162 U/mL were not (HR, 1.39; 95% CI, 0.59 to 3.28; p=.45). Serum HE4 concentrations were also found to be significantly higher at baseline in patients with a higher tumor burden compared to those with less extensive tumor growth (p<.0001) while CA 125 concentrations were not (p=.067). At baseline after cytoreductive surgery, neither CA 125 (p=.641) or HE4 (p=.054) concentrations were significantly associated with the amount of residual disease. Nadir CA 125 and HE4 levels were both found to be significantly elevated in patients who developed platinum-resistant disease (p<.0001).

Rong et al (2021) conducted a retrospective study that assessed the prognostic value of HE4 and CA 125 in 89 patients with EOC.[12,] All patients received 6 to 8 cycles of platinum-based chemotherapy after surgery. HE4 (cutoff, 70 pmol/L) and CA 125 (cutoff, 35 U/mL) were measured before treatment, after each cycle, and at the time of recurrence. After a median follow-up of 35 months, 73 patients were platinum-sensitive and 16 patients were platinum-resistant. The sensitivity and specificity of HE4 in predicting platinum responsiveness after the third chemotherapy cycle were 75% and 80.8%, respectively. HE4 had a positive predictive value (PPV) of 54.5% and negative predictive value (NPV) of 93.7%. The sensitivity, specificity, and PPV, and NPV of CA 125 after the first chemotherapy cycle were 75%, 71.2%, 36.4%, and 92.9%, respectively. The combination of both biomarkers had a sensitivity and specificity for predicting platinum responsiveness of 50% and 94.5%, respectively, with a PPV of 66.7% and NPV of 89.6%. HE4 predicted 2-year PFS after the third and sixth chemotherapy cycles (p=.001 and p=.011, respectively). CA 125 predicted 2-year PFS only after the first chemotherapy cycle (p=.023). Prolonged PFS and OS were significantly associated with HE4 after the third cycle (p<.0001) and CA 125 after the first cycle (p<.0001).

Samborski et al (2022) retrospectively examined the utility of HE4 in comparison to CA 125 in women undergoing surveillance after treatment for EOC between January 1997 to October 2010.[13,] A total of 129 women with a diagnosis of EOC were identified and included in the analysis, of which 11 women had stage I disease (8.5%), 12 had stage II disease (9.3%), 94 had stage III disease (72.9%), and 12 had stage IV (9.3%) disease. At a threshold of 25% change in serum biomarker level indicating progressive disease, HE4 had an overall accuracy for change in disease status of 81.8% (95% CI, 79.7% to 83.7%) with a specificity of 90.5% (95% CI, 88.7% to 92.1%), sensitivity of 45.2% (95% CI, 39.2% to 51.2%), PPV of 53.2% (95% CI, 46.6% to 59.7%) and a NPV of 87.4% (95% CI, 85.4% to 89.2%). The concordance comparison of HE4 accuracy (81.8%)/CA 125 accuracy (82.6%) was 0.990, indicating HE4 was not inferior to CA 125 (McNemar's test p-value=.522).

Table 2. Summary of Observational Study Characteristics

Study	Study Type	Country	Dates	Participants N	Treatment 1 n	Treatment 2 n
Vallius et al (2017)[9,]	Observational cohort	Finland	2009-2014	FIGO Stage III to IV EOC 49	PDS + platinum-based chemotherapy 22	NACT + IDS 27

EOC: epithelial ovarian cancer; FIGO: Federation of International Gynecology and Obstetrics; IDS: interval debulking surgery; NACT: neoadjuvant chemotherapy; PDS: primary debulking surgery.

Table 3. Summary of Observational Study Results

Study	Treatment 1[a]	Treatment 2[b]
Vallius et al (2017)[9,]		
	Median (range)	Median (range)
HE4 (pmol/L)		
At diagnosis	573 (59 to 1391)	1070 (156 to 12,128)
Preoperative	N/A	104 (35 to 477)
Postoperative	96 (34 to 856)	99 (39 to 384)
Nadir	48 (25 to 204)	69 (31 to 257)
Post-primary therapy	48 (25 to 431)	61 (31 to 175)
CA 125 (U/mL)		
At diagnosis	1094 (17 to 17,992)	1078 (156 to 20,897)
Preoperative	N/A	43 (7 to 464)
Postoperative	181 (32 to 2023)	42 (6 to 589)
Nadir	12 (4 to 162)	15 (4 to 447)
Post-primary therapy	12 (4 to 127)	15 (4 to 37)

[a] Primary debulking surgery and platinum-based chemotherapy
[b] Neoadjuvant chemotherapy and interval debulking surgery
CA 125: cancer antigen 125; HE4: human epididymis protein 4; N/A: not applicable;.

Relevance and relevance design and conduct limitations are reported in Tables 4 and 5.

Table 4: Study Relevance Limitations

Study	Population	Intervention	Comparator	Outcomes	Duration of Follow-Up
Vallius et al (2017)[9,]	1. Study population is mixed regarding risk factors 2. Clinical context for primary debulking surgery + chemotherapy differs from neoadjuvant chemotherapy + interval debulking surgery				

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose
[d] Outcomes key: 1. Study dose not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests)
[e] Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true positives, true negatives, false positives, false negatives cannot be determined).

Table 5: Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Completeness of Follow-Up[e]	Statistical[f]
Vallius et al (2017)[9,]			1. Broad date range for obtaining samples	1. Assessment of residual disease solely based on surgeon evaluation		1. Not uniformly reported

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.

[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).

[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.

[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for
interpreting tests not described; 4. Expertise of evaluators not described.

[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.

[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to followup
or missing data.

[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison to other tests not reported.

Clinically Useful

The available observational studies have used HE4 alone or in combination with CA 125 to predict residual tumor mass and association with recurrence after primary chemotherapy. In addition, HE4 alone or in combination with CA 125 has been assessed for its association with residual disease and tumor progression during the course of primary chemotherapy after tumor debulking as well as during neoadjuvant chemotherapy followed by interval debulking surgery. Improvement in health outcomes would depend on demonstrating that further assessment and management decisions on patients with ovarian cancer were initiated that would improve health outcomes. There is no clear chain of evidence demonstrating that incremental changes in ovarian cancer recurrence detection would lead to improved health outcomes. No prospective studies were identified that compared health outcomes in patients who had ovarian cancer managed with and without HE4 testing, alone or in combination with CA 125 or other disease markers.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials (RCTs).

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Section Summary: Individuals with Ovarian Cancer

For individuals who have ovarian cancer who receive a measurement of serum biomarker HE4, the evidence includes 7 nonrandomized prospective and retrospective studies comparing the diagnostic accuracy of HE4 with CA 125 for predicting disease progression and/or recurrence. Data submitted to the FDA for approval of commercial HE4 tests found that HE4 was not inferior to CA 125 for detecting ovarian cancer recurrence. Although a single prospective observational study found that elevated levels of HE4, but not CA 125, at the time of cancer progression was significantly associated with reduced OS, a direct comparison between biomarkers was not provided. Overall, the superiority of HE4 to CA 125 (alone or in combination), the key question in the evidence review, was not demonstrated in the available literature. In addition, there is no established cutoff in HE4 levels for monitoring disease progression, and cutoffs in studies varied. There is no direct evidence from prospective controlled studies on the impact of HE4 testing on health outcomes, and no clear chain of evidence that changes in management based on HE4 would lead to an improved health outcome.

Biomarker Human Epididymis Protein 4 Testing for Adnexal Masses
Clinical Context and Test Purpose

The purpose of testing serum biomarker HE4 levels is to provide an alternative to or an improvement on existing testing in individuals with adnexal masses.

The following PICO was used to select literature to inform this review.

Populations
The relevant populations of interest are individuals with adnexal masses.

Interventions
The test being considered is testing serum biomarker HE4 levels. These levels are used to evaluate individuals with adnexal masses who are undergoing diagnostic workup for ovarian cancer.

Comparators
Comparators of interest include measurement of CA 125 and measurement of the combination CA 125 plus HE4.

Outcomes
The general outcomes of interest are OS, disease-specific survival, test accuracy, test validity, and other test performance measures.

Evaluation of an adnexal mass would be determined by whether or not the individual has surgical management, and typical clinical follow-up in the absence of a pathological diagnosis would be every 6 months.

Study Selection Criteria
For the evaluation of clinical validity of HE4 testing, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology.
- Included a suitable reference standard.
- Patient/sample characteristics were described.
- Patient/sample selection criteria were described.

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
Systematic Reviews and Meta-Analyses
A number of meta-analyses have assessed studies on the accuracy of HE4 for diagnosing ovarian cancer. Table 6 presents the pooled sensitivities and specificities of HE4 from meta-analyses that conducted quality assessments of individual studies and that limited their selections to studies using pathologic findings as to the reference standard for ovarian cancer diagnosis.[14,-23,]

Table 6. Meta-Analyses of Human Epididymis Protein 4 for Diagnosing Ovarian Cancer

Meta-Analyses (Year)	No. of Studies	Pooled Sensitivity, % (95% CI)	Pooled Specificity, % (95% CI)
Olson et al (2021)[14,]	7	79.4 (74.1 to 83.8)	84.1 (79.6 to 87.8)
Suri et al (2021)[15,]	25	73 (71 to 75)	90 (89 to 91)
Huang et al (2018)[16,]	18	81 (77 to 85)	91 (86 to 93)
Dayyani et al (2016)[17,]	5	82 (68 to 90)	85 (72 to 93)
Macedo et al (2014)[18,]	45	78 (77 to 79)	86 (85 to 87)
Wang et al (2014)[19,]	28	76 (72 to 80)	93 (90 to 96)
Zhen et al (2014)[20,]	25	74 (72 to 76)	90 (89 to 91)
Yang et al (2013)[21,]	31	73 (71 to 75)	89 (88 to 90)
Ferraro et al (2013)[22,]	14	79 (76 to 81)	93 (92 to 94)
Yu et al (2012)[23,]	12	80 (77 to 83)	92 (90 to 93)

CI, confidence interval.

Meta-analyses differed somewhat in their study inclusion criteria, search dates, and other factors, but, as shown in Table 6, had similar results in terms of the diagnostic value of HE4; pooled sensitivities ranged from 73% to 82%, and pooled specificities ranged from 84.1% to 93%.

Several of the previous meta-analyses also pooled data from studies on the diagnostic accuracy of CA 125, alone and/or in combination with HE4 and findings are shown in Table 7.

Table 7. Meta-Analyses of Cancer Antigen 125 Alone or Combined With Human Epididymis Protein 4 for Diagnosing Ovarian Cancer

Meta-Analyses (Year)	No. of Studies	Pooled Sensitivity, % (95% CI)	Pooled Specificity, % (95% CI)
CA 125 alone			
Olson et al (2021)[14,]	8	81.4 (74.6 to 86.2)	56.8 (47.9 to 65.4)
Suri et al (2021)[15,]	26	84 (82 to 85)	73 (72 to 74)
Dayyani et al (2016)[17,]	5	80 (66 to 89)	83 (66 to 92)
Wang et al (2014)[19,]	28	79 (74 to 84)	82 (77 to 87)
Zhen et al (2014)[20,]	25	74 (72 to 76)	83 (81 to 84)
Ferraro et al (2013)[22,]	13	79 (77 to 82)	78 (76 to 80)
Yu et al (2012)[23,]	10	66 (62 to 70)	87 (85 to 89)
HE4 and CA 125			
Zhen et al (2014)[20,]	9	90 (87 to 92)	85 (82 to 87)
Ferraro et al (2013)[22,]	4	82 (78 to 86)	76 (72 to 80)

CA 125: cancer antigen 125; CI: confidence interval; HE4: human epididymis protein 4.

All meta-analyses included in Table 7, except Dayyani et al (2016), Olson et al (2021), and Suri et al (2021), reported statistical comparisons between the diagnostic performance of HE4 and CA 125. None found that the performance (a combination of sensitivity and specificity) of HE4 and CA 125 differed significantly. However, both Wang et al (2014)[19,] and Zhen et al (2014)[20,] found that the specificity (but not sensitivity) of HE4 was significantly higher than CA 125.

Findings differed in the 2 meta-analyses that compared the diagnostic performance of HE4 and CA 125 with CA 125 alone. Ferraro et al (2013) did not find that the sensitivity and specificity of HE4 in combination with CA 125 differed significantly from that of CA 125 alone.[22,] Zhen et al (2014) found that both the sensitivity and specificity of HE4 combined with CA 125 were significantly better than CA 125 alone.[20,] In the subgroup of 9 studies that made direct comparisons in the Zhen et al (2014) meta-analysis, the sensitivity of HE4 plus CA 125 was 90% (95% CI, 87% to 92%) and for CA 125 alone was 74% (95% CI, 69% to 78%); the specificity of HE4 plus CA 125 was 85% (95% CI, 82% to 87%) and for CA 125 alone was 73% (95% CI, 69% to 76%). In addition, in the Zhen et al (2014) meta-analysis, the overall diagnostic accuracy (measured by the diagnostic odds ratio) was significantly higher for the combination of HE4 and CA 125 than for HE4 alone. Pooled diagnostic odds ratio were 10.31 (95% CI, 6.18 to 17.21) for CA 125 and 53.92 (95% CI, 26.07 to 111.54) for HE4 plus CA 125. Zhen et al (2014) noted several limitations to their meta-analysis, including substantial publication bias for HE4, heterogeneity among studies, and a lack of consideration given to clinical factors such as menopausal status.

Nonrandomized Observational Studies

Several studies have evaluated the diagnostic performance of HE4 as a second-line test after the subjective assessment of transvaginal ultrasound (Tables 8 and 9). The final histologic diagnosis was used as the reference standard.

Kaijser et al (2014) enrolled 389 patients with a suspicious pelvic mass who were scheduled for surgery.[24,] Data on 360 (93%) patients were available for analysis. Experienced ultrasonographers categorized each mass as benign, borderline, or invasive malignant. Serum samples were obtained before surgery, and HE4 levels were measured, using a cutoff of at least 70 pmol/L to indicate malignancy. Overall, subjective ultrasound evaluation by an experienced examiner had

higher sensitivity and specificity than serum HE4. Sensitivity was 97% with subjective assessment ultrasound and 74% with HE4, and specificity was 90% and 85%, respectively. The additional consideration of HE4 levels after sonographers categorized a mass as benign resulted in a slight increase in sensitivity and a large increase in the number of false positives. Moreover, the sequential use of serum HE4 after sonographers categorized a mass as malignant resulted in lower sensitivity and an increase in specificity.

Moszynski et al (2013) retrospectively reviewed records on 253 women with adnexal masses.[25,] Women were examined with transvaginal ultrasound by an experienced examiner before surgery. The sonographer categorized masses as certainly benign, probably benign, uncertain, probably malignant, and certainly malignant. Tumors in the certainly benign and certainly malignant categories were excluded from further analysis, and the remainder (n=145) were considered suspicious tumors. HE4 and CA 125 levels were measured in serum, and a cutoff of 65 pmol/L was used for HE4. The sensitivity and specificity of ultrasound evaluation for diagnosing the suspicious tumors were 93.3% and 90.6%, respectively. Neither HE4 nor CA 125 improved the diagnostic accuracy for suspicious tumors. The sensitivity and specificity of HE4 were 80.0% and 91.7%, respectively, and the sensitivity and specificity of CA 125 were 85.8% and 74.7%, respectively. A logistic regression analysis confirmed that neither HE4 nor CA 125 improved the diagnostic accuracy beyond that of subjective assessment of ultrasonography.

Nikolova et al (2017) conducted a study to measure the effectiveness of HE4 compared with CA 125 for differentiating ovarian endometriosis from EOC in premenopausal women. In the observational study, 164 patients were divided into 4 study groups: ovarian endometriosis (n=37), other benign pelvic masses (n=57), EOCs (n=11), and a control group (n=59).[26,] Analysis of biomarkers in blood samples from all 4 groups determined that HE4 performed the best at differentiating endometriosis from EOC (specificity, 100%; accuracy, 95.83%), while the Copenhagen Index (CPH-I) also performed well (specificity, 97.30%; accuracy, 93.75%). CA 125 was found to have significantly lower specificity and accuracy. Limitations of the study include the relatively small cohort.

Gentry-Maharaj et al (2020) performed a cohort study nested within the screening population of a larger multicenter RCT to assess the ability of HE4 and CA 125 to diagnose ovarian cancer in postmenopausal women with adnexal masses.[27,] The initial trial (United Kingdom Collaborative Trial of Ovarian Cancer Screening; UKCTOCS) randomized 202,638 postmenopausal women to be screened for ovarian cancer to multimodal screening with CA 125 levels, transvaginal ultrasound, or no screening. Women who were randomized to 1 of the screening groups who had an abnormality received a repeat of the initial screening test that they received, and those with persistent abnormalities were further assessed by a clinical team subsequently managed with surgery or conservative management, and had serum CA 125 and HE4 levels taken within 6 months of the scan. A total of 1590 women met these criteria and were found to have adnexal masses. Follow-up occurred for a median of 10.9 years. Reported area under the curve (AUC) values at a specificity of 90% were as follows: 0.896 (95% CI, 0.847 to 0.935) for ultrasound plus CA 125 plus HE4, 0.893 (95% CI, 0.844 to 0.933) for ultrasound plus CA 125, and 0.854 (95% CI, 0.802 to 0.9) for ultrasound plus HE4. Reported AUC values were significantly lower for the ultrasound plus HE4 group compared to the ultrasound plus CA 125 plus HE4 group (p=.033); AUC values were not significantly different when comparing the ultrasound plus CA 125 plus HE4 group to the ultrasound plus CA 125 group (p=.4527). These 2 groups were also reported to have a similar sensitivity at 90% specificity (p=.564); comparison of sensitivity among other groups was not provided.

Carreras-Dieguez (2022) retrospectively evaluated the performance of several serum biomarkers, including CA 125 and HE4, to preoperatively identify EOC or metastatic ovarian cancer in women with a diagnosis of an adnexal mass based on pelvic imaging (N=1071).[28,]In this study, the AUC for HE4 was higher than for CA 125 (0.91 vs. 0.87). Subgroup analysis showed that in premenopausal women (n=629), HE4 performed better than CA 125 (AUC, 0.86 vs 0.76, respectively; p<.05). Conversely, in postmenopausal women (n=442), HE4 and CA 125 AUCs did not significantly differ (0.91 and 0.93,

respectively). In a subgroup of patients with inconclusive diagnosis (n=348), the AUC for HE4 and CA 125 was 0.84 and 0.810, respectively. Lastly, in a subgroup of patients with stage 1 EOC (n=58), the AUC for HE4 and CA 125 was 0.86 and 0.81, respectively.

Lof et al (2022) evaluated the role of HE4 in discriminating benign from malignant tumors in patients who presented with a pelvic mass on ultrasound that was suspected of ovarian origin.[29,] A total of 316 patients were included, of which 195 had a benign, 39 had a borderline and 82 had a malignant ovarian mass. HE4 performed better when age-based cut-offs were applied (sensitivity, 65%; specificity, 79%) instead of one cut-off at 70 pmol/L (sensitivity, 68%; specificity, 65%) or 150 pmol/L (sensitivity, 38%; specificity, 96%). CA 125 performed slightly better when menopausal-based cut-offs were applied (sensitivity, 72%; specificity, 53%) compared with one cut-off at 35 kU/L (sensitivity, 71%; specificity, 50%).

Table 8. Characteristics of Diagnostic Studies Evaluating Human Epididymis Protein 4 as a Test for Ovarian Cancer

Study	Country	Participants	Evaluated Tests
Kaijser et al (2014)[24,]	EU	Women with adnexal masses scheduled for surgery (N=289)	HE4
Moszynski et al (2013)[25,]	EU	Women with adnexal masses (N=253)	HE4, CA 125, individually and combined in ROMA score
Nikolova et al (2017)[26,]	Macedonia, Serbia	Women with ovarian endometriosis, benign pelvic masses, EOCs (N=164)	HE4, CA 125, individually individually and combined in ROMA score and CPH-I
Gentry-Maharaj et al (2020)[27,]	England, Wales, Ireland	Postmenopausal women with adnexal masses (N=1590)	HE4, CA 125, individually and combined
Carreras-Dieguez (2022)[28,]	Spain	Pre- and post-menopausal women with a diagnosis of an adnexal mass based on pelvic imaging (N=1071)	HE4, CA 125, individually and combined in ROMA score and CPH-I
Lof et al (2022)[29,]	Netherlands	Pre- and post-menopausal women with a pelvic mass that was suspected of ovarian origin on ultrasound (N=316)	HE4, CA 125, individually and combined in ROMA score

CA 125: cancer antigen 125; CPH-I: Copenhagen Index; EOC: epithelial ovarian cancer; HE4: human epididymis protein 4; ROMA: Risk of Ovarian Malignancy Algorithm.

Table 9. Clinical Validity of Human Epididymis Protein 4 as a Test for Ovarian Cancer

Study	Initial N	Final N	Excluded Samples	Prevalence of Condition	Clinical Validity (95% Confidence Interval)			
					Sensitivity	Specificity	PPV	NPV
Kaijser et al (2014)[24,]	389	360	29	40%				
HE4 (≥70 pmol/L)					74%	85%		
SA					97%	90%		
Moszynski et al (2013)[25,]		253		41.4%				
HE4 (≥65 pmol/L)					80.0%	91.7%	87.3%	86.7%
SA					93.3%	90.6%	87.5%	95.1%
Nikolova et al (2017)[26,]		164						
CA 125 (≥35 U/mL)					81.8% (48.2 to 97.7)	48.7% (31.9 to 65.6)	32.1% (15.9 to 52.4)	90.0% (68.3 to 98.8)
HE4 (≥70 pmol/L)					81.8% (48.2 to 97.7)	100% (90.5 to 100)	100% (66.4 to 100)	94.87% (82.7 to 99.4)

Study	Initial N	Final N	Excluded Samples	Prevalence of Condition	Clinical Validity (95% Confidence Interval)			
Gentry-Maharaj et al (2020)[27,]	1590							
CA 125[a]					74.4% (63.8 to 83.3)		27.8% (21.8 to 34.3)	98.6% (97.8 to 99.1)
HE4[a]					67.9% (57.5 to 78.4)		26% (20.1 to 32.6)	98.2% (97.3 to 98.8)
CA 125 plus HE4[a]					75.6% (65.4 to 84.3)		28.1% (22.1 to 34.7)	98.6% (97.9 to 99.2)
Carreras-Dieguez (2022)[28,]	1071							
CA 125 (≧100 U/mL)					61.86 (55.21 to 68.09)	92.71% (90.73 to 94.29)	68.91% (62.07 to 75.02)	90.30% (88.11 to 92.11)
HE4 (≧70 pmol/L)					83.25% (77.31 to 87.88)	86.11% (83.40 to 88.43)	61.15% (55.11 to 66.87)	95.14% (93.22 to 96.53)
HE4 (≧120 pmol/L)					69.11 %(62.23 to 75.23)	96.29% (94.65 to 97.44)	83.02% (76.42 to 88.06)	92.23% (90.10 to 93.93)
Lof et al (2022)[29,]	316							
CA 125 (≥35 kU/L)					71%	50%	34%	83%
HE4 (≥70 pmol/L)					68%	65%	41%	86%
HE4 (≥150 pmol/L)					38%	96%	76%	81%

CA 125: cancer antigen 125; HE4: human epididymis protein 4; NPV: negative predictive value; PPV: positive predictive value; SA: subjective assessment.

[a]All clinical validity measures reported at a fixed 90% specificity

Study relevance and study design and conduct limitations are reported in Tables 10 and 11.

Table 10: Study Relevance Limitations

Study	Population	Intervention	Comparator	Outcomes	Duration of Follow-Up
Kaijser et al (2014)[24,]	1. 7% of data unavailable from population				
Moszynski et al (2013)[25,]					
Nikolova et al (2017)[26,]					
Gentry-Maharaj et al (2020)[27,]					
Carreras-Dieguez (2022)[28,]					
Lof et al (2022)[29,]	1. General hospital population. Also, the final histological diagnosis was missing for 30 patients (~9%)				

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.

[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.

[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose
[d] Outcomes key: 1. Study dose not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests)
[e] Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true positives, true negatives, false positives, false negatives cannot be determined).

Table 11: Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Completeness of Follow-Up[e]	Statistical[f]
Kaijser et al (2014)[24,]	2. Selection retrospective and not randomized	1. Results were not blinded				1. p –values/CI not reported
Moszynski et al (2013)[25,]	2. Selection retrospective and not randomized	1. Results were not blinded				1. p –values/CI not reported
Nikolova et al (2017)[26,]	2. Selection not randomized; small cohort	1. Results were not blinded				1. p –values not reported
Gentry-Maharaj et al (2020)[27,]						1. p –values not reported for all comparisons
Carreras-Dieguez (2022)[28,]	2. Selection retrospective and not randomized	1. Results were not blinded				1. p –values not reported for all comparisons
Lof et al (2022)[29,]	2. Selection not randomized	1. Results were not blinded				1. p –values not reported

CI: confidence interval.
The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for
interpreting tests not described; 4. Expertise of evaluators not described.
[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to followup
or missing data.
[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison to other tests not reported.

Clinically Useful

Although HE4 levels are associated with the presence of ovarian cancer, the test does not have high sensitivity or specificity. Thus it cannot be used to rule in or rule out ovarian cancer before surgery. No prospective studies were identified that compared health outcomes in patients with adnexal masses managed with and without HE4 testing, alone or in combination with CA 125 or other disease markers. There is no strong chain of evidence demonstrating that clinical decisions based on HE4 testing would improve patient outcomes.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Section Summary: Individuals with Adnexal Masses

For individuals who have adnexal masses who receive a measurement of serum biomarker HE4, the evidence includes diagnostic accuracy studies and meta-analyses. Meta-analyses have generally found that HE4 and CA 125 have a similar overall diagnostic accuracy (i.e., sensitivity, specificity), and several found that HE4 has significantly higher specificity than CA 125, but not sensitivity. Two meta-analyses had mixed findings on whether the combination of HE4 and CA 125 is superior to CA 125 alone for the initial diagnosis of ovarian cancer. The number of studies evaluating the combined test is relatively low, and publication bias in studies of HE4 has been identified. In addition, studies have not found that HE4 improves diagnostic accuracy beyond that of subjective assessment of transvaginal ultrasound. There is no direct evidence from prospective controlled studies on the impact of HE4 testing on health outcomes, and no clear chain of evidence that changes in management based on HE4 would lead to an improved health outcome.

Biomarker Human Epididymis Protein 4 Screening for Asymptomatic Individuals Not at High Risk of Ovarian Cancer

Clinical Context and Test Purpose

The purpose of testing serum biomarker HE4 levels is for diagnosis in individuals who are asymptomatic and not at high-risk of ovarian cancer.

The following PICO was used to select literature to inform this review.

Populations

The relevant populations of interest are asymptomatic individuals not at high-risk of ovarian cancer.

Interventions

The test being considered is testing serum biomarker HE4 levels. These levels are used for screening in asymptomatic individuals.

Comparators

Comparators of interest include no ovarian cancer screening (asymptomatic individuals).

Outcomes

The general outcomes of interest are OS, disease-specific survival, test accuracy, test validity, and other test performance measures.

Though not completely standardized, follow-up for individuals who are asymptomatic and not at high-risk of ovarian cancer would typically occur in the years before diagnosis.

Study Selection Criteria

For the evaluation of clinical validity of HE4 testing, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology.
- Included a suitable reference standard.
- Patient/sample characteristics were described.

- Patient/sample selection criteria were described.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
Observational Studies

Several retrospective studies aimed at determining the potential value of using HE4 and other biomarkers in early identification of ovarian cancer in asymptomatic women. Anderson et al (2010) published data on 34 women with ovarian cancer and 70 matched controls, all of whom were participating in an unrelated RCT on smokers at increased risk of lung cancer.[30,] Blood samples were available for the women between 0 years and 18 years before ovarian cancer diagnosis. In descriptive analyses, individual serum markers, including HE4, CA 125, and mesothelin, showed increasing accuracy over time approaching the diagnosis of ovarian cancer. Mean concentrations of these markers, which were measured by visually read immunoassays, began to increase approximately 3 years before diagnosis but attained detectable levels only within the final year before diagnosis. The study had a small sample size, limiting the ability to conduct quantitative analysis, and included only heavy smokers and therefore may not be representative of the population of women at risk of ovarian cancer.

Urban et al (2011) retrospectively reviewed preclinical serum samples to evaluate the potential utility of HE4 and other markers as a secondary screening test in women found to have epithelial ovarian cancer.[31,] There were samples from 112 ovarian cancer patients and 706 matched controls. Individuals participated in the Prostate, Lung, Colorectal, and Ovarian trial and had been screened annually for 6 years with CA 125. Serum samples to evaluate potential markers were taken from the year proximate to that in which women were diagnosed with ovarian cancer. Serum samples were not available for the fourth screen, so they were taken from the third year for the women diagnosed with ovarian cancer between the third and fourth screens. Investigators evaluated the associations between CA 125, HE4, and levels of 5 other markers with malignancy, accounting for increasing CA 125 levels and adjusting for demographic characteristics. Increase in CA 125 levels was associated with statistically significant increases in all of the markers. Levels of HE4 were most elevated compared to controls (i.e., the highest average HE4 level was 4.26 standard deviations above the mean HE4 level in control samples).

Terry et al (2016) retrospectively analyzed prospectively collected data from the European Prospective Investigation into Cancer and Nutrition study, a multicenter cohort study investigating the relationship between diet and cancer.[32,] The analysis used a nested case-control design. A total of 197 women who developed invasive ovarian cancer were matched with 725 randomly selected ovarian cancer-free controls. Baseline and follow-up blood samples were analyzed for levels of several biomarkers (i.e., CA 125, HE4, cancer antigen 15.3, cancer antigen 72.4) and the sensitivity, specificity, and area under the receiver operating characteristic curve were calculated. CA 125 was best able to discriminate between cases and controls within 6 months of ovarian cancer diagnosis (C statistic, 0.92), followed by HE4 (C statistic, 0.84). The ability of the markers to discriminate between cases and controls decreased with longer intervals between blood draws and cancer diagnosis. For example, with a 1- to 2-year time lag, C statistic values were 0.72 for CA 125 and 0.65 for HE4; for a 3- to 6-year time lag, the C statistic was 0.55 for CA 125. Data on HE4 were not available for the 3- to 6-year time lag analysis.

Clinically Useful

No RCTs or nonrandomized comparative studies evaluating the clinical utility of screening asymptomatic women with HE4 were identified. The studies have not estimated the sensitivity and specificity of HE4 in the screening setting, and thus the chain of evidence supporting screening is incomplete.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Section Summary: Screening Asymptomatic Individuals Not at High-Risk of Ovarian Cancer

For individuals who are asymptomatic and not at high-risk of ovarian cancer who receive screening with a serum biomarker HE4 test, the evidence includes several retrospective comparative studies and no prospective studies comparing health outcomes in asymptomatic women managed with and without HE4 screening. The retrospective studies found that HE4 levels increased over time in women ultimately diagnosed with ovarian cancer. Prospective comparative studies are needed to determine definitively whether HE4 testing is a useful screening tool.

Supplemental Information

The purpose of the following information is to provide reference material. Inclusion does not imply endorsement or alignment with the evidence review conclusions.

Practice Guidelines and Position Statements

Guidelines or position statements will be considered for inclusion in 'Supplemental Information' if they were issued by, or jointly by, a US professional society, an international society with US representation, or National Institute for Health and Care Excellence (NICE). Priority will be given to guidelines that are informed by a systematic review, include strength of evidence ratings, and include a description of management of conflict of interest.

American College of Obstetricians and Gynecologists

Guidelines from the American College of Obstetricians and Gynecologists (ACOG) on evaluation and management of adnexal masses (2016, reaffirmed 2021) state that measurement of cancer antigen 125 (CA 125) is the most extensively studied serum marker to be used in combination with imaging to determine the likelihood of malignancy.[33,] The authors also suggest that measurement of CA 125 is most useful for identification of nonmucinous epithelial cancer in postmenopausal women. Although the guideline mentions that human epididymis protein 4 (HE4) has recently been identified as a biomarker that may be useful for distinguishing between benign and malignant masses, no further recommendations regarding HE4 are provided.

In 2017 (reaffirmed 2021), a committee opinion document from ACOG and the Society of Gynecologic Oncology stated that tumor markers such as CA 125 and transvaginal ultrasound, alone or in combination, have not improved early detection or survival in women with average risk for ovarian cancer.[34,] There is also a potential for harm if surgery is performed in response to a positive test result.

National Comprehensive Cancer Network

The National Comprehensive Cancer Network (NCCN) ovarian cancer guidelines (v. 2.2023) state that, for monitoring and follow-up of patients with stage I to IV ovarian cancer with a complete response to initial treatment, "CA-125 [cancer antigen 125] or other tumor marker" should be used at "every visit if initially elevated".[35,] The guidelines do not specify any marker other than CA 125 for monitoring patients after treatment. The guidelines also recommend "CA-125 or other tumor markers as clinically indicated" for patients referred with newly diagnosed ovarian cancer after recent surgical procedure.

Elsewhere, the NCCN guidelines provides the following comment about screening using HE4: "Some evidence suggests that HE4 [human epididymis protein 4] may be a useful prognostic marker in

patients with ovarian cancer, decreases during response to treatment, and may improve early detection of recurrence relative to CA-125 alone." The NCCN guidelines currently do not recommend routine HE4 as part of preoperative workup because results vary across studies.

Several biomarker combination tests have received Food and Drug Administration approval for estimating the risk of ovarian cancer in patients with adnexal masses and planned surgery. The Risk of Ovarian Malignancy Algorithm (ROMA) test includes HE4 plus CA-125 plus menopausal status , the OVA1 test includes 5 markers including CA-125 (but not HE4), and the OVERA test includes 5 markers including both CA-125 and HE4. The NCCN guidelines state the following about using these biomarker tests: "Currently, the NCCN Panel does not recommend the use of these biomarker tests for determining the status of an undiagnosed adnexal/pelvic mass."

The NCCN guidelines state the following on screening for ovarian cancer: "Very few biomarkers have been tested prospectively to determine whether they can detect ovarian cancer or predict development of ovarian cancer in women who have no other signs or symptoms of cancer. Data show that several markers (including CA-125, HE4, mesothelin, B7-H4, decoy receptor 3 [DcR3], and spondin-2) do not increase early enough to be useful in detecting early-stage ovarian cancer."

National Institute for Health and Care Excellence
In 2011, NICE recommended using CA 125 to test for ovarian cancer in patients presenting to primary care providers with symptoms of ovarian cancer.[36,] No other biomarker tests are mentioned in the NICE guidance.

U.S. Preventive Services Task Force Recommendations
The U.S. Preventive Services Task Force updated its recommendations for screening for ovarian cancer in February 2018.[37,] The Task Force recommended against screening for ovarian cancer in asymptomatic women (D recommendation). HE4 was not specifically discussed.

Medicare National Coverage
There is no national coverage determination. In the absence of a national coverage determination, coverage decisions are left to the discretion of local Medicare carriers.

Ongoing and Unpublished Clinical Trials
Some currently ongoing and unpublished trials that might influence this review are listed in Table 12.

Table 12. Summary of Key Trials

NCT No.	Trial Name	Planned Enrollment	Completion Date
Ongoing			
NCT02595281	Determination of the Interest of HE4 as a Relapse Biomarker in Ovarian Cancers Stages IIIb, IIIc and IV After Neo-adjuvant Chemotherapy and Surgery	90	July 2022
Unpublished			
NCT01768156	Determination of the Prognostic and Predictive Value of the New Marker HE4 in Metastatic Ovarian Cancer Monitoring	101	Nov 2016
NCT03982914	The Use of a New Biomarker, HE4, in Combination With Simple Ultrasound Rules in the Prediction of Malignancy in a Pelvic Mass Detected on Ultrasound	814	Aug 2021

NCT: national clinical trial.

References

1. Surveillance Epidemiology and End Results Program (SEER). SEER Stat Fact: Ovarian Cancer. n.d.; http://seer.cancer.gov/statfacts/html/ovary.html. Accessed October 31, 2023.

2. Ledermann JA, Raja FA, Fotopoulou C, et al. Newly diagnosed and relapsed epithelial ovarian carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. Oct 2013; 24 Suppl 6: vi24-32. PMID 24078660

3. Harris HR, Guertin KA, Camacho TF, et al. Racial disparities in epithelial ovarian cancer survival: An examination of contributing factors in the Ovarian Cancer in Women of African Ancestry consortium. Int J Cancer. Oct 15 2022; 151(8): 1228-1239. PMID 35633315

4. Rustin GJ, van der Burg ME, Griffin CL, et al. Early versus delayed treatment of relapsed ovarian cancer (MRC OV05/EORTC 55955): a randomised trial. Lancet. Oct 02 2010; 376(9747): 1155-63. PMID 20888993

5. Han Y, Jiang L, Liu K, et al. Predictive Value of HE4 in Platinum-Based Chemotherapy for Ovarian Cancer : A Systematic Review. Front Oncol. 2021; 11: 703949. PMID 34307173

6. Food and Drug Administration, 510(k) substantial equivalence determination decision summary: assay only (K072939). n.d.; http://www.accessdata.fda.gov/cdrh_docs/reviews/K072939.pdf. Accessed October 31, 2023.

7. Food and Drug Administration. 510(k) substantial equivalence determination decision summary: assay only (K093957). n.d.; http://www.accessdata.fda.gov/cdrh_docs/reviews/K093957.pdf. Accessed October 30, 2023.

8. Nassir M, Guan J, Luketina H, et al. The role of HE4 for prediction of recurrence in epithelial ovarian cancer patients-results from the OVCAD study. Tumour Biol. Mar 2016; 37(3): 3009-16. PMID 26419591

9. Vallius T, Hynninen J, Auranen A, et al. Postoperative human epididymis protein 4 predicts primary therapy outcome in advanced epithelial ovarian cancer. Tumour Biol. Feb 2017; 39(2): 1010428317691189. PMID 28218038

10. Potenza E, Parpinel G, Laudani ME, et al. Prognostic and predictive value of combined HE-4 and CA-125 biomarkers during chemotherapy in patients with epithelial ovarian cancer. Int J Biol Markers. Dec 2020; 35(4): 20-27. PMID 33126819

11. Salminen L, Gidwani K, Grènman S, et al. HE4 in the evaluation of tumor load and prognostic stratification of high grade serous ovarian carcinoma. Acta Oncol. Dec 2020; 59(12): 1461-1468. PMID 33030975

12. Rong Y, Li L. Early clearance of serum HE4 and CA125 in predicting platinum sensitivity and prognosis in epithelial ovarian cancer. J Ovarian Res. Jan 04 2021; 14(1): 2. PMID 33397458

13. Samborski A, Miller MC, Blackman A, et al. HE4 and CA125 serum biomarker monitoring in women with epithelial ovarian cancer. Tumour Biol. 2022; 44(1): 205-213. PMID 36189508

14. Olsen M, Lof P, Stiekema A, et al. The diagnostic accuracy of human epididymis protein 4 (HE4) for discriminating between benign and malignant pelvic masses: a systematic review and meta-analysis. Acta Obstet Gynecol Scand. Oct 2021; 100(10): 1788-1799. PMID 34212386

15. Suri A, Perumal V, Ammalli P, et al. Diagnostic measures comparison for ovarian malignancy risk in Epithelial ovarian cancer patients: a meta-analysis. Sci Rep. Aug 27 2021; 11(1): 17308. PMID 34453074

16. Huang J, Chen J, Huang Q. Diagnostic value of HE4 in ovarian cancer: A meta-analysis. Eur J Obstet Gynecol Reprod Biol. Dec 2018; 231: 35-42. PMID 30317143

17. Dayyani F, Uhlig S, Colson B, et al. Diagnostic Performance of Risk of Ovarian Malignancy Algorithm Against CA125 and HE4 in Connection With Ovarian Cancer: A Meta-analysis. Int J Gynecol Cancer. Nov 2016; 26(9): 1586-1593. PMID 27540691

18. Macedo AC, da Rosa MI, Lumertz S, et al. Accuracy of serum human epididymis protein 4 in ovarian cancer diagnosis: a systematic review and meta-analysis. Int J Gynecol Cancer. Sep 2014; 24(7): 1222-31. PMID 25078339

19. Wang J, Gao J, Yao H, et al. Diagnostic accuracy of serum HE4, CA125 and ROMA in patients with ovarian cancer: a meta-analysis. Tumour Biol. Jun 2014; 35(6): 6127-38. PMID 24627132

20. Zhen S, Bian LH, Chang LL, et al. Comparison of serum human epididymis protein 4 and carbohydrate antigen 125 as markers in ovarian cancer: A meta-analysis. Mol Clin Oncol. Jul 2014; 2(4): 559-566. PMID 24940495

21. Yang Z, Wei C, Luo Z, et al. Clinical value of serum human epididymis protein 4 assay in the diagnosis of ovarian cancer: a meta-analysis. Onco Targets Ther. 2013; 6: 957-66. PMID 23901285

22. Ferraro S, Braga F, Lanzoni M, et al. Serum human epididymis protein 4 vs carbohydrate antigen 125 for ovarian cancer diagnosis: a systematic review. J Clin Pathol. Apr 2013; 66(4): 273-81. PMID 23426716

23. Yu S, Yang HJ, Xie SQ, et al. Diagnostic value of HE4 for ovarian cancer: a meta-analysis. Clin Chem Lab Med. Feb 03 2012; 50(8): 1439-46. PMID 22868811

24. Kaijser J, Van Gorp T, Smet ME, et al. Are serum HE4 or ROMA scores useful to experienced examiners for improving characterization of adnexal masses after transvaginal ultrasonography?. Ultrasound Obstet Gynecol. Jan 2014; 43(1): 89-97. PMID 23828371

25. Moszynski R, Szubert S, Szpurek D, et al. Usefulness of the HE4 biomarker as a second-line test in the assessment of suspicious ovarian tumors. Arch Gynecol Obstet. Dec 2013; 288(6): 1377-83. PMID 23722285

26. Nikolova T, Zivadinovic R, Evtimovska N, et al. Diagnostic performance of human epididymis protein 4 compared to a combination of biophysical and biochemical markers to differentiate ovarian endometriosis from epithelial ovarian cancer in premenopausal women. J Obstet Gynaecol Res. Dec 2017; 43(12): 1870-1879. PMID 29027715

27. Gentry-Maharaj A, Burnell M, Dilley J, et al. Serum HE4 and diagnosis of ovarian cancer in postmenopausal women with adnexal masses. Am J Obstet Gynecol. Jan 2020; 222(1): 56.e1-56.e17. PMID 31351062

28. Carreras-Dieguez N, Glickman A, Munmany M, et al. Comparison of HE4, CA125, ROMA and CPH-I for Preoperative Assessment of Adnexal Tumors. Diagnostics (Basel). Jan 17 2022; 12(1). PMID 35054393

29. Lof P, van de Vrie R, Korse CM, et al. Can serum human epididymis protein 4 (HE4) support the decision to refer a patient with an ovarian mass to an oncology hospital?. Gynecol Oncol. Aug 2022; 166(2): 284-291. PMID 35688656

30. Anderson GL, McIntosh M, Wu L, et al. Assessing lead time of selected ovarian cancer biomarkers: a nested case-control study. J Natl Cancer Inst. Jan 06 2010; 102(1): 26-38. PMID 20042715

31. Urban N, Thorpe JD, Bergan LA, et al. Potential role of HE4 in multimodal screening for epithelial ovarian cancer. J Natl Cancer Inst. Nov 02 2011; 103(21): 1630-4. PMID 21917606

32. Terry KL, Schock H, Fortner RT, et al. A Prospective Evaluation of Early Detection Biomarkers for Ovarian Cancer in the European EPIC Cohort. Clin Cancer Res. Sep 15 2016; 22(18): 4664-75. PMID 27060155

33. Eskander R, Berman M, Keder L. Practice Bulletin No. 174: Evaluation and Management of Adnexal Masses. Obstet Gynecol. Nov 2016; 128(5): e210-e226. PMID 27776072

34. Committee Opinion No. 716: The Role of the Obstetrician-Gynecologist in the Early Detection of Epithelial Ovarian Cancer in Women at Average Risk. Obstet Gynecol. Sep 2017; 130(3): e146-e149. PMID 28832487

35. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: Ovarian Cancer Including Fallopian Tube Cancer and Primary Peritoneal Cancer. Version 2.2023. https://www.nccn.org/professionals/physician_gls/pdf/ovarian.pdf. Accessed October 31, 2023.

36. National Institute for Health and Care Excellence (NICE). Ovarian cancer: recognition and initial management [CG122]. 2011; https://www.nice.org.uk/guidance/cg122. Accessed October 31, 2023.

37. U.S. Preventive Services Task Force. Recommendation Statement: Screening for Ovarian Cancer. 2018; file:///C:/Users/alt/Downloads/ovarian-cancer-final-rec-statement.pdf. Accessed October 31, 2023.

Documentation for Clinical Review

- No records required

Coding

This Policy relates only to the services or supplies described herein. Benefits may vary according to product design; therefore, contract language should be reviewed before applying the terms of the Policy.

The following codes are included below for informational purposes. Inclusion or exclusion of a code(s) does not constitute or imply member coverage or provider reimbursement policy. Policy Statements are intended to provide member coverage information and may include the use of some codes for clarity. The Policy Guidelines section may also provide additional information for how to interpret the Policy Statements and to provide coding guidance in some cases.

Type	Code	Description
CPT®	81500	Oncology (ovarian), biochemical assays of two proteins (CA-125 and HE4), utilizing serum, with menopausal status, algorithm reported as a risk score
	86305	Human epididymis protein 4 (HE4)
HCPCS	None	

Policy History

This section provides a chronological history of the activities, updates and changes that have occurred with this Medical Policy.

Effective Date	Action
07/02/2010	New Policy
01/11/2013	Policy revision with position change effective March 11, 2013
01/23/2013	Administrative update
03/11/2013	Policy revision with position change
06/30/2015	Policy title change from Human Epididymis Protein 4 (HE4) Testing
04/01/2016	Policy revision without position change
02/01/2017	Policy revision without position change
02/01/2018	Policy revision without position change
03/01/2019	Policy revision without position change
02/01/2020	Annual review. No change to policy statement. Literature review updated.
02/01/2021	Annual review. No change to policy statement. Literature review updated.
02/01/2022	Annual review. No change to policy statement. Literature review updated.
02/01/2023	Annual review. No change to policy statement. Literature review updated.
02/01/2024	Annual review. No change to policy statement. Policy guidelines and literature review updated.

Definitions of Decision Determinations

Medically Necessary: Services that are Medically Necessary include only those which have been established as safe and effective, are furnished under generally accepted professional standards to treat illness, injury or medical condition, and which, as determined by Blue Shield, are: (a) consistent with Blue Shield medical policy; (b) consistent with the symptoms or diagnosis; (c) not furnished

primarily for the convenience of the patient, the attending Physician or other provider; (d) furnished at the most appropriate level which can be provided safely and effectively to the patient; and (e) not more costly than an alternative service or sequence of services at least as likely to produce equivalent therapeutic or diagnostic results as to the diagnosis or treatment of the Member's illness, injury, or disease.

Investigational/Experimental: A treatment, procedure, or drug is investigational when it has not been recognized as safe and effective for use in treating the particular condition in accordance with generally accepted professional medical standards. This includes services where approval by the federal or state governmental is required prior to use, but has not yet been granted.

Split Evaluation: Blue Shield of California/Blue Shield of California Life & Health Insurance Company (Blue Shield) policy review can result in a split evaluation, where a treatment, procedure, or drug will be considered to be investigational for certain indications or conditions, but will be deemed safe and effective for other indications or conditions, and therefore potentially medically necessary in those instances.

Prior Authorization Requirements and Feedback (as applicable to your plan)

Within five days before the actual date of service, the provider must confirm with Blue Shield that the member's health plan coverage is still in effect. Blue Shield reserves the right to revoke an authorization prior to services being rendered based on cancellation of the member's eligibility. Final determination of benefits will be made after review of the claim for limitations or exclusions.

Questions regarding the applicability of this policy should be directed to the Prior Authorization Department at (800) 541-6652, or the Transplant Case Management Department at (800) 637-2066 ext. 3507708 or visit the provider portal at www.blueshieldca.com/provider.

We are interested in receiving feedback relative to developing, adopting, and reviewing criteria for medical policy. Any licensed practitioner who is contracted with Blue Shield of California or Blue Shield of California Promise Health Plan is welcome to provide comments, suggestions, or concerns. Our internal policy committees will receive and take your comments into consideration.

For utilization and medical policy feedback, please send comments to: MedPolicy@blueshieldca.com

Disclaimer: This medical policy is a guide in evaluating the medical necessity of a particular service or treatment. Blue Shield of California may consider published peer-reviewed scientific literature, national guidelines, and local standards of practice in developing its medical policy. Federal and state law, as well as contract language, including definitions and specific contract provisions/exclusions, take precedence over medical policy and must be considered first in determining covered services. Member contracts may differ in their benefits. Blue Shield reserves the right to review and update policies as appropriate.

Appendix A

POLICY STATEMENT (No changes)	
BEFORE	**AFTER**

Serum Biomarker Human Epididymis Protein 4
 	
Policy Statement:
  I.   Measurement of human epididymis protein 4 (HE4) is considered **investigational** for all indications.
	Serum Biomarker Human Epididymis Protein 4
	
	Policy Statement:
	  I. Measurement of human epididymis protein 4 (HE4) is considered **investigational** for all indications.