

# Guidance for Industry and Food and Drug Administration Staff

---

## Class II Special Controls Guidance Document: Ovarian Adnexal Mass Assessment Score Test System

Document issued on: March 23, 2011

For questions regarding this document contact Donna Roscoe at 301-796-6183  
([Donna.Roscoe@fda.hhs.gov](mailto:Donna.Roscoe@fda.hhs.gov)), or Marina Kondratovich at 301-796-6036  
([Marina.Kondratovich@fda.hhs.gov](mailto:Marina.Kondratovich@fda.hhs.gov)).



U.S. Department of Health and Human Services  
Food and Drug Administration  
Center for Devices and Radiological Health  
Office of In Vitro Diagnostic Device Evaluation and Safety  
Division of Immunology and Hematology Devices

# Preface

## Public Comment

You may submit written comments and suggestions at any time for Agency consideration to the Division of Dockets Management, Food and Drug Administration, 5630 Fishers Lane, rm. 1061, (HFA-305), Rockville, MD, 20852. Submit electronic comments to <http://www.regulations.gov>. Identify all comments with the docket number listed in the notice of availability that publishes in the *Federal Register*. Comments may not be acted upon by the Agency until the document is next revised or updated.

## Additional Copies

Additional copies are available from the Internet at: <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm237299.htm>. You may also send an e-mail request to [dsmica@fda.hhs.gov](mailto:dsmica@fda.hhs.gov) to receive an electronic copy of the guidance or send a fax request to 301-827-8149 to receive a hard copy. Please use the document number (1707) to identify the guidance you are requesting.

## Table of Contents

<b>I. INTRODUCTION .....</b>	<b>4</b>
<b>II. BACKGROUND .....</b>	<b>5</b>
<b>III. SCOPE .....</b>	<b>6</b>
<b>IV. RISKS TO HEALTH.....</b>	<b>6</b>
<b>V. DEVICE DESCRIPTION .....</b>	<b>7</b>
A. Background .....	7
B. Quality Systems Regulation (QS Reg) .....	7
C. Intended Use/Indications for Use.....	8
D. Test Rationale .....	8
E. Test Components and Methodology .....	8
F. Test Results .....	10
<b>VI. ANALYTICAL PERFORMANCE VALIDATION .....</b>	<b>10</b>
A. Specimen.....	11
B. Repeatability / Reproducibility .....	11
C. Linearity of Individual Analytes .....	13
D. Performance at Low Levels .....	13
E. Interference .....	14
F. Cross-reactivity/non-specific binding.....	14
G. Hook Effect of the Individual Analytes .....	14
H. Carry-Over Contamination .....	15
I. Matrix comparison .....	15
J. Stability .....	15
K. Calibration and Controls .....	15
<b>VII. SOFTWARE.....</b>	<b>15</b>
<b>VIII. CLINICAL PERFORMANCE EVALUATION.....</b>	<b>17</b>
A. Study population/samples .....	17
B. Cut-Off/ Clinical Decision Points.....	18
C. Clinical Reference Standard (“Gold Standard”).....	18
D. Study Design.....	19
E. Expected Values in Other Benign and Malignant Conditions .....	24
F. Reference Intervals .....	24
G. Relevance of the Individual Analytes Included in the Score .....	24
<b>IX. LABELING .....</b>	<b>25</b>
<b>X. REFERENCES.....</b>	<b>27</b>

# Guidance for Industry and Food and Drug Administration Staff

---

## Class II Special Controls Guidance Document: Ovarian Adnexal Mass Assessment Score Test System

*This guidance represents the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.*

### I. INTRODUCTION

This document was developed as a special controls guidance to support the classification of ovarian adnexal mass assessment score test system into class II (special controls). An ovarian adnexal mass assessment score test system measures one or more analytes in serum and combines the values into a single score that is then used to determine the likelihood that the pre-surgical adnexal mass in a woman not yet referred to an oncologist, is malignant. The test is used in conjunction with a clinical and radiological evaluation of the patient by physicians in determining whether the patient should be referred to a gynecologic oncologist for surgery.

This guidance provides recommendations to manufacturers for planning premarket notifications and labeling for ovarian adnexal mass assessment score test systems. The recommendations in this document are applicable to tests that measure separately one or more proteins obtained from whole blood preparations. The result, or score, is used by physicians as an adjunctive test to complement, not replace, other diagnostic and clinical procedures. A woman for whom surgical intervention is planned should be referred to a gynecologic oncologist when either the physician's independent pre-surgical assessment, or the ovarian adnexal mass assessment score, or both, suggest the likelihood of malignancy.

An ovarian adnexal mass assessment score test system is not indicated as a screening test or for the diagnosis of ovarian cancer. (Refer to Section IX for additional provisions required in

labeling associated with these tests.) It is intended for use in those patients for whom surgery is planned, and should not be used to decide whether or not a patient should receive surgery. This guidance does not apply to gene expression assays or tissue-based assays.

This guidance is issued in conjunction with a Federal Register notice announcing the classification of ovarian adnexal mass assessment score test system. Any firm submitting a 510(k) premarket notification for an ovarian adnexal mass assessment score test system will need to address the issues covered in this special controls guidance. The firm must show that its device addresses the issues of safety and effectiveness identified in this guidance, either by meeting the recommendations of this guidance or by some other means that provides equivalent assurances of safety and effectiveness.

Designation of this document as a special control means that any firm currently marketing, or intending to market, ovarian adnexal mass assessment score test system will need to address the issues covered in this special controls guidance. The firm will need to show that its device addresses the issues of safety and effectiveness identified in this guidance, either by meeting the recommendations of this guidance or by some other means that provide equivalent assurances of safety and effectiveness.

## **II. BACKGROUND**

Physicians routinely find pelvic adnexal masses in women of all ages, either incidentally during the course of a standard gynecological evaluation, or following an examination due to the woman's presentation of symptoms. Approximately 5 to 10% of women will undergo surgery for a suspected ovarian malignancy, and 13 to 21% of these masses will be diagnosed as ovarian cancer (Ref. 1). Guidelines for the differential diagnosis and management of patients with adnexal masses have been established, and include the referral to a gynecological oncologist for women with suspected ovarian cancer. Studies have shown that patients with ovarian cancer have improved progression-free survival and overall survival when the surgery is performed by gynecologic oncologists as opposed to general gynecologists and surgeons. These published observations and guidelines support the clearance of tests that augment patient referral to a gynecologic oncologist through the supplemental assessment of malignancy.

A manufacturer who intends to market a device of this generic type must:

- conform to the general controls of the Federal Food, Drug, and Cosmetic Act (FD&C Act), including the premarket notification requirements described in 21 CFR 807 Subpart E,
- conform to the special control developed for this device, by addressing the specific risks to health associated with the ovarian adnexal mass assessment score test system identified in this guidance, and
- obtain a substantial equivalence determination from FDA prior to marketing the device. (21 CFR 807.81, 21 CFR 807.87, and 21 CFR 807.100).

FDA believes that special controls, when combined with the general controls of the act, are sufficient to provide reasonable assurance of the safety and effectiveness of these devices.

This special control guidance document identifies the classification regulation and product codes for the ovarian adnexal mass assessment score test system (please refer to Section III Scope). Other sections of this guidance document provide recommendations to manufacturers on addressing risks related to these devices.

### **III. SCOPE**

The scope of this document is limited to the following device described in 21 CFR 866.6050 (product code ONX):

21 CFR 866.6050 An ovarian adnexal mass assessment test system is a device that measures one or more proteins in serum. It yields a single result for the likelihood that an adnexal pelvic mass in a woman, for whom surgery is planned, is malignant. The test is for adjunctive use, in the context of a negative primary clinical and radiological evaluation, to augment the identification of patients whose gynecologic surgery requires oncology expertise and resources.

### **IV. RISKS TO HEALTH**

The ovarian adnexal mass assessment score test system is not indicated for use as a screening or diagnostic test for ovarian cancer. Failure of the assay to perform as indicated could lead to inappropriate assessment and improper management of patients with ovarian malignancies. Specifically, a falsely low ovarian adnexal mass score could result in a determination that the patient may not have ovarian malignancy, which could lead to less than optimal surgical expertise and resources. A falsely high ovarian adnexal mass score could result in a determination that the patient may have ovarian malignancy which could lead to inappropriate surgical decisions and unnecessary patient anxiety. Off-label use of the test (e.g., in patients who are not already identified as needing surgery for pelvic mass or without reference to an independent clinical/radiological evaluation of the patient), may lead to a high frequency of unnecessary further testing and surgery due to false positive results, or to delay in tumor diagnosis due to false negative results.

In the table below, FDA has identified the risks to health generally associated with the use of this device. The measures recommended to mitigate the identified risks are described in this guidance document, as shown in the table below. You should conduct a risk analysis, prior to submitting your premarket notification, to identify any other risks specific to your device. Risks may vary depending on the detection and measurement method used. The premarket notification should describe the risk analysis method. If you elect to use an alternative approach to address the risks identified in this document, or have identified risks additional to those in this document, you should provide sufficient detail to support the approach you have used to address that risk.

Identified risk	Recommended mitigation measures
False negative result	Section VI-IX
False positive result	Section IX
Off-label use as a screening test, stand-alone diagnostic test, or as a test to determine whether or not to proceed with surgery.	Section IX Black Box warning

## V. DEVICE DESCRIPTION

### A. Background

We recommend that you identify your device by regulation and product code described in Section III, above. You must identify a legally marketed predicate device (21 CFR 807.87(f)). You should outline in a table the similarities and differences between the predicate and your device so that FDA can efficiently determine whether comparisons can be made between your device and the predicate and identify aspects of your device that may or may not need additional performance studies.

The values obtained from the multiple analytes in your test are derived from individual assays for each analyte. If your test uses values obtained from individual assays sold by other manufacturers (materials required but not provided), you should indicate the classification of the test, whether the assays have been cleared or approved by the FDA (i.e., formally reviewed in the 510(k) or PMA process as a class II or class III device, respectively), and the intended use for which they have been cleared or approved.

Individual assays sold by other manufacturers (i.e., items you have listed as materials required for assay results, but are sold separately by other manufacturers), are considered components of your device. For cleared or approved assays, you should provide a copy of the FDA reviewed labeling and summarize the analytical performance characteristics for each assay. If the assay(s) sold by other manufacturers have not been cleared or approved (e.g., labeled class I exempt), you should also provide the manufacturers labeling, summarize the information requested in this section. However, you will need to conduct and submit the analytical performance studies for these assays separately (refer to Section VI Analytical Performance Validation for more information) and have them reviewed as components of your class II device.

Your submission should adequately describe the following features of your ovarian adnexal mass assessment score test system:

### B. Quality Systems Regulation (QS Reg)

All manufacturers of in vitro diagnostic tests must adhere to the requirements in 21 CFR Part 820, Quality Systems Regulation. If you have designed an algorithm to be used with tests that are purchased from other manufacturers, you are responsible for those assays to the

extent that they are components of your device, whether or not the assay has been previously cleared by FDA. Your responsibilities include, but are not limited to, device design, equipment, purchase and handling of components, production and process controls, packaging and labeling control, device evaluation, installation, complaint handling, servicing, and records. You should provide a statement of compliance with the Quality Systems Regulation for your device, including all components, and a brief description of how you have, or will, accomplish this prior to marketing of your test (i.e., through established relationships with the manufacturer(s) of other components). While establishing compliance with Quality Systems Regulation is not part of a 510(k) premarket notification review, FDA seeks assurance from manufacturers of these types of assays that they have adequately assumed responsibility for all aspects of the test prior to clearance.

### **C. Intended Use/Indications for Use**

Your submission must include an intended use/indications for use statement that summarizes how you, the manufacturer, intend the product to be used, and the clinical purpose of the test (21 CFR 807.87(e) and 21 CFR 807.92 (a)(5)). The intended use for ovarian adnexal assessment score test systems should include the name of the test, a description of the nature of the test result output, the specimen type, the test method, the target population (may be limited by demographics of the study population e.g., age, race, and clinical conditions such as menopausal status and stage of the disease). Ovarian adnexal assessment score test systems are indicated for women with an ovarian/adnexal mass for which surgery is planned, and not yet referred to gynecologic oncologists, as an aid to further assess the likelihood that malignancy is present when the physician's independent clinical and radiological evaluation does not indicate malignancy.

### **D. Test Rationale**

Provide a summary of the test analytes and the rationale behind including the measurement of each analyte in the assay. Published information may be submitted in support of the individual analytes for the test's indication for use. Include summaries of any unpublished studies which lead to the conclusion to include the individual test analytes. Your description should include relevant information about different protein states as a result of different RNA splicing, or post-translational modifications, and how these may be different for the population with disease under investigation when compared to the non-diseased population.

### **E. Test Components and Methodology**

You should describe in detail the reagents, assay format/methodology, instruments and software used in your device.

#### ***1. Test Reagents***

You should provide a description of all reagents and components (including calibrators, controls, instruments) provided or recommended for use. Include a description of the source of each reagent (e.g., mouse, cell line), its purification method, and the verification process

for use in the test. You should provide certificates of analysis if the reagent is obtained from an outside vendor. If the reagents and components in your test consist of individual test kits supplied by other manufacturers, we consider these test kits reagents in your test and the specific manufacturer test kit and instrument for use with your test should be specified in the labeling. A summary of this information in table format for the individual assays is requested.

## ***2. Test Methodology***

You should provide a description of the test methodology employed by your device. This should include test platform(s) and method of measurement. A brief table summarizing this information for all assays included in the test is requested. For devices that include novel analytes or immunoassays never reviewed and cleared or approved by FDA, additional information is requested:

- Explain how you screened, selected, and determined the optimal antibody-analyte-antibody combinations for capture and detection of the target analytes in the test.
- Methods used to attach the capture antibodies to the substrate. Include a description of how this was optimized (e.g., maximum antibody and antigen concentrations)
- Description of the secondary antibody conjugates including the labels and conjugation procedures.
- Reaction components and conditions, washing procedures, and signal detection components and methods and a description of how these conditions were optimized including antibody concentrations.
- Description of how cross-reactivity and non-specific binding are minimized in the test. Provide Western blots as evidence.
- Sample collection requirements and sample handling from time of collection to use in the test, including any requirements (e.g., preservatives) for ensuring stability of the analytes.
- Description of all controls and calibrators and how they function in the system.
- Provide a description of how background signals are minimized or normalized (e.g., ratios of signal to background).
- Instrumentation and instrumentation software required for your device, including the components and their function within the system.

### **3. *Score Algorithm***

You should provide

- A description of the training data set(s) (inclusion/exclusion criteria, description of clinical sites, number of subjects, prevalence of malignancy, and so on);
- A description of the classifier/algorithm development including the sample types and the statistical models and techniques used;
- A description of performance measures (internal validation).

If the cutoff for the score was selected in the training data set(s), you should provide information about how the cutoff was determined.

You should provide a brief description of the final computational method indicating how the individual analytes values outside of measuring intervals were used in the calculation of the score values and software employed to obtain the score result.

If the algorithm for calculation of the score test results is using the clinical decision points for the individual analytes in your test, you should provide information about these cutoffs. For example, your test may incorporate an analyte typically used as an aid to diagnose inflammation. The cut-offs recommended for this analyte may be different in women with benign and malignant pelvic masses. You should explain whether known reference intervals were used or values were obtained from literature or findings from your research studies.

### **F. Test Results**

You should state the description of the nature of the test result output (e.g., patient classification or a continuous numerical value to which a cut-off is applied) and interpretation of the score result. You should provide examples of the test reports (e.g., printouts) that are generated for the clinician.

## **VI. ANALYTICAL PERFORMANCE VALIDATION**

In your 510(k), you should detail the study design you used to evaluate each of the performance characteristics outlined below. All analytical performance studies should be conducted using the final version of your ovarian adnexal assessment score test system device.

For each of the analytical performance studies described below, you should state your predetermined acceptance criteria for each analyte individually and the impact on the overall result.

If your test uses values obtained from individual quantitative assays sold by other manufacturers (materials required but not provided), that have been cleared or reviewed by FDA, you may not need to submit additional performance studies beyond precision, as

described below, for these assays individually, provided the ovarian adnexal mass assessment score test system is based on the performance characteristics described in the FDA reviewed labeling for these tests. For example, if your test does not incorporate individual quantitative values outside of the claimed measuring interval for the individual assay, you will not need to conduct additional linearity studies. Tests that have not been reviewed by FDA (e.g., class I exempt) should demonstrate analytical performance as components of your class II device. It is important to note that these individual class I exempt assays are not receiving separate clearance as class II assays, nor should they be considered as having been reviewed for their class I exempt intended use. The analytical performance studies for these assays are demonstrating these assays are fit for the purpose of use in the ovarian adnexal mass assessment score test system.

FDA intends to limit the scope of the clearance of a test to the specific test kits and instruments evaluated.

#### **A. Specimen**

*Pre-analytical factors:* If your test includes novel analytes (e.g., analytes that have never been incorporated into an in vitro lab test before) you should indicate whether specific instructions for blood collection (e.g., position of patient during blood collection) and blood collection tube handling are required (due to labile nature of the analyte).

*Stability:* You should demonstrate the stability of the specimens across the extremes of these parameters (e.g., temperature, time to freezing, freeze-thaw, and shipping) for use in your test. You should describe how you selected the acceptance criteria for each analyte. The specimen stability claims for your test are limited to the least stable analyte in your score test system.

#### **B. Repeatability / Reproducibility**

You should provide a description of how you assessed and determined the acceptable variability in the individual analytes based on the acceptable impact of variability on the test score overall.

We recommend you provide an evaluation of the precision of your score test system with samples using samples that span the range of the score test results. The CLSI documents EP5-A2 “Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline” and EP12-A2 “User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline” include guidelines that may be helpful for developing design and computations of the data in the precision studies.

The samples in the precision study should span the range of the score numerical values; you should include a few samples (3-5 samples) with score values close to the cutoff(s) of the score test due to different combinations of the analytes.

Ideally, you should identify all sources of the score test variability and include them in the precision study. You should provide a demonstration of the precision of your score test across three laboratories and provide an evaluation of the repeatability (within-run precision), between-run, between-day, between-operator, and between-site components of imprecision. Include a detailed description of the number of days, number of operators, assays, instruments, lots, and calibration cycles evaluated in the study.

You should provide the acceptance criteria and demonstrate the precision for the test score using samples that span the range of test score results. You should include the data from the precision studies for each analyte as well, to demonstrate that the precision meets predefined acceptance criteria for each analyte. You should indicate how you concluded that the allowable analyte imprecision would not diminish the accuracy of the index value reported. You should also provide an evaluation of lot-to-lot precision using 3 different test lots. This includes multiple lots of each individual assay, calibrators, and controls that comprise your test system.

In addition to the precision studies for the test score system described above, you should provide a simulation of possible results for test score system precision based on the precision profiles of each individual analyte. The usual precision study provides information about precision for some particular combinations of the amounts of individual analytes that were present in the samples of the precision studies described above. There are, however, many possible combinations of the amounts of individual analyte that give the same value of the test score but have different precisions. Acknowledging that it would be impossible to evaluate the precision of all possible combinations of analytes, the additional simulation provides information about possible precision profile of the test score system for different combinations of individual analyte values. If the simulation predicts an unacceptable level of precision at the clinical decision point, it may be important to evaluate contrived samples reflective of that particular scenario.

An example of one such simulation method you may elect to use is presented here. The precision data from previously performed precision studies (for already cleared/approved analytes and for the data provided for the novel analytes) should be used for building precision profiles of each individual analyte. The precision profile for repeatability, and the precision profile for within-laboratory precision for each individual analyte, should be constructed by performing linear interpolation using the known precision data from the repeatability and within-laboratory precision studies with actual samples. For each possible combination of the values of the individual analytes, estimate the value of the score corresponding to this combination of values of analytes and repeatability and within-laboratory precision of the score based on the corresponding precision profiles. Because the score is based on separate measures of individual analytes in a sample, random measurement errors of each analyte can be considered as uncorrelated. The basic steps of additional statistical simulations are the following (for sake of simplicity, consider two individual analytes  $X_1$  and  $X_2$  with  $\text{Score}=F(X_1, X_2)$  and repeatability precision data):

- 1) Provide repeatability precision results (mean value, standard deviation (SD), and percentage coefficient of variation (%CV)) from previously performed precision

- 2) Consider a combination of two analytes with values  $X_1=U$  and  $X_2=V$ . Using repeatability precision profiles, obtain  $SD_1(U)$  for  $X_1=U$  and  $SD_2(V)$  for  $X_2=V$ .
- 3) Generate  $X_1^*$  using normal distribution with mean value of  $U$  and standard deviation of  $SD_1(U)$  and generate  $X_2^*$  using normal distribution with mean value of  $V$  and standard deviation of  $SD_2(V)$ . Calculate  $Score^*=F(X_1^*,X_2^*)$ . After performing this step  $K$  times (for example, 100), calculate the mean value of score of  $K$  measurements  $Score^*_{mean}$  (corresponding to mean value of the score for  $X_1=U$  and  $X_2=V$ ) and standard deviation  $SD$  and  $\%CV$  of the  $K$  score measurements .
- 4) Provide repeatability precision profile for the Score: values of the mean score  $Score^*_{mean}$  with the  $SD$  and  $\%CV$  from the previous step for all possible combinations of  $U$  and  $V$  for which precision profiles are available. Repeatability precision profile should be provided in the form of table (EXCEL) and graphically (X-axis is mean value  $Score^*_{mean}$  and Y-axis is corresponding  $\%CV$ ).

Perform similar statistical simulations for evaluation of the within-laboratory precision profile for the Score using within-laboratory precision profiles of individual analytes.

### **C. Linearity of Individual Analytes**

A demonstration of linearity for each individual analyte is based on the measuring range incorporated into the algorithm. For test score systems that use individual immunoassays, measuring range claims that extend beyond the cleared/approved analyte test range you should provide a new demonstration of linearity. Uncleared/unapproved test analytes should demonstrate linearity as well. You should indicate if the values obtained with your patient population are likely to fall below the claimed measuring range of the individual analyte assay and how you control for the impact of out-of range results on your score. We recommend you refer to CLSI document EP6-A “Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline” for more information about conducting linearity studies for individual analytes.

#### ***Score Values Range***

You should provide information about range of the numerical score values of the test score system based on the measuring ranges of the individual analytes. If, in addition to providing the qualitative test score results based on the cutoff(s), you plan to report the numerical values of the test score, the data should demonstrate that the higher numerical values of the test score system are related to the progressively higher or progressively lower probabilities of malignancy.

### **D. Performance at Low Levels**

You should demonstrate the limit of detection and limit of quantitation for any uncleared/unapproved analyte tests and for any changes that increase the lower end measuring range claims for cleared/approved analyte tests. You should explain how the low level values related to limit of detection and/or limit of quantitation are incorporated into the

algorithm such that results outside of the measuring interval are not imported and do not yield a test result. We recommend you refer to CLSI document EP17-A “Protocols for Determination of Limits of Detection and Limits of Quantitation; Approved Guideline” for more information about conducting limit of detection and limit of quantitation studies for individual analytes.

#### **E. Interference**

Substances that interfere with any of the analytes in your test are likely to interfere with the test result. You should indicate (preferably in table format stating the concentrations evaluated) whether any of the analytes are subject to interference by hemoglobin, bilirubin (conjugated and unconjugated), triglycerides, total protein, heterophilic antibodies (HAMA) and rheumatoid factor. For the interferents described above, you should demonstrate the % difference in assay results by comparing a sample with interferent to the same sample without interferent for its impact on the index overall along with 95% confidence interval. Reporting the % difference for each analyte from this analysis is helpful as well. Ideally, the analyte concentrations evaluated would be near the clinical decision points for the test score system.

You should indicate whether any known sources of interference occur for the analytes in your test and, if so, demonstrate the impact of that interferent on the score.

You should demonstrate that common medications do not interfere with the test.

#### **F. Cross-reactivity/non-specific binding**

An evaluation of known cross-reactants and their potential impact on the test score system should be performed. The test can be influenced by several factors such as effects on ligand binding due to antibody immobilization to a substrate, nonspecific adsorption of proteins, and the influence of other proteins in the matrix. If your test is a multiplex immunoassay, you should demonstrate that cross-reactivity, non-specific binding, and cross-interference between the analytes does not occur. Indicate whether there are any potential cross-reactants for the analytes in your test. You should also provide a demonstration that the detection of the analytes by your antibodies is specific. Western blots should be provided toward this demonstration.

#### **G. Hook Effect of the Individual Analytes**

When applicable, you should demonstrate that excess analyte does not cause a hook (prozone) effect. This demonstration should be performed for uncleared/unapproved individual assays (components) of your test and for each analyte in a score test system.

## **H. Carry-Over Contamination**

You should provide a description of the potential for carry-over contamination for test systems that use previously cleared instrumentation.

### **I. Matrix comparison**

For some analytes, matrix effects can occur when testing plasma samples with various anticoagulants, which lead to changes in the performance of the test. If your test recommends more than one sample type, you should evaluate the possibility of matrix effects on the test. The impact of matrix effects should be presented for each individual analyte in the test score and for its impact on the score result overall.

### **J. Stability**

You should describe your study design for determining the real-time stability of the reagents and instruments and, if applicable, for open vial and on-board stability. Your stability studies should include information about the times, temperatures, and storage of your test system and reagents. For each study, you should provide your acceptance criteria and a description of how you selected the acceptance criteria values (i.e., concluded the limit of the acceptance criteria did not impact the results). We recommend you refer to CLSI document EP25-A “Evaluation of Stability of In Vitro Diagnostic Method Products; Approved Guideline” for more information about conducting stability studies for individual analytes.

### **K. Calibration and Controls**

For ovarian adnexal mass assessment score test systems whose components are made up of cleared/approved individual immunoassays, the calibrators and controls for each assay should be described based on their use in the assay.

For all other tests, you should describe the following for your control and calibration materials:

- The nature and function of the various controls that you include with, or recommend for, your system.
- The methods for value assignment and validation of control and calibrator material. Include certificates of analysis if any reagents incorporated into your test system are supplied by a vendor.
- The control parameters that could be used to detect failure of the instrumentation to meet required specifications.

## **VII. SOFTWARE**

You should provide detailed information about the software used in your device in accordance with the level of concern. For additional information refer to the FDA document “Guidance for the Content of Premarket Submissions for Software Contained in Medical

Devices.”<sup>1</sup> You should determine the level of concern prior to the mitigation of hazards. In vitro diagnostic devices of this type are typically considered a moderate level of concern because software flaws could result in false results reported to clinician and patient, which could cause harm to the patient.

You should include the following points, as appropriate, in preparing software documentation for FDA review:

- Full description of the software design. Your software should not include utilities that are specifically designed to support uses beyond those in your intended use. You should also consider privacy and security issues in your design. Information about some of these issues may be found at the following website regarding the Health Insurance Portability and Accountability Act (HIPAA)  
<http://www.hhs.gov/ocr/privacy/hipaa/understanding/index.html>.
- Hazard analysis based on critical thinking about the device design and the impact of any failure of subsystem components, such as signal detection and analysis, data storage, system communications, and cybersecurity in relationship to incorrect patient reports, instrument failures, and operator safety.
- Documentation of complete verification and validation (V&V) activities for the version of software that will be submitted to demonstrate substantial equivalence. You should also submit information regarding validation of the compatibility of test software with any instrumentation software.
- If the information you include in the 510(k) is based on a version other than the release version, identify all differences in the 510(k) version and detail how these differences (including any unresolved anomalies) impact the safety and effectiveness of the device.

Below are additional references to help you develop and maintain your device under good software life cycle practices consistent with FDA regulations.

- General Principles of Software Validation; Final Guidance for Industry and FDA Staff; available on the FDA Web site at:  
<http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm085371><http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm085371.pdf>.
- Guidance for Off-the-Shelf Software Use in Medical Devices; Final; available on the FDA Web site.<sup>2</sup>
- [21 CFR 820.30](#) Subpart C – Design Controls of the Quality System Regulation.

---

1

<http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm073779.pdf>

<sup>2</sup><http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm073779.pdf>

- ISO 14971-1; Medical devices - Risk management - Part 1: Application of risk analysis.
- AAMI SW68:2001; Medical device software - Software life cycle processes.

## VIII. CLINICAL PERFORMANCE EVALUATION

The data from your clinical studies should support the indications for use and claims for your device. The clinical validation study should use patient samples that are obtained from the intended use population and that are different from the training sets (specimens you used to develop the algorithm (e.g., score)). You should describe the protocol of each clinical study, including the inclusion and exclusion criteria, study design, statistical analysis method, and statistical justification of the sample size. You should submit the data with the values of the individual analytes along with the score test results from your clinical validation studies.

### A. Study population/samples

The intended use population for an ovarian adnexal mass assessment score test system consists of those patients with pelvic masses known to require surgery having undergone an evaluation in a primary care setting (i.e., gynecologist, internist, family practitioner but not a gynecologic oncologist). The ovarian adnexal mass assessment score test system, in conjunction with pre-surgical clinicopathologic information, augments the identification of patients whose gynecologic surgery requires oncology expertise and resources. You should provide your inclusion and exclusion criteria. Patients should be representative of the intended use population. FDA recommends that you enroll patients from several distinct geographical locations within the U.S. population. Samples used in the training sets should not be included as part of your validation set.

You should plan to evaluate your results in both the pre-menopausal and post-menopausal women separately and combined. You should indicate how menopausal status will be identified. If self-identified, you should provide patients with a definition and provide a plan for those cases that do not provide an answer. It is preferable to have an objective method for defining menopausal status that can be applied uniformly to the patient population. For example, age, date of last menses or follicle stimulating hormone (FSH) serum levels.

For all samples you should provide summary patient information (age, race/ethnicity, menopausal status, current medical conditions) overall and by enrollment site.

You should provide justification for the number of samples used in the study. Provide a detailed accounting of samples that were excluded and the specific reasons for exclusion. Samples should not be excluded based on post-surgical findings.

While samples collected in a prospective clinical study are preferred, well-characterized retrospective samples from specimen banks may be used in your clinical validation study, provided the following conditions are met in addition to those listed above:

- The samples in your study are from patients who are representative of your intended use population.
- There are no biases due to selection methods (i.e., sample procurement should be with consideration to time and sites)
- There are no biases due to analytical artifacts (e.g., due to storage conditions, multiple freeze-thaws).
- Specimens retrieved from the bank meet predefined criteria in a sample collection protocol.
- Samples are annotated with the following information: patient demographics (age, menopausal status) pre-surgical assessment (malignant vs. benign) by the non-gynecologic oncologist and surgical pathology (histological diagnosis, and if malignant, tumor stage).

We recommend you consult with FDA prior to performing validation studies using banked samples.

## **B. Cut-Off/ Clinical Decision Points**

In your submission, you should explain how the cut-off (the value used to distinguish the probable presence of malignancy versus the absence of malignancy) was determined. Selection of the appropriate clinical cutoff can be justified by the relevant levels of sensitivity and specificity that are based on Receiver Operating Curve (ROC) analysis of training or/and pilot studies with clinical samples. The clinical performance at the selected clinical cutoff is easily estimated using a pivotal clinical study (validation data set). In some circumstances, the clinical cutoff can be determined during the pivotal clinical study using an unbiased procedure and an appropriate sample size. If the level of sensitivity (or specificity) that is clinically acceptable is pre-specified then the pivotal study can be used to establish the clinical cutoff corresponding to the pre-specified level of sensitivity and to obtain an unbiased estimation of the clinical performance of the score test with this selected cutoff (Ref. 2). If the test has a range of results for which retesting is recommended or for which a determination of a “positive or negative” result cannot be made (i.e., equivocal zone), you should explain how you determined the limits of the equivocal zone. You should also justify the clinical implications for patients whose samples give equivocal results.

If your test has one cutoff for use with all patients, you should demonstrate that the cut-off is appropriate for both pre-menopausal and post-menopausal women. We recommend you investigate the need to select cut-offs specific to menopausal status in your pilot (or training) studies. If menopausal status is relevant to the interpretation of the test results, you should plan to adequately represent the distinct menopausal groups and validate cut-offs based on menopausal status.

## **C. Clinical Reference Standard (“Gold Standard”)**

To evaluate the performance of the ovarian adnexal mass assessment score test system in distinguishing a benign or malignant adnexal mass, the result of the test should be compared

to histopathological information obtained following surgery. For each patient/sample, you should indicate whether the mass was benign or malignant, and when malignant indicate the pathological diagnosis (i.e., epithelial ovarian cancer, other primary ovarian malignancy, ovarian malignancy of low malignant potential (LMP), non-primary ovarian malignancies with involvement of the ovaries, or non-primary ovarian malignancies with no involvement of ovaries), the tumor stage and histology (serous, mucinous, endometrioid etc.). The classification of pathological findings into two categories (i.e., malignant and non-malignant) should be comprehensive and prespecified. You should state who ultimately performed the surgery (gynecologic oncologist (GO) or physician other than a gynecological oncologist (non-GO)).

#### **D. Study Design**

Because the ovarian adnexal mass assessment test system is used in conjunction with the clinical evaluation of patients presenting ovarian masses *selected for surgery* but not yet referred to a GO, it is essential to have a well-organized and complete accounting of the pre-surgical clinical evaluations by the non-GOs. For each patient in the clinical study, pre-surgical and pre-referral clinicopathologic information (e.g., patients' symptoms, physical findings, imaging, CA 125 value) should be collected and integrated into the statement of a binary pre-surgical assessment identifying the mass as "Benign" or "Malignant."

The evaluation of the score test as an aid in the evaluation of the patients in addition to the pre-surgical information is accomplished by considering a combination of the pre-surgical clinical assessment made by the non-GO and the results of your test. This is referred to as the "OR" decision rule, i.e., a case would be considered positive if either the non-GO presurgical assessment is positive, *or* if the result of the test find it positive. Using this rule, if the score test result is negative and pre-surgical information is positive for a patient, then the patient should be considered to have a high probability for ovarian malignancy.

Likewise, if the score test result is positive and the pre-surgical assessment is negative for a woman, the woman should be considered to have a high probability for ovarian malignancy. Use in this manner is designed to improve the referral to gynecological oncologists for patients with malignant pelvic masses while still assuring referral for all women who would otherwise be referred based on pre-surgical information alone (Ref. 3).

In order to demonstrate that the score test provides additional information beyond pre-surgical assessment alone, you should compare the performance of pre-surgical assessment alone and performance of the pre-surgical assessment and score test when combined by the "OR" rule. Positive predictive values (PPV) and negative predictive values (NPV) (or, equivalently, positive and negative likelihood ratios) are the basis for this comparison. For additional information, see Biggerstaff.<sup>3</sup>

You should demonstrate that your test provides additional information for biologically relevant subpopulations (e.g., pre-menopausal, post-menopausal) or provide an acceptable justification for why such a demonstration is not needed.

---

<sup>3</sup> Biggerstaff, B.J. Comparing diagnostic tests: a simple graphic using likelihood ratios. *Statistics in Medicine* 2000, 19: 649-663.

Consider a general scheme of comparison of test  $T_1$  and combination OR of tests  $T_1$  and  $T_2$ . The sensitivity for the “OR” combination is at least as large as the sensitivity for  $T_1$  alone. The specificity for the “OR” combination is the same or worse than the specificity for  $T_1$  alone. Thus, the combination “OR” has an inherent trade-off between sensitivity and specificity. An increase in combined sensitivity alone does not prove that the combination (OR) of the  $T_1$  and  $T_2$  tests is effective if the combined specificity is shown to decrease appreciably.

The data of the clinical study should demonstrate that

- There is a statistically and clinically significant improvement in NPV with the combination OR (pre-surgical assessment and the score test) vs. NPV of the pre-surgical assessment alone; and
- If there is a loss in PPV with the combination OR of the pre-surgical assessment and the score test vs. PPV of the pre-surgical assessment alone, this loss in PPV should be clinically acceptable.

The logic for these success criteria is developed below. Consider a straight line connecting the point (0,0) through the point corresponding to test  $T_1$ , on a plot of Sensitivity vs. 1-Specificity (see Figure 1 below). This line denotes performance characteristics for tests that have the same PPV as test  $T_1$ . A straight line connecting the point (1,1) through the point corresponding to test  $T_1$  denotes the performance characteristics of tests that have the same NPV as test  $T_1$ <sup>4</sup>. In comparing the performance of test  $T_1$  with the performance of the OR combination of tests  $T_1$  and  $T_2$ , there are three possible scenarios:

Scenario A: Both predictive values (PPV and NPV) for “OR” combination ( $T_1$  or  $T_2$ ) are larger than predictive values of  $T_1$  alone (see the green region). Thus, it is easy to draw the conclusion that the combination OR is better than the  $T_1$  alone.

---

<sup>4</sup> PPV depends on a positive likelihood ratio (PLR) and prevalence of malignancy, and NPV depends on negative likelihood ratio (NLR) and prevalence. For a comparison of two tests within the same population, comparison of PPV and NPV is equivalent to comparison of PLR and NLR.

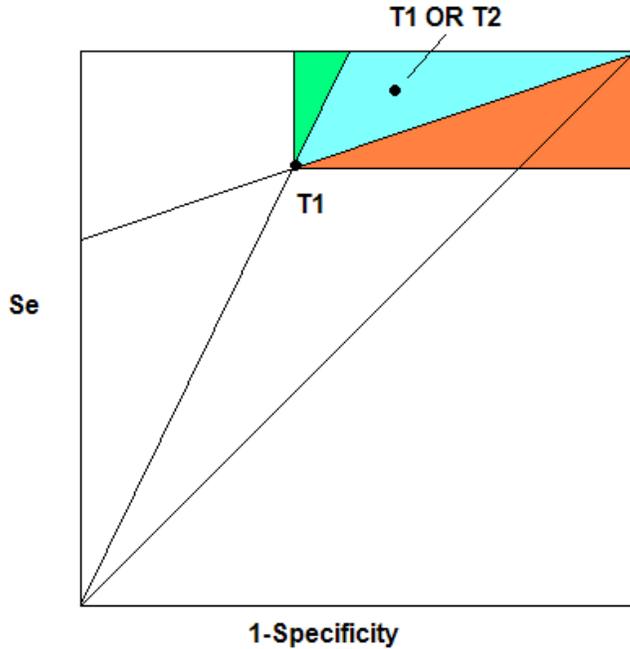


Figure 1

Scenario B: The PPV of combination OR is worse than PPV of  $T_1$  but the NPV of combination OR is better than NPV of  $T_1$  (see the Blue region). In this region there is a trade-off between the amount by which NPV increased and the amount by which PPV decreased. Success can be concluded if the lowered PPV remains consistent with safe and effective use of the test.

Scenario C: Both PPV and NPV of combination OR are worse than the PPV and NPV of test  $T_1$  (see the red region). Thus, it is easy to draw the conclusion that the combination OR is worse than the  $T_1$  alone.

We recommend you summarize the results of the clinical study based on pathology results in tables similar to the ones below and provide an assessment of probability of malignancy (along with 95% CI) based on the various outcomes shown in the tables below:

		Non-GO Pre-surgical Assessment		
		Positive	Negative	
Score test	Positive	A	B	
	Negative	C	D	
				N

Present the performance of the score test and pre-surgical assessment for the subjects with malignancy by pathology and for subjects with no malignancy by pathology separately.

Malignancy by Pathology				
		Non-GO Pre-surgical Assessment		
		Positive	Negative	
Score test	Positive	A <sub>1</sub>	B <sub>1</sub>	
	Negative	C <sub>1</sub>	D <sub>1</sub>	
				N <sub>1</sub>

No Malignancy by Pathology				
		Non-GO Pre-surgical Assessment		
		Positive	Negative	
Score test	Positive	A <sub>0</sub>	B <sub>0</sub>	
	Negative	C <sub>0</sub>	D <sub>0</sub>	
				N <sub>0</sub>

The Table below shows performance characteristics for the test applied to all subjects evaluated by non-GO physicians. For Single Assessment, only the pre-surgical assessment is used, without reference to a score test result. For Dual Assessment (i.e., “OR” combination) the adnexal mass is declared potentially malignant if the pre-surgical clinical assessment, the score test, or both were positive.

Performance	Single Assessment (Pre-surgical Assessment)	Dual Assessment (Pre-surgical Assessment “OR” Score Test)
Sensitivity		
Specificity		
PPV		
NPV		
Prevalence		

- Provide sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) along with the 95% confidence intervals for the pre-surgical assessment alone;
- Provide sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) for the Score test performance in conjunction with the pre-surgical assessment using the decision rule “OR” along with 95% confidence intervals.
- Calculate the difference in NPVs and difference in PPVs along with 95% two-sided confidence intervals (the bootstrap technique can be used for calculation of the confidence intervals). Improvement in NPV should be statistically and clinically significant, and, if a loss in PPV is observed, you should justify the clinical acceptability of this loss.

In addition, you should present the observed frequencies of malignancy for different results of the pre-surgical assessment and the Score test results from the patients evaluated by non-GO in the table below along with 95% confidence intervals:

	<b>Frequency of Malignancy</b>	<b>95% CI</b>
Prevalence of malignancy among patients with adnexal mass assessed by non-GO physicians:		
Pre-surgical assessment alone “Positive”		
Pre-surgical assessment alone “Negative”		
Score test alone “Positive”		
Score test alone “Negative”		
Pre-surgical assessment “Positive” and Score “Positive”		
Pre-surgical assessment “Positive” and Score “Negative”		
Pre-surgical assessment “Negative” and Score “Positive”		
Pre-surgical assessment “Negative” and Score “Negative”		

The same information should be presented as likelihood ratios along with their 95% CI, tabulated as illustrated above for frequencies of malignancy. Likelihood ratio (Result) =  $\text{Pr}(\text{Result}|\text{Malignancy}) / \text{Pr}(\text{Result}|\text{No Malignancy})$ . Likelihood ratio, unlike predictive value, is independent of the prevalence of disease.

*Subgroup Analyses:* You should demonstrate statistically and clinically significant improvement in NPV of dual assessment vs. NPV of single assessment for pre-menopausal and post-menopausal patients separately analysis analogous to the one described for the overall population.

*Additional Information:* You should provide a tabulation of the descriptive statistics for your Score test within patients grouped according to tumor stage or histopathological findings.

***Results in Patient Populations Evaluated by Gynecologic Oncologists***

Ovarian adnexal mass assessment score test system is intended for women with pelvic masses who will be having surgery. The test is indicated as an aid in making referral decisions. Your clinical study should avoid possible bias in results from evaluating patients who may have been selectively enrolled at non-GO sites. (For example, some non GO

physicians may automatically refer patients to a GO for various reasons regardless of the pre-surgical evaluation. The enrollment in your clinical trial at such a site would lead to a potentially biased representation of patients.) You may opt to provide additional data in GO-evaluated patients to demonstrate a positive bias is not occurring in your test performance. This data is reviewed by FDA with the expectation that the performance is not diminished in the GO-evaluated group.

#### **E. Expected Values in Other Benign and Malignant Conditions**

The target population may have a wide variety of conditions unrelated to cancer but present at the time an ovarian mass has been identified. These other conditions (for which in some cases the actual measurements of the analytes [e.g., immunoassays that are components of the test] are indicated) could dramatically affect the Score test result and confound its interpretation. You should demonstrate the results of your test score in patients with the disease conditions indicated by the individual analyte assays, as well as benign and malignant conditions that may be occurring concurrently. Examples of these conditions are: endometriosis, pelvic inflammatory disease, diabetes, anemia, autoimmune diseases such as Crohn's, SLE and rheumatoid arthritis, cardiac disease, hepatitis, kidney diseases and malnutrition, and various cancers such as cervical cancer, lung cancer, breast cancer, and colorectal cancer.

#### **F. Reference Intervals**

Reference values in apparently healthy women may be provided, though such women are not part of the intended use population for the Score test. For each analyte and for the Score test result, any reference values should include women that span the age range of your test and should evaluate a minimum of 120 premenopausal women and 120 postmenopausal women unless you are able to demonstrate that there are not any differences between the two populations. You should include other ethnicities if possible (Latino and Asian) in addition to Caucasian and African American. You should provide the score for each woman and investigate the relationship of the score vs. age.

#### **G. Relevance of the Individual Analytes Included in the Score**

You should justify inclusion of each individual analyte in the use of the Score test. One option is to demonstrate that the individual analytes included in the calculation of the score are informative for ovarian malignancy using the data from the clinical study. For this, perform ROC analyses: for each individual analyte, present an ROC curve of the individual analyte and calculate the areas under ROC curve of the individual analyte along with confidence interval (multiplicity issue should be properly addressed). In addition, for each individual analyte, present an ROC curve of the individual analyte and the ROC curve of the score on the same graph. If the data of the clinical study did not demonstrate that some individual analytes are informative for ovarian malignancy, you should justify why these analytes were included in the calculation of the Score test.

## **IX. LABELING**

The premarket notification must include labeling in sufficient detail to satisfy the requirements of 21 CFR 807.87(e). Final labeling for in vitro diagnostic devices must also comply with the requirements of 21 CFR 809.10 before an in vitro diagnostic device is introduced into interstate commerce. The following list below is not inclusive of all the elements required in a labeling, but it is aimed at assisting you in preparing labeling that satisfies these requirements.

### **Intended use**

The intended use should specify what the test measures, the clinical indications for which the test is to be used and the specific population, as applicable, for which the test is intended. The intended use should specify whether the test is qualitative or quantitative.

### **Black Box Warning**

Considering the history and currently unmet medical needs for ovarian cancer testing, FDA concludes that there is a risk of off-label use of this device. To address this risk, manufacturers should provide notice concerning the risks of off-label uses in the labeling, advertising and promotional material of ovarian adnexal mass assessment score test systems. Manufacturers must address the following risks:

- Women without adnexal pelvic masses (i.e., for cancer "screening") are not part of the intended use population for the ovarian adnexal mass assessment score test systems. Public health risks associated with false positive results for ovarian cancer screening tests are well described in the medical literature and include morbidity or mortality associated with unneeded testing and surgery. The risk from false negative screening results also includes morbidity and mortality due to failure to detect and treat ovarian malignancy.
- Analogous risks, adjusted for prevalence and types of disease, arise if test results are used to determine the need for surgery in patients who are known to have ovarian adnexal masses.
- If used outside the "OR" rule that is described in this special control guidance, results from ovarian adnexal mass assessment score test systems pose a risk for morbidity and mortality due to non-referral for oncologic evaluation and treatment.

To address the risks of off-label use, labeling, advertising and promotional materials for ovarian adnexal mass assessment score test systems should contain a precaution box with text using the following template or equivalent:

**PRECAUTION:** The [test name] should not be used without an independent clinical/radiological evaluation and is **not** intended to be a screening test or to determine whether a patient should proceed to surgery. Incorrect use of the [test name] carries the risk of unnecessary testing, surgery, and/or delayed diagnosis.

### **Test Principle**

You should describe the test components (specific assays, calibrators, and instruments) or test methodology used in this type of device.

### **Warnings and Precautions**

You should include any warnings and precautions specific to your test, which include conditions that affect the sample, conditions specified in any other applicable manufacturer package insert for components of your test, and potential laboratory hazards.

### **Specimen and Reagent – Stability and Storage**

You should state the sample matrix used with your test, instructions for sample handling, and stability information (including storage and temperature). If your test system is comprised of individual immunoassays, specimen stability and storage claims should be limited to the performance claims of the most unstable component (assay) in your test system, unless you have provided validation data to demonstrate otherwise. You additionally should summarize the storage and stability date for each individual assay.

### **Test Components**

You should provide a list of the specific assays required for your test system including the calibrators and controls. You should provide the user a summary of any expectations for the performance of these assays that are relevant to your test performance, including but not limited to measuring ranges, measurement units, and quality control measures.

### **Procedure**

This section should include clear and concise instructions for the procedure, from specimen handling through to result reporting. Specific and sufficient instructions, including any troubleshooting recommendations for software installations, should be provided. Users can be referred to component package inserts; however, a general summary of the procedure for each assay should be included.

### **Interpretation of Results**

You should clearly define the possible range of results, the specific cut-points and/or equivocal zones used, the meaning of the results across these cut-points, and explain what the user should do in the event they have any equivocal results (e.g., repeat). You should indicate

whether the results should be interpreted differently based on age, menopausal status, or other factors.

### **Limitations**

You should clearly describe any and all limitations in the labeling. This section should include the appropriate limitations that an operator or physician needs to know prior to using the test.

In addition to any limitations and warnings that are relevant to your test, an ovarian adnexal mass assessment score test system should contain a statement that a negative test result, in the setting of a positive pre-surgical assessment, should not preclude oncology referral.

### **Clinical Performance Studies**

You should include in the package insert a summary of the demographic characteristics and pathology for all evaluable subjects in your study. You should include a summary of your study designs and the results from the studies. This section should include a description performance (sensitivity, specificity, NPV, PPV and 95% CI) for the pre-surgical assessments (single assessment), the test (single assessment) and the two combined (dual assessment). It should include, as applicable, results based on menopausal status, pathology, and stage. You should summarize your conclusions from these studies.

### **Analytical Performance Results**

You should provide summaries of the analytical performance results for your score test system. The results you provide should only be for the overall result, not the individual analytes. This data should include, when appropriate, precision (repeatability/reproducibility), range of numerical test score results, interference, cross-reactivity and matrix comparison.

### **Reference Values and Expected Values**

These sections should include the 5<sup>th</sup> and 95<sup>th</sup> percentile ranges of your test results in non-diseased women, and women with other benign and malignant conditions. The information should also include the number of samples, age, conditions, and demographics of the population used to determine the values.

## **X. References**

1. ACOG Practice Bulletin No. 83: Management of Adnexal Masses. *Obstetrics & Gynecology*. 110(1):201-214, July 2007.
2. Kondratovich M., Yousef WA. Evaluation of accuracy and optimal cutoff of diagnostic devices in the same study. *Joint Statistical Meeting*. 2005. ASA Section on Statistics in Epidemiology; p.2547-2551.

3. Gostout BS, Brewer MA. Guidelines for Referral of the Patient with an Adnexal Mass. 2006. *Clinical Obstetrics and Gynecology* 49(3): 448-459.