

Guidance for Industry and FDA Staff

Class II Special Controls Guidance Document: Gene Expression Profiling Test System for Breast Cancer Prognosis

Document issued on: May 9, 2007

For questions regarding this document contact Reena Philip at 240-276-1286 or by email at reena.philip@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

Written comments and suggestions may be submitted at any time for Agency consideration to Division of Dockets Management, Food and Drug Administration, 5630 Fishers Lane, Room 1061, (HFA-305), Rockville, MD, 20852. Alternatively, electronic comments may be submitted to <http://www.fda.gov/dockets/ecomments>. Please identify all comments with the docket number 2007D-0137. 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/cdrh/oivd/guidance/1627.pdf>. You may also send an e-mail request to dsmica@fda.hhs.gov to receive an electronic copy of the guidance or send a fax request to 240-276-3151 to receive a hard copy. Please use the document number 1627 to identify the guidance you are requesting.

Table of Contents

1. INTRODUCTION.....	4
2. BACKGROUND	5
3. SCOPE	6
4. RISKS TO HEALTH.....	7
5. DEVICE DESCRIPTION	8
INTENDED USE	8
TEST METHODOLOGY	8
TEST ALGORITHMS	9
TEST RESULTS	10
6. PERFORMANCE CHARACTERISTICS	10
PREANALYTICAL FACTORS	10
QUALITY CONTROL.....	11
ANALYTICAL PERFORMANCE.....	11
CLINICAL VALIDATION	13
7. SOFTWARE.....	16
8. LABELING	17

Guidance for Industry and FDA Staff

Class II Special Controls Guidance Document: Gene Expression Profiling Test System for Breast Cancer Prognosis

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.

1. Introduction

This guidance document was developed as a special controls guidance to support the classification of gene expression profiling test systems for breast cancer prognosis into class II (special controls). A gene expression profiling test system for breast cancer prognosis is a device that measures the RNA expression level of multiple genes and combines this information to yield a signature (pattern or classifier or index) to aid in prognosis of previously diagnosed breast cancer.

This guidance provides recommendations to manufacturers regarding preparation of premarket notifications and labeling for a gene expression profiling test system for breast cancer prognosis. The recommendations in this document are applicable to RNA expression assays used for cancer prognosis, such as reverse-transcriptase polymerase chain reaction (RT-PCR) and gene expression microarrays. In gene expression test systems for breast cancer prognosis, an algorithm is applied to such measurements to yield a result that can be used by physicians as a prognostic marker, in combination with clinicopathological factors, to assess the risk of cancer recurrence (e.g., distant metastasis).

This type of prognostic test is one for which test results explain the variation in outcomes for patients who are otherwise alike in terms of a predefined set of characteristics such as biological features (e.g., women over age 50 at a specific stage of disease) or a previously defined treatment (e.g., women receiving no adjuvant therapy).

A gene expression profiling test system for breast cancer prognosis is not intended for diagnosis, or to predict or detect response to therapy, or to select the optimal therapy for

Contains Nonbinding Recommendations

patients. This guidance does not address predictive markers, which are distinguished from prognostic markers, because predictive markers predict response to therapy¹.

This guidance is issued in conjunction with a *Federal Register* notice announcing the classification of gene expression profiling test systems for breast cancer prognosis. Any firm submitting a 510(k) premarket notification for a gene expression profiling test system for breast cancer prognosis will need to address the issues covered in this special controls guidance. However, the firm need only show that its device meets the recommendations of the guidance or in some other way provides equivalent assurances of safety and effectiveness.

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidance documents describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidance documents means that something is suggested or recommended, but not required.

The Least Burdensome Approach

The issues identified in this guidance document represent those that we believe need to be addressed before your device can be marketed. In developing the guidance, we carefully considered the relevant statutory criteria for Agency decision-making. We also considered the burden that may be incurred in your attempt to follow the statutory and regulatory criteria in the manner suggested by the guidance and in your attempt to address the issues we have identified. We believe that we have considered the least burdensome approach to resolving the issues presented in the guidance document. If, however, you believe that there is a less burdensome way to address the issues, you should follow the procedures outlined in the document, **“A Suggested Approach to Resolving Least Burdensome Issues.”** It is available on our Center web page at: <http://www.fda.gov/cdrh/modact/leastburdensome.html>.

2. Background

FDA believes that special controls, when combined with the general controls, will be sufficient to provide reasonable assurance of the safety and effectiveness of a gene expression profiling test system for breast cancer prognosis. A manufacturer who intends to market a device of this generic type should: (1) conform to the general controls of the Federal, Food, Drug, and Cosmetic Act (the Act), including the premarket notification requirements described in 21 CFR 807 Subpart E, (2) address the specific risks to health associated with the device identified in this guidance, and (3) obtain a substantial equivalence determination from FDA before marketing the device.

¹ Sargent DJ, Conley BA, Allegra C, Collette L. Clinical trial designs for predictive marker validation in cancer treatment trials. *J Clin Oncol.* 2005;23(9):2020 – 2027.

Contains Nonbinding Recommendations

This guidance document identifies the classification regulation and product code for gene expression profiling test system for breast cancer prognosis (refer to Section 3 – Scope). In addition, other sections of this guidance document identify the risks to health and describe measures that, if followed by manufacturers and combined with the general controls, will generally address the risks associated with gene expression profiling test systems for breast cancer prognosis and lead to a timely premarket notification (510(k)) review and clearance. This document supplements other FDA documents regarding the specific content requirements of a premarket notification submission. You should also refer to 21 CFR 807.87 and other FDA documents on this topic, such as **Premarket Notification: 510(k)**, available on the web at: <http://www.fda.gov/cdrh/devadvice/314.html>.

There are three types of Premarket Notification 510(k)s that may be submitted to FDA: Traditional, Special, and Abbreviated. The Special and Abbreviated 510(k) methods were developed to help streamline the 510(k) review process and are explained in “**The New 510(k) Paradigm – Alternate Approaches to Demonstrating Substantial Equivalence in Premarket Notifications; Final Guidance**” (<http://www.fda.gov/cdrh/ode/parad510.html>). An Abbreviated 510(k) provides a means to simplify the review of data in a 510(k) through a reliance on FDA-recognized consensus standards, special controls, or FDA guidance documents, and provides the least burdensome means of demonstrating substantial equivalence for a new device. Guidance on the content and format for abbreviated and traditional 510(k)s is available at <http://www.fda.gov/cdrh/ode/guidance/1567.html>. Also, see Section 514(c)(1)(B) of the Act and the FDA guidance, “**Use of Standards in Substantial Equivalence Determinations**” (<http://www.fda.gov/cdrh/ode/guidance/1131.pdf>) for additional information. The Special 510(k) is available for manufacturers considering modifications to their own cleared devices. Information on how to prepare a Special 510(k) is available at <http://www.fda.gov/cdrh/devadvice/3144.html>.

3. Scope

The scope of this document is limited to the following devices described in 21 CFR 866.6040 (product code NYI).

21 CFR 866.6040– A gene expression profiling test system for breast cancer prognosis is a device that measures the RNA expression level of multiple genes and combines this information to yield a signature (pattern or classifier or index) to aid in prognosis of previously diagnosed breast cancer.

Traditionally, prognosis is a term reserved for patients who are untreated (in this context, those that do not receive any adjuvant therapy). However, providing information on predicted outcomes for women within a single therapy regime (e.g., estrogen receptor (ER) positive women treated solely with tamoxifen) can also have clinical utility in terms of breast cancer prognosis and falls within the scope of this guidance.

A gene expression profiling test system for breast cancer prognosis is a test that may require instrumentation for clinical multiplex test systems. Instrumentation for clinical multiplex test systems is regulated under 21 CFR 862.2570. Guidance for such

Contains Nonbinding Recommendations

instrumentation is available in the FDA Guidance for Industry and FDA Staff, “**Class II Special Controls Guidance Document: Instrumentation for Clinical Multiplex Test Systems.**”² If your gene expression profiling test system for breast cancer prognosis includes instrumentation for clinical multiplex test systems for that assay, you may submit the information for both the assay and the instrumentation within one 510(k). If instrument manufacturers prefer to submit a 510(k) for instrumentation only, they may submit it in conjunction with the assay premarket notification.

4. Risks to Health

A gene expression profiling test system for breast cancer prognosis is intended to provide prognostic information to aid in clinical evaluation of breast cancer patients. Failure of this device to perform as indicated may lead to erroneous test results. False positive results will misclassify the patient into a higher risk group and false negative results will misclassify the patient into a lower risk group. Misclassification of cancer recurrence risk may lead to incorrect prognosis with attendant psychological distress, inaccurate counseling, and suboptimal patient care.

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 type of expression assays used, the intended use of the test, the sample type, and how the result will be 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
Failure of the test to perform properly, for example, inaccurate or absent results due to failure of reagents, instrumentation, data management, or software, may lead to false positive results, or false negative results, and an incorrect prognosis.	Section 6-7
Failure to properly interpret test results	Sections 5 (see Test Results section) and 8

² <http://www.fda.gov/cdrh/oivd/guidance/1546.pdf>

5. Device Description

In your 510(k) submission, you should identify the regulation, the product code, and a legally marketed predicate device. In order to help FDA efficiently review all the aspects of your device compared with the predicate, you should include a table that outlines the similarities and differences between the predicate and your device.

Key issues in the review of a new device are the specific intended use, the type of specimens tested, and the technology utilized. You may submit appropriate peer-reviewed literature references relevant to the technology of the device, in addition to the descriptive information, to adequately describe the new device.

You should include the following descriptive information to adequately characterize your gene expression profiling test system for breast cancer prognosis.

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 for which the test is intended. It should include clinical and demographic description of patients (e.g., gender, age, lymph node status, stage, tumor type, tumor size) for whom clinical performance has been demonstrated. The intended use should specify whether the test is qualitative or quantitative. If the test is intended for use at a single laboratory site, this information should be included in the intended use.

Test Methodology

You should describe in detail the methodology used by your device. For example, you should describe the following elements, as applicable to your device:

- Test platform (for example, RT-PCR or expression arrays).
- Composition and spatial layout of arrays or other spatially fixed platforms.
- Description of the assay elements, particularly with respect to parameters such as the genes used for normalization, indicators of hybridization, and quality control.
- How you evaluated the potential for sample carryover or contamination.
- Limiting factors of the assay (e.g., saturation level of hybridization, maximum cycle number).
- For arrays:
 - Methods used in attaching the probe material to a solid surface.
 - Hybridization conditions, washing procedures and drying conditions (e.g., temperature, length of time).
- Specificity of probes for the sequence of interest, especially when pseudogenes or sequence-related genes exist.
- Sample collection and handling methods from the time the tumor or alternative specimen is extracted until the processing of the sample.

Contains Nonbinding Recommendations

- Methods for RNA extraction that you perform, provide, or recommend to users.
- Methods for ensuring RNA integrity in sample extracts.
- Reagent components provided or recommended for use, and their function within the system (e.g., buffers, enzymes, fluorescent dyes, chemiluminescent reagents, other signaling/amplification reagents).
- Instrumentation required for your device, including the components and their function within the system.
- Types of output generated by the instrumentation and system parameters (e.g., measurement ranges).
- The computational path from raw data to the final prognostic result (e.g., how raw signals are converted into a prognostic signal). This would include sufficient software controls for identifying and dealing with missing values and obvious problems in the dataset. Describe adjustment for background for normalization.
- External controls that you recommend or provide to users.
- Internal controls and a description of their specific function in the system.
- Related peer-reviewed literature references describing the test methodology, if applicable.
- Illustrations or photographs of non-standard equipment or methods, if available.

Where applicable for your device, you should describe the quality control design specifications used to address the following concerns:

- Correct placement and identity of assay features (e.g., probes).
- For multiplexed tests in which the target molecules will contact a number of different probes, the potential for specific and non-specific probe cross-hybridization.
- Prevention of probe cross-contamination, for multiplexed tests in which many probes are handled during the manufacturing process.

Test Algorithms

The algorithms that are used in these types of test systems to predict breast cancer prognosis may often be novel, proprietary, and complex, and are among the most critical elements of the test system. You should provide the following, when applicable:

- A detailed description of the algorithm architecture and implementation.
- A detailed description of the datasets that were used to discover and validate the patterns or classifiers that are used in your test (often referred to as “training” and independent “test” sets), including the principles used to select the samples from which the data were derived (clinical history, demography, matrix, geographical origin, etc.), the statistical justification for sample size, and any assumptions you made when assembling the datasets.

Contains Nonbinding Recommendations

- A detailed description of performance measures (internal validation and external validation using an independent clinical dataset) and how they were obtained.

In some cases the device and the algorithm evolve over time during product development. You should provide the data obtained using the final device and final algorithm for the device described in the submission.

Test Results

You should provide examples of the test reports (e.g., printouts) that are generated for the clinician. These reports should contain adequate information to allow interpretation by the ordering physician or other healthcare professional. The test report should reference the performance of the test in the clinical validation dataset (e.g., “The analysis of this test in a clinical population revealed low risk patients have a probability of 92% of metastasis free survival at 5 years. High risk patients have a probability of 60% metastasis free survival at 5 years”). The report may contain other descriptive information such as Kaplan-Meier survival curves for low risk and high risk patients as calculated using the clinical validation data set.

6. Performance Characteristics

In your 510(k), you should detail the study design you used to evaluate each of the performance characteristics outlined below.

Preanalytical Factors

Consideration of preanalytical factors is critical for high-quality genomic tests.

Specimen collection

You should evaluate all sample collection, transport, and storage options you recommend (e.g., RNA preserving fixatives, frozen, fixed paraffin-embedded tumor tissue). You should ensure that the test is validated using specimens that are handled in the same manner as will be recommended in the test label (e.g., collection, storage, shipment methods). You should validate that the allowable elapsed time between tumor resection and preservation (e.g., by snap freezing, fixation or other methods) results in uniformly acceptable specimens. You should specify the specimen transport conditions. You should validate that the transport conditions are adequate to ensure sample integrity, and to determine the limits of transport variability that are acceptable (e.g., time in transit, quantity of coolant required).

Your validation of appropriate storage conditions should include both the sample and the extracted RNA product.

RNA extraction

Contains Nonbinding Recommendations

If you intend to provide reagents in your test kit for extraction and preparation of RNA, you should validate each step in the preanalytical process for its effects on reproducibility, accuracy, and stability of product, and describe your study design and results in the 510(k) submission. Your external site studies (e.g., reproducibility, method comparison) should include evaluation of preanalytical processes.

If you do not intend to provide reagents in your test kit for RNA extraction and preparation, you should provide adequate specifications to ensure RNA quality sufficient to generate correct test results. Examples for specifications include OD₂₆₀/OD₂₈₀ ratio, ribosomal RNA ratio (28S/18S), and measurement of RNA integrity. You should not recommend any research-use-only (RUO) reagents.

Quality Control

Several levels of quality controls should be considered for gene expression profiling test system devices of this type. Controls should provide information about 1) sample/biopsy quality, 2) RNA quality, and 3) process quality. The process quality controls should reflect the whole process, including but not limited to, RNA labeling, amplification, hybridization, scanning, and normalization.

Controls should approximate the composition and RNA concentration of a sample in order to adequately challenge the system, as well as address reproducibility around the cut-off.

You should describe the following concerning quality control and calibration:

- The nature and function of the various controls that you include with, or recommend for, your system. These controls should enable the user to determine if all steps and critical reactions have proceeded properly without contamination or cross-hybridization.
- Your methods for value assignment (relative or absolute) and validation of control and calibrator material, if applicable.
- The control parameters that could be used to detect failure of the instrumentation to meet required specifications.

Analytical performance

All analytical performance studies should be conducted using the final version of your device rather than a prototype. You should evaluate performance of your assay, including RNA extraction, from all the sources of RNA that you recommend for your assay (e.g., tissue biopsy, needle biopsy). We recommend that you describe the following performance characteristics:

Specimen requirements

You should validate that the specimen requirements you specify are sufficient to identify the diagnostic patterns or classifiers of your test within your stated accuracy and precision criteria. You should determine:

Contains Nonbinding Recommendations

- The minimum amount of tissue required to perform an acceptable assay with your device.
- The minimum percentage of tumor cells in the specimen (e.g., as determined by haematoxylin and eosin (H&E) stain) required to generate an acceptable result.
- The maximum acceptable percentage of necrotic or hemorrhagic tissue, if applicable.
- The lower and upper limit of the assay, in terms of RNA/cRNA concentration and amount of tumor specimen, for which the device can give reliable results with a given accuracy and precision.

For assays using a complex algorithm to generate a signature (pattern or classifier or index), the upper and lower limits of RNA concentration and/or percentage of tumor cells should not compromise the assay outcome as indicated by precision measures.

Analytical Specificity/Interference

Where applicable, you should evaluate potential for non-specific amplification, non-specific hybridization, and cross-hybridization of your device.

Potential interfering substances may exist in the specimen (e.g., adipose tissues, blood) and may be introduced during specimen collection (e.g., environmental effects such as crush artifact) and sample preparation. Therefore, your RNA specifications should be adequate to exclude the presence of any effect from likely interfering substances.

Cut-off

In your submission, you should explain how the cut-off was determined and how this cut-off value was validated. The cut-off should be established using statistical methods that are appropriate for your classifier development strategy. If the assay has an equivocal zone, you should explain how you determined the limits of the equivocal zone. The performance of your device using the established cut-off (and equivocal zone, if applicable) should be validated in an independent population consistent with the defined intended use of your device.

Precision (Repeatability/Reproducibility)

You should provide data demonstrating the precision (i.e., repeatability/reproducibility) of your system. The CLSI documents, "Evaluation of Precision Performance of Clinical Chemistry Devices" (CLSI Guideline EP5-A) and "User Protocol for Evaluation of Qualitative Test Performance" (CLSI Guideline EP-12A), include guidelines that may be helpful for developing experimental design, computations, and a format for establishing performance claims. Ideally, you should identify all sources of assay variability in the precision study. You should establish the performance characteristics for each classifier across the entire range (e.g., high risk, low risk, borderline) of each classifier that can be reported. Additional factors influencing precision that you should consider include the following:

Contains Nonbinding Recommendations

- Ensure that samples used in reproducibility testing are processed from clinical specimens (e.g., tissue biopsy) at the test site, using the procedure you plan to recommend in the test labeling.
- If the assay is intended to be performed in more than one laboratory, include three or more sites with multiple operators at each site. Operators should reflect potential users of the assay in terms of education and experience. You should provide training only to the same extent that you intend to train users after marketing the test.
- If the assay is intended to be performed in a single laboratory, include multiple operators from that laboratory.
- Include multiple product lots (e.g., multiple lots of reagents, multiple lots of primers and probes for RT-PCR, multiple lots of arrays), and multiple instruments.
- Use appropriate test samples representing every class that the test can detect (e.g., high risk, low risk, borderline).
- If applicable, perform dye-reverse experiments to ensure that there is no bias in dye incorporation.
- If applicable, demonstrate reproducibility of sample labeling procedures.

In the study design description in your 510(k), you should identify which factors (e.g., instrument calibration, reagent lots, and operators) were held constant and which were varied during the evaluation, and describe the computations and statistical analyses used to evaluate the data.

Stability studies

You should describe your study design for determining the real-time stability of the reagents and instruments, and if applicable, for accelerated stability and stress test conditions and results. For each study, you should describe how you selected the acceptance criteria values.

Validation of instrumentation

For instruments and systems that measure and sort multiple signals, and other complex laboratory instrumentation that has not been previously cleared, refer to the guidance document: "**Class II Special Controls Guidance Document: Instrumentation for Clinical Multiplex Test Systems**,"³ for details on the types of data you should provide to support instrument clearance.

Clinical Validation

You should provide data from clinical studies to support the indications for use and claims for your device. The clinical validation study should use patient samples that are derived from the intended use population and that are independent of the specimens you used to develop the signature (pattern or classifier or index). You should describe the protocol of each clinical study (including the inclusion and exclusion criteria, study endpoints, acceptance criteria), and a description of how the studies support the proposed

³ <http://www.fda.gov/cdrh/oivd/guidance/1546.pdf>

Contains Nonbinding Recommendations

intended use. You should submit the raw data along with the processed data (i.e., prognostic results) from your clinical validation studies.

For the clinical validation study, the validation dataset should consist of clinical samples collected from at least three different clinical sites in different geographical locations. Preferably, studies would be conducted within the U.S. population. If the studies are conducted outside the U.S, you will need to document the relevance of your studies to U.S. clinical practice and demographics.

If the clinical validity and utility of your specific device is supported by an established scientific framework and a sufficient body of evidence, then you may submit peer-reviewed references to support your claim. These should include multiple studies that test appropriate populations. In cases where the literature does not sufficiently support your indications for use, you should conduct studies to support claims for your device. Retrospective analysis of prospectively collected banked samples may be acceptable if appropriate measures are taken to identify and either remove or mitigate any biases in the study set. We recommend that you discuss with FDA your specific proposed study to determine whether it is adequate.

Accuracy using comparison to clinical outcome:

Clinical truth: In order to allow FDA to judge the performance of your device, you should define the measure of clinical outcome used for all patients in the clinical validation study, and the method by which the measure was obtained.

End points: You should describe the appropriate prognostic endpoints for your device. Examples include 1) time from surgery to distant metastases, 2) overall survival (defined as the time from surgery to death from any cause), and 3) disease-free survival (defined as time from surgery to any recurrence - local or regional, second breast primary, distant metastasis, or death from any cause). For example, a Kaplan-Meier, product-limit estimator can be used to display time-to-event curves for one or more of these three endpoints. Ninety-five percent, two-sided confidence intervals for fixed time intervals may also be included, but the actual times may differ with the intended use population (e.g., events at 5 years may be relevant for some patient groups but less relevant for others). Alternatively, continuous-valued risk descriptors (e.g., hazard ratios) may be used if model assumptions are met.

Validation Strategy: You should provide the method used to validate the gene signature. This should include a clinical protocol and statistical analysis plan. The clinical data should be a new data set not used in the development of the gene signature and the patients should be representative of the intended use population for the device. For the statistical approach, one can consider the estimation of "hazard ratios" (an estimate calculated using statistical methods for time to event data) to quantify the relative risk of an event in the high-risk group compared with the low-risk group. The statistical analysis plan for validation should include a hypothesis about the relative risk that is of interest in the clinical study, e.g., the risk of developing metastatic cancer within 5 years can be estimated by the gene expression profile x . The hypothesized relative risk should be a clinically relevant difference that validates the gene signature as a prognostic marker.

Contains Nonbinding Recommendations

The clinical study should be sized to obtain sufficient statistical power to demonstrate this hypothesis. Note that in a longitudinal study some patients will be censored, e.g., if a woman dies of unrelated causes, such as heart disease before the end of the study; however, we would expect all such cases to be included in the analysis. Many statistical methods rely on assumptions that you should check prior to submission of your 510(k) (e.g., proportional hazards in a Cox regression model). You should provide summaries of this clinical validation study, including descriptive statistics for patients within the study as well as either survival curves for specific groups of patients or estimates of risk associated with your endpoint (e.g., estimated proportion of patients that develop metastatic disease within 5 years.⁴)

Prognostic performance can be measured in terms of the probability or risk of metastatic disease as follows:

- (1) P(no metastatic disease within 5 years given the device outcome is “low risk for metastatic disease”) and,
- (2) P(metastatic disease within 5 years given the device outcome is “high risk for metastatic disease”).

Note that (1) is consistent with a definition of Negative Predictive Value and (2) is consistent with Positive Predictive Value. We ask that you report a 95% confidence interval for each. The performance will be impacted by the prevalence of “metastatic disease within 5 years” in the pivotal clinical study. Therefore, you should report prevalence of the target endpoint in the cohort studied.

In addition to a primary analysis using the results of your device, you should provide an analysis that demonstrates your device is “value added” and provides additional information concerning prognosis even after considering clinical data available to a physician. In breast cancer, there is information available from a variety of sources that provides prognostic value. (For example, the age of the patient, ER status, tumor size and grade, are routinely assessed.). You should provide information that demonstrates added prognostic value in comparison with routine information obtained in current clinical practice. A Cox regression model may be considered.

The clinical information appropriate for consideration may vary with the study group of interest. We recommend that you discuss with FDA your specific proposed study prior to conducting your study.

Study Samples

While prospective samples are preferred, well-characterized samples from banks can be used in your clinical validation study, provided that there is no collection or selection bias, and patient history and appropriate outcome information are available.⁵ You should

⁴ Five years is used in this section, as an example of a minimum time point. It is possible that some studies may have endpoints exceeding five years.

⁵ The use of banked leftover specimens is discussed in FDA’s guidance “Guidance on Informed Consent for *In Vitro* Diagnostic Device Studies Using Leftover Human Specimens that are Not Individually Identifiable.” <http://www.fda.gov/cdrh/oivd/guidance/1588.html>.

Contains Nonbinding Recommendations

fully describe selection (inclusion/exclusion) criteria and characterize any relevant features or limitations of the samples (whether prospective or from banks). You should describe patient demographics and disease characteristics and the prevalence of relevant outcomes in the intended use and study populations. You should select samples in a way that minimizes the sources of bias such as sample integrity, storage duration, and tumor size. We recommend you consult with FDA prior to performing pivotal studies using banked samples.

You should use clinical samples from all matrices you claim in your intended use (e.g., frozen, or formalin-fixed, paraffin-embedded (FFPE), or collected in any nucleic acid preservative) to demonstrate that correct results can be obtained from clinical material. Appropriate sample size depends on factors such as precision/reproducibility, interference, and other performance characteristics of the test. We recommend that you provide a justification using statistical methods to support your study sample size. For samples you use in your clinical studies, you should provide data demonstrating that storage and transport of retrospectively examined samples have not affected assay results.

Sample collection and handling conditions

You should assess the effect of recommended storage times and temperatures on sample stability and recovery using an analysis of specimen aliquots stored/transported under the recommended conditions of time and temperature, and which have undergone a specified number of freeze/thaw cycles (if appropriate). For these types of studies, you should state your acceptance criteria for all sample stability parameters.

7. Software

If your system includes software, you should submit software documentation detailed in accordance with the level of concern (See: **“Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices.”**⁶). 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 indirectly affect the patient and potentially result in injury because the healthcare provider and patient do not get accurate information.

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://aspe.os.dhhs.gov/admnsimp>.

⁶ <http://www.fda.gov/cdrh/ode/guidance/337.pdf>

Contains Nonbinding Recommendations

- 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 assay 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) 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/cdrh/ode/510kmod.pdf>.
- Guidance for Off-the-Shelf Software Use in Medical Devices; Final; available on the FDA Web site at: <http://www.fda.gov/cdrh/ode/guidance/585.pdf>.
- 21 CFR 820.30 Subpart C – Design Controls of the Quality System Regulation.
- ISO 14971-1; Medical devices - Risk management - Part 1: Application of risk analysis.
- AAMI SW68:2001; Medical device software - Software life cycle processes.

8. Labeling

The premarket notification should include labeling in sufficient detail to satisfy the requirements of 21 CFR 807.87(e). Although final labeling is not required for 510(k) clearance, final labeling for in vitro diagnostic devices must comply with the requirements of 21 CFR 809.10 before an in vitro diagnostic device is introduced into interstate commerce. The recommendations below are aimed at assisting you in preparing labeling that satisfies these requirements.

For a test intended to be performed at one laboratory site that does not distribute a package insert as part of a packaged device, the manufacturer should make the labeling information available to users by providing a reference link to the 510(k) summary and/or decision summary documents posted at the publicly accessible FDA 510(k) database at the <http://www.accessdata.fda.gov> website in their test report form.

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 for which the test is intended. It should include a description of patients, e.g., gender, age, lymph node status, stage, tumor type,

Contains Nonbinding Recommendations

tumor size for whom clinical performance has been demonstrated. The intended use should specify whether the test is qualitative or quantitative. If the test is intended for use at a single laboratory site, this information should be included in the intended use.

Device Description

You should describe the test methodology used in your device.

General Procedure

You should include a general description of the analysis procedure, from physician sampling up to, and including, result reporting.

Directions for use

You should present clear and concise instructions that delineate the technological features of the specific device and how the device is to be used. Instructions should encourage users to familiarize themselves with the features of the device and how to use it in a safe and effective manner.

You should include handling and storage instructions. You should describe stability (i.e., expiration dating) under the opened and closed storage conditions that you recommend to users.

Quality Control

You should provide quality control recommendations in the package insert. This should include a clear explanation of what controls are to be used in the assay and the expected results for the control material.

Precautions, Warnings, and Limitations

You should clearly describe any assay limitations in the labeling. This section should include the appropriate limitations and warnings that a physician needs to know prior to ordering the test.

In addition to any limitations and warnings that are relevant to your assay, a gene expression profiling test system for breast cancer prognosis should contain the following limitations:

- Results from this assay should not be used for diagnosis.
- Results from this assay should not be used to predict response to therapy regimens or to select the optimal therapeutic regimen.
- Results from this assay should not be used to exclude a therapeutic regimen.
- A statement explaining that the results are limited to the pool of patient samples that were used in the study, e.g., a statement that the study used only banked samples from women who did not receive adjuvant therapy, or that the women in the study represented only certain populations.

Contains Nonbinding Recommendations

Performance Characteristics

You should include in the package insert a summary of the study designs and the results from the studies described in Section 6 that would aid users in interpreting test results. This section should include a description of the clinical (i.e., medical) and analytical (i.e., technical) performance characteristics. Clinical performance characteristics should include clinical study validation summaries. Analytical performance characteristics should include descriptions of the results and methodologies used for the studies.

Interpretation of Results

You should clearly define the “classification,” “pattern,” “score,” or “index” used to convey the patient-specific result. Prognostic endpoints cited in the report (such as time to distant-metastasis or overall survival and disease-free survival) should be based on the results from the clinical trial that was used to clinically validate the device.

Expected Values

This section should include the expected values of the test and the explanation of the result (e.g., High Risk means that x% of the patients in a reference group have developed distant metastasis within 5 years, Recurrence Score 7 means that ...). It should also include the number of samples, age, gender, and demographics of the population used to determine the expected values.