Guidance for Industry and FDA Staff

Pharmacogenetic Tests and Genetic Tests for Heritable Markers

Document issued on: June 19, 2007

The draft of this guidance was issued on February 9, 2006



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health
Center for Drug Evaluation and Research
Center for Biologics Evaluation and Research

GRZAN

Preface

Comments

Written comments and suggestions may be submitted at any time for Agency consideration to the Division of Dockets Management, Food and Drug Administration, 5630 Fishers Lane, Room 1061, (HFA-305), Rockville, MD, 20852. When submitting comments, please refer to the docket number 2006D-0012. Comments may not be acted upon by the Agency until the document is next revised or updated.

For questions regarding this document contact Robert Becker at 240-276-0843, email robert.becker1@fda.hhs.gov, or Maria Chan at 240-276-0848, email maria.chan@fda.hhs.gov.

For questions regarding proposed use of this document in relation to applications to CBER, contact CBER's Office of Communications, Training, and Manufacturers Assistance at 800-835-4709 or 301-827-1800. For questions relating to applications to CDER, contact Felix Frueh, Office of Clinical Pharmacology and Biopharmaceutics, HFD-850 10903 New Hampshire Avenue, Silver Spring, MD, 20993, or at 301-796-1530.

Additional Copies

Additional copies are available from the Internet at:
http://www.fda.gov/cdrh/oivd/guidance/1549.pdf or http://www.fda.gov/cber/guidelines.htm
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 1549 to identify the guidance you are requesting.

Table of Contents

I. Introduction 1

II. Pharmacogenetic Testing versus Genetic Testing 3

III. Recommendations for the Preparation of the Pharmacogenetic or Genetic Test Applications 3

A. Intended Use of a Device 4

B. Device Design 5

C. Analytical Studies 6

D. Software and Instrumentation 9

E. Comparison studies using clinical specimens 11

F. Clinical Evaluation Studies Comparing Device Performance to Accepted Diagnostic Procedure(s) 12

G. Effectiveness of the Device 13

IV. Labeling 14

Appendix I: General considerations for planning and evaluating clinical studies 17

Guidance for Industry and FDA Staff

Pharmacogenetic Tests and Genetic Tests for Heritable Markers

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 guidance document is intended to facilitate progress in the field of pharmacogenomics and genetics by helping to shorten development and review timelines, facilitate rapid transfer of new technology from the research bench to the clinical diagnostic laboratory, and encourage informed use of pharmacogenomic and genetic diagnostic devices. It provides recommendations to sponsors and FDA reviewers in preparing and reviewing premarket approval applications (PMA) and premarket notification (510(k)) submissions[1] for pharmacogenetic and other human genetic tests, whether testing is for single markers or for multiple markers simultaneously (multiplex tests). Array-based tests (commonly referred to as microarrays) are a subset of multiplex tests and are included in the scope of this document. The recommendations within this guidance for elements of a genetic test submission apply to pharmacogenetic (e.g., drug-metabolizing enzyme allele tests, single nucleotide polymorphism (SNP) analysis) and other types of genetic tests. Tests of gene expression and tests for non-heritable (somatic) mutations are not specifically addressed, although many of the same principles may apply. In addition, this guidance considers nucleic acid-based analysis only, but the principles may be applied to other matrices (e.g., protein) when the purpose is to provide genetic information.

This document is intended to recommend a basic framework for the types of data and regulatory issues that we believe should be addressed in a genetic test submission and

[1] For information about PMA and 510(k) submissions, refer to the following website: http://www.fda.gov/cdrh/devadvice

provide a common baseline from which both manufacturers and scientific reviewers can operate. The recommendations contained in this document are purposefully general. It is well-known that each testing system will have an associated unique set of concerns, and we expect to identify and discuss these unique concerns with individual manufacturers, for example, through pre-IDE submission meetings.

Although this document focuses on information to include in a 510(k), the general types of information are likely to be the same for PMAs. However, we may request different types of data and statistical analyses in PMAs. The appropriate information depends on the following:

- intended use (e.g., to detect cytochrome P450 enzyme alleles)
- indications for use (e.g., predictive or prognostic for disease, treatment response, or drug sensitivity)
- methodology (e.g., polymerase chain reaction)
- technical interpretation of results (e.g., positive for variant alleles)
- quality control and assay limitations
- performance (see sections C-E below)
- clinical validity (e.g., false positives and negatives, see sections F-G below)
- clinical interpretation
- benefits and risks
- claims made by the manufacturer (e.g., effectiveness)

Technical aspects of this guidance may also be useful for other FDA applications that utilize these assay formats in support of product development, e.g., Investigational New Drug Applications (INDs), Biologics License Applications (BLAs), and New Drug Applications (NDAs). We recommend that the sponsor or manufacturer consult with the appropriate review Office within the Center for Biologics Evaluation and Research (CBER) or Center for Drug Evaluation and Research (CDER) for these types of applications.

The appropriate type of submission depends on claims and information available regarding the specific device. We expect that most pharmacogenetic and genetic device submissions will be traditional 510(k)s or de novo classifications. However, some devices will require submission of a PMA (see sections 513 and 515 of the Federal Food, Drug, and Cosmetic Act (Act) (21 U.S.C. 360c, 360e)). We recommend that the sponsor or manufacturer consult with the Office of In Vitro Diagnostic Device Evaluation and Safety (OIVD), to determine the appropriate type of submission. We also suggest that sponsors consider submitting protocols ("pre-IDEs") before carrying out studies to ensure review issues are addressed and resolved prior to submission of a 510(k) or PMA. Additional information on submission of pharmacogenomic information can be found in the "Guidance for Industry: Pharmacogenomic Data Submissions" (http://www.fda.gov/cder/guidance/6400fnl.pdf). FDA has also issued special controls guidance documents for some specific types of genetics tests, including drug metabolizing enzyme genotyping systems,

(http://www.fda.gov/cdrh/oivd/guidance/1551.pdf) and cystic fibrosis transmembrane regulator gene mutation detection systems (http://www.fda.gov/cdrh/oivd/guidance/1564.pdf).

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances 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 guidances 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.

II. Pharmacogenetic Testing versus Genetic Testing

Fundamentally, testing for pharmacogenetic polymorphisms and genetic mutations is the same and yields the same general types of results. The target populations and how the test results are used, however, are expected to be quite different. We consider pharmacogenetic tests for clinical use to be mostly those that are intended to provide information that may aid in selection of certain therapeutics. When sufficient clinical information is available, they may also aid in dosage selection of the therapeutic. Therefore, a pharmacogenetic test target population will typically be composed of candidates for a particular therapeutic. Target populations of genetic tests, on the other hand, will usually be composed of those who are suspected of having, or are at risk of developing, a particular disease or condition. The following recommendations will apply to both types of tests unless noted otherwise.

III. Recommendations for the Preparation of the Pharmacogenetic or Genetic Test Applications

The following are areas that you should address in the preparation of a submission for a medical device that measures pharmacogenetic or genetic information.

A. Intended Use of a Device

An application for premarket approval or clearance of a device must include a statement of the intended use of the device. 21 CFR 807.92(a)(5), 814.20(b)(3)(i). The intended use of the device for which approval or clearance is sought should specify the marker the device is intended to measure, the clinical purpose of measuring the marker, and the populations to which the device is targeted, where appropriate.

Some devices may have multiple intended uses. We encourage separate applications for each intended use, if each has unique and separate supporting studies; however, in certain cases of pharmacogenetic tests, we would consider application of test results in multiple therapeutic settings as a single intended use. For example, determination of CYP2D6 alleles for the purpose of providing information to aid in drug selection, without reference to a particular drug, would be an appropriate single intended use, given that it is well known that CYP2D6 affects the metabolism of many drugs. In other cases, it may be necessary to identify multiple intended uses. For example, a genetic test for a disease-causing mutation could be used for testing for carriers, prenatal testing, or for diagnosis. Each of these scenarios would have studies using different populations. In addition, the different uses might have different risk profiles and, therefore, might have separate intended use claims and submissions. In these cases, you should provide appropriate data to support each claimed intended use. You should consult the appropriate review divisions in OIVD for advice on submitting tests with multiple intended uses.

In this document, "screening" as an intended use is considered to be an indication to test an asymptomatic individual who is not necessarily at increased risk due to a positive family history. We recommend that if you are presenting data to support this type of intended use, you carefully consider the issues listed below.

- Some alleles, genotypes, and mutations will have very low prevalence in given populations. In these cases, samples from many patients should be obtained in order to detect a significant number of positives. Furthermore, some alleles, genotypes, or mutations might only be present in particular ethnic groups, which should therefore participate in the study in significant numbers. Enrichment can sometimes be appropriate to address these types of problems. However, one of the drawbacks of enrichment is that sensitivity can be affected by spectrum bias due to irregular retrospective selection of cases. Also, predictive values are dependent on the prevalence in the intended use population, which cannot be characterized from a study in which enrichment is used. We recommend that you contact OIVD for feedback if you are considering using sample enrichment in your studies.
- When many samples are tested for rare events, false positive results could become problematic in that they may be more common than true positives, due to test error and low prevalence.

- In some cases, properly banked samples may be studied to establish a predictive screening indication in healthy or asymptomatic individuals. In other cases, a study including long-term follow-up may be the only way to prove that the test was indeed predictive and to evaluate issues such as penetrance. In select cases, it may be possible to use postmarket studies to support this type of indication.

We recommend that you consult OIVD about study design if your device is intended for screening or detection of rare mutations or variants.

B. Device Design

An application for premarket approval or clearance of a device must include information on the design of the device. 21 CFR 807.92(a)(4), 814.20(b)(3)(ii), 814.20(b)(4). We recommend that you carefully characterize design of pharmacogenetic and genetic testing devices. For example, you should describe the following elements where applicable:

- Test platform (e.g., flow cytometry, instrumentation for clinical multiplex test systems).
- Composition and layout in spatially fixed platforms, including feature (e.g., probe) identity and placement, where applicable.
- Methods used in attaching the probe material to a solid surface, if applicable.
- Sequence or identity of oligonucleotides, primers, probes, or other capture elements.
- Hybridization conditions, washing procedures, and drying conditions (e.g., temperature, length of time).
- Assay components such as buffers, enzymes, fluorescent dyes, chemiluminescent reagents, other signaling and signal amplification reagents, instruments, software, etc.
- Specificity of probes for locus of interest; this is especially important when pseudogenes or sequence-related genes exist.
- Methodology for DNA extraction that you provide or that you recommend for users, and other applicable preanalytical elements.
- Range of input sample concentrations that meet performance specifications.
- Internal controls and external controls that you recommend or provide.
- Stability and reproducibility of the platform when used for its intended use.
- For multiplex tests in which the target molecules will contact a number of different probes, the methods used to mitigate the risk for specific and non-specific probe cross-hybridization.

- For multiplex tests that utilize many probes, the methods used to address the potential for probe cross-contamination..

We recommend that you describe in detail the test system's methodology for detecting alleles, genotypes, or mutations. You should briefly outline your risk analysis relating to the test system methodology and describe device design elements that resulted from optimization of the test system for the analyte to be tested, if applicable.

We recommend that you include illustrations or photographs of non-standard equipment or methods because these can be helpful in understanding novel methodologies and your approach to risk management, including incorporation of features to minimize potential device failures and user errors.

C. Analytical Studies

For performance data that is included in your 510(k) (21 CFR 807.92(b)), or PMA (21 CFR 814.20(b)(3)(v), we recommend that you describe the analytical studies you used to evaluate the following performance characteristics, including protocols and results. Where applicable, you should include the following information for each allele, genotype, or mutation and for each matrix claimed in the intended use statement:

1. General analytical performance considerations

You should demonstrate the device's ability to accurately and reproducibly differentiate genotypes, alleles, or mutations using both the lowest and highest nucleic acid input concentrations recommended in product labeling. When fresh samples for rare alleles, genotypes, or mutations are scarce, we will consider the use of archived or retrospective samples. Although natural samples are preferred, we will also consider artificially prepared materials, such as plasmid DNA or amplified gene segments. These artificially prepared materials should mimic natural matrices to the greatest degree possible. In particular, when using cloned or amplified material, the copy number tested should approximate that found in a natural sample. If appropriate, you should demonstrate that your assay can distinguish between hetero- and homozygotes, since this is one of the critical aspects in assessing analytical performance of a genetic assay.

2. Sample characterization and specifications

If you intend to provide reagents for specimen processing, you should demonstrate that the chosen sample preparation method consistently provides quality nucleic acid samples that yield reproducible test results for each specimen type with which your test is intended to be used. (See also Section

4, Precision.) If your sample preparation method involves preparation of an RNA intermediate, you should evaluate your procedure to ensure that residual contaminating genomic DNA is either absent or, if present, will not interfere with assay results. If you do not intend to provide sample preparation reagents in your kits, you should provide specifications for assessing the quality of the assay input sample so that users can validate their own sample preparation method and reagents. You should provide justification for these specifications in the submission. We also recommend that you carefully characterize sample stability and validate your storage and handling recommendations.

3. Effect of excess sample and limited sample

You should investigate the range of nucleic acid sample concentrations that reproducibly yield acceptable results. You should also determine the minimum amount of testable input DNA sample that provides acceptable performance and approximate the amount of patient specimen needed to generate this minimum amount of sample. We recommend that you determine the minimum amount of input nucleic acid needed to obtain a correct genotype, such that the lower bound of the 95% confidence interval for the estimated "correct call" fraction is greater than 95%. The recommended minimum sample input in your labeling may fall above this amount in order to improve performance of the test.

4. Precision (Repeatability/Reproducibility)

You should perform studies to determine estimates of total variability for each specimen type. For information on precision studies, we recommend that you consult "Evaluation of Precision Performance of Clinical Chemistry Devices;" Approved Guideline-2nd Edition, CLSI (Clinical Laboratory Standards Institute), EP5-A2 and "User Protocol for Evaluation of Qualitative Test Performance;" Approved Guideline, CLSI, EP12-A. Information on obtaining these documents is available at http://www.clsi.org/. You should include, as appropriate, repeatability (same day, site, operator, instrument, and lot) and reproducibility (between runs, days, sites, operators, instruments, and lots) studies. Precision panels should be designed to evaluate the lowest and highest nucleic acid input concentrations recommended in product labeling. We recommend that you carry out reproducibility at three or more sites. Multiple operators with skill levels the same as those of your intended users should perform the test, preferably using multiple lots of devices and reagents. You should also perform testing over several weeks and at different times of the day to maximize detection of potential sources of variability. The protocol should include evaluation of sample preparation reagents provided with the kit. If you do not include sample preparation reagents in the test kit, each site should use and validate its own specimen processing procedures and demonstrate that the resulting sample meets manufacturer-supplied

specifications. Likewise, if you do not include or recommend a specific instrument with the kit, each site should use its own instrument during testing, as appropriate.

5. Controls and calibrators

For external controls and calibrators, you should describe the following:

- nucleic acid levels
- matrix
- your method of preparation, value assignment, and validation
- your protocol and acceptance criteria for determining stability
- how you established the recommended calibration and control testing frequency

If you do not provide external controls or calibrators, you should indicate commercial availability of these materials or describe a method that users can follow to prepare them (or both). You should describe the reactions and functions monitored by internal controls. For different technologies, these controls may differ, but the controls should enable users to determine if critical reactions have proceeded properly. Controls should contain nucleic acid levels at the low end of the input concentrations recommended in product labeling in order to adequately stress the system.

6. Cut-off

We recommend you provide the following to support an analytical characterization of your cut-off(s), if applicable:

- study design and analytical data to support the established cut-off
- rationale for the units, cut-off, and/or categories of the results
- a description of specimen preparation, including analyte levels, matrix, and how levels were established
- statistical methods used [e.g., Receiver Operating Characteristic (ROC) Analysis]

7. Analytical specificity (interference and cross-reactivity studies)

Potential inhibitors present in patient specimens may not be efficiently removed by sample preparation procedures and may even interfere with sample preparation itself. We recommend that you examine potential interfering substances commonly present in the indicated patient specimens for their effects on sample preparation and assay performance. Test samples should be chosen to evaluate the lowest and highest nucleic acid input concentrations recommended in product labeling. For more information on interference studies we recommend that you consult “Interference Testing in

Clinical Chemistry;" Approved Guideline, CLSI, EP-7A. For both cross-reactive and interfering substances tested, we recommend you include the following:

- the concentrations at which these substances were present in the samples
- sample description and preparation, including matrix and nucleic acid levels
- the number of replicates tested for each substance
- how interference and cross-reactivity were defined in relation to the results obtained for the reference positive and negative control samples
- a description of the degree of interference or cross-reactivity observed
- results demonstrating that your test rejects sequences similar to the target sequence, at nucleic acid levels that include the lowest and highest input concentrations recommended in product labeling, where applicable

8. Assay conditions

As applicable, you should verify hybridization conditions (for thermocycling, cycling conditions), concentration of reactants, and control of non-specific activity. In the case of multiplex tests, you should examine and describe optimization of multiple simultaneous target detection. When thermocycling is used, you should verify optimization, specificity, and robustness of amplification.

9. Potential for sample carryover and cross-hybridization

We recommend that you assess the potential for sample carryover and cross-hybridization, and that you provide instructions in your labeling for preventing carryover and reducing or eliminating cross-hybridization.

10. Limiting factors of the device

You should describe any known limitations of the device. Examples are when the device does not measure all possible alleles, genotypes, or mutations, or when the range of alleles, genotypes or mutations is not known.

D. Software and Instrumentation

1. Data processing

If your device includes software, you should include specific information about the software in your submission. We recommend that computational methods be developed and verified using the CDRH software development and validation guidance documents that are available at http://www.fda.gov/search/databases.html.

You should provide support for your selection of the appropriate level of concern. You also must demonstrate that the software design has been verified as required by 21 CFR 820.30(f). For more information on "levels of concern," you should refer to FDA's Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices, May 11, 2005. (http://www.fda.gov/cdrh/ode/guidance/337.pdf). If applicable, you should describe how computational concerns such as probe saturation level, background correction, normalization, etc., are addressed by the software.

2. Validation of instrumentation

You should provide specifications in your labeling for any generic instrument needed to run the test, so that users may select an instrument that is suitable for their purposes. You should base your recommendations on performance testing of various instruments with your device (see Precision, Section C4, above). If you provide, or recommend, specific instrumentation for your device, whether manufactured by you or by another company, you should include specific information about the instrument(s) in your submission, and you should perform testing described in this guidance document using this instrument. We recommend that you describe the following and include test results to support your descriptions, where appropriate:

a. Characterization: You should characterize the instruments. We recommend that you include information on how the instrument assigns values to, or interprets, assay variables such as feature location, size, concentration, volume, drying of small samples, and effects of small volume reactions. You should also include information on how these types of variables impact test performance, especially results.

b. Calibration: You should describe how the instrument is calibrated and the materials used in calibration. You should indicate the recommended calibration frequency and how it was established.

c. Uncertainties: You should describe potential sources and estimates of uncertainties in results introduced by hardware components such as scanners, LCD cameras, etc.

If you specify a particular instrument (by manufacturer or brand), you should assure that any changes made to the instrument (by you or the manufacturer) are tracked and evaluated to determine whether there is any effect on assay

performance, in accordance with the quality system regulation (21 CFR Part 820). If changes in instrumentation introduce new or different assay performance issues, you will be responsible for validation of your device under the changed conditions and you should determine whether you need to submit this information to FDA. 21 CFR 820.30(g). (See http://www.fda.gov/cdrh/ode/510kmod.html and 21 CFR 807.81, 814.39.)

E. Comparison studies using clinical specimens

Where comparison studies are appropriate to establish performance of a device, we recommend that you describe your evaluation protocol and results, including the items listed below. You should include in your evaluation all matrix types with which your device is intended to be used.

1. Comparison to a Reference Method

For pharmacogenetic and genetic tests, we recommend that you validate your assay by performing studies that compare results obtained with your device to those obtained with bidirectional sequencing. Bidirectional sequencing is considered the reference method[2] for sequence analysis (sometimes also referred to as the "gold standard"). For large deletion, rearrangement or insertion mutations, or other cases where bidirectional sequencing may not be an appropriate comparator, we recommend that you consult with OIVD to discuss an appropriate study design. In your description of this study, you should include your protocol, the sample types you used, any selection criteria you applied, and results. If the population tested is representative of the population for which the device is intended, results may be reported as clinical sensitivity and specificity. You should address the quality of the bidirectional sequencing by an appropriate metric and include it in your submission.

2. Comparison to another device

You may also choose to describe comparison studies with another well-characterized or predicate device, in addition to comparison with the reference method. You should generally report results as positive and negative percent agreement. While comparison to another device can be useful, FDA believes that the best measure of test performance will come from comparison to bidirectional sequencing.

3. Resolution of Comparison Discrepancies

[2] In this document, we use the term "reference method" to refer to a well-validated analytical procedure sufficiently free of systematic or random error to make it useful for validating proposed new analytical procedures for the same analyte.

You should identify discrepant results. If you perform discrepancy resolution, you should report the result of that testing. FDA strongly discourages the use of resolved results in calculations of device performance unless unbiased statistical techniques can be used. FDA has developed a guidance document on statistical analysis that you may refer to for further discussion of this subject. See the guidance, "Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests; Guidance for Industry and FDA Staff" (http://www.fda.gov/cdrh/osb/guidance/1620.pdf).

4. Failure Rates

You should identify incorrect results obtained due to device failure. You should also provide estimates of expected failure rates (e.g., when result is "no call" due to device failure or sample inadequacy).

5. Evaluation of devices employing quantitative measurement techniques

You should evaluate the effects of random and systematic error in comparison to the reference method. You should calculate bias when possible and appropriate. For more information, you should also refer to "Method Comparison and Bias Estimation Using Patient Samples" Approved Guideline, 2nd Edition, CLSI, EP-9A2.

6. Confidence Intervals

For any device, you should calculate and report confidence intervals around the point estimates of performance measures.

F. Clinical Evaluation Studies Comparing Device Performance to Accepted Diagnostic Procedure(s)

Where clinical studies are needed to establish safety and effectiveness of a pharmacogenetic or genetic testing device, you should address the points listed below. You should include in your evaluation all matrix types with which your device is intended to be used. In addition, you may refer to Appendix I for more points to consider in designing studies. You should provide appropriate clinical data to support each intended use.

a) You should define "clinical truth" as it will be used in evaluating the clinical performance of the device. For the purposes of this guidance, we define clinical truth as the best clinical evidence for a specific diagnosis or allele assignment. If you use discrepant resolution in your analysis, you should describe the strategy used.

b) We recommend that you validate genotype/phenotype correlations, if necessary, on a statistically determined number of specimens for each intended use. You should include the following information, when defining the population(s) used:

- Number of specimens from the normal population, summarized according to appropriate demographic characteristics.
- Number of specimens included in each disease, condition, genotype, or group summarized according to appropriate demographic characteristics.

c) You should include clinical samples for all matrices included in your intended use statement. For more information on evaluating matrix effects, see "Evaluation of Matrix Effects" Approved Guideline, CLSI, EP-14A.

d) Clinical cut-off (where applicable): You should describe clinical validation of the established cut-off and its validation for the new device. You should identify clinical cut-off points in a training set and validate these in a separate, independent test data set. You may also provide literature references that support clinical cut-offs.

e) You should describe statistical methods used and confidence intervals for calculations, where appropriate.

f) If you are establishing clinical validity through retrospective or prospective studies using a genotyping method other than the method you are submitting, we recommend you consult with OIVD to determine the appropriateness of this approach for your device.

g) When you plan to use literature to support clinical validity, we recommend that you consult OIVD to determine the suitability of literature, and techniques for its evaluation, to supplement or substitute for clinical performance studies.

G. Effectiveness of the Device

For PMAs, you must provide valid scientific evidence to establish reasonable assurance of the safety and effectiveness of the new device and for de novo classification submissions, you must provide valid scientific evidence to establish which general and special controls are necessary to provide a reasonable assurance of safety and effectiveness of the device (see sections 513 and 515 of the Act (21 U.S.C. 360c, 360e)). For both PMA and de novo submissions, we recommend you submit clinical data.

1. New markers

When you demonstrate the clinical validity and test performance[3] of a device that utilizes new markers, such as mutations, patterns, and/or other outputs of pharmacogenetic and genetic tests, you must meet the requirements for determining safety and effectiveness for the tests' intended use, as outlined in 21 CFR 860.7.

2. Established Markers

For established markers, you may use appropriate information in the medical literature as evidence of the effectiveness of the marker or mutation. If you use peer-reviewed literature to support clinical validity and/or test performance, you should provide copies of all relevant articles, as well as a justification for the use of the literature in place of clinical studies. You should establish comparability between the new device and the device used in the published literature in order to ensure that the data can be confidently extrapolated. We recommend that you consult OIVD to determine the suitability of literature, and techniques for its evaluation, to supplement or substitute for clinical validity and/or test performance studies.

IV. Labeling

For 510(k)s, the submission must include labeling in sufficient detail to satisfy the requirements of 21 CFR 807.87(e). For PMAs, your application must include copies of all proposed labeling for the device. (see 21 CFR 814.20(b)(10)). The following suggestions are aimed at assisting you in preparing proposed labeling that satisfies these requirements and final labeling that satisfies 21 CFR Parts 801 and 809.

Directions for Use

[3] For the purpose of this guidance, clinical validity of the marker is defined as the determination that the device marker(s) are truly informative (e.g., predictive) as claimed in the intended use; and test performance is defined as the determination that the test reliably and reproducibly identifies the marker(s) in a clinical setting.

You should provide clear instructions that delineate the technological features of the specific device and how the device is to be used on patients. Instructions should encourage local/institutional training programs to familiarize users with the features of the device and how to use it in a safe and effective manner. Devices incorporating nucleic acid amplification should provide sample work-flow recommendations in the labeling.

Quality Control

We recommend that you provide a description of quality control measures that the laboratory should follow to help ensure proper device performance.

Interpretations and Precautions

We recommend that you provide the key for interpretations of results and specify the language to be used in reporting results. You should provide a reference for the nomenclature system that you are using to describe alleles, genotypes, and mutations.

We recommend that you address the limitations of your device with statements in the labeling, for example:

- This test does not identify all alleles of CYP2D6.
- It is suspected that some mutations, alleles, or genotypes are private (found only in a small population or single family) and may not be detected by this test. Therefore, you should interpret the results of the test with caution.
- This test should not be used for (name a reason or reasons).
- The presence of other (rare) mutations or polymorphisms may result in false positive or false negative results for this test.

Stability

You must include information on the reagent shelf life. (See 21 CFR 809.10(a)(5) and 809.10(a)(6) regarding labeling; 809.10(b)(5)(i) regarding the package insert 814.20(b)(6)(i) regarding additional information for premarket approval applications.) You should also include recommendations for handling and collecting specimens based on your specimen stability data

Performance

You should describe device performance in comparison to the reference method (bidirectional sequencing). Useful formats include 2 x 2 (or other N x N) tables, sensitivity and specificity, percent agreement, or other illustrative examples. You should calculate the sensitivity and specificity or percent agreement with their

respective confidence intervals using all tested samples. The number of failed (invalid) assays by your device (e.g., inability to genotype the sample) should be reported in your performance characteristics.

Appendix I: General considerations for planning and evaluating clinical studies

We recommend that you consult with the appropriate OIVD review divisions to determine the most appropriate strategies for your clinical studies. The following are some general recommendations that may be used when planning and evaluating clinical studies. An additional resource to consider when seeking guidance on reporting clinical and/or method comparison studies is the STARD (Standards for Reporting of Diagnostic Accuracy) statement (1), which was published in 2003 and is a roadmap for improving the quality of reporting of studies of diagnostic accuracy.[4]

1) Plan studies to support the intended use claim for the device with data that are representative of the population for whom the device is intended (e.g., ethnicity, sex, clinical condition, as appropriate).

2) Describe all protocols for internal and external evaluation studies. Clearly define the study population and inclusion and exclusion criteria and the chosen clinical endpoint. If literature is to be used to support your intended use, you should clearly explain the study population, inclusion/exclusion criteria, and endpoints in the publication and reflect how the device will be used in practice.

3) Establish uniform protocols for all external evaluation sites prior to study and follow them consistently throughout the course of data collection.

4) Use investigational sites and populations appropriate to the intended use and claims being sought. You should clearly outline efforts to define population sampling bias when this issue may impact performance.

5) Determine sample size prior to beginning the clinical study. The sample size should have sufficient statistical power to detect differences of clinical importance for each marker, mutation, or pattern. FDA will consider other approaches in cases with a small available sample size, for example, a disease allele having a low prevalence in the intended use population.

6) Describe the sampling method used in the selection and exclusion of patients. If it is necessary to use archived specimens or a retrospective design, provide pre-specified inclusion and exclusion criteria for samples, and adequate justification for why the sampled population is relevant to the patient population targeted for the intended use.

[4]Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem. 2003;49(1):7-18.

7) For genetic tests, you should include samples from individuals with diseases or conditions that may cause false positive or false negative results with the device (i.e., within the differential diagnosis), if appropriate.

8) Analyze data for each individual test site and pooled over sites, if statistically and clinically justified. Justification of data pooling over sites should address variation between sites in prevalence, age, sex, and race/ethnicity.

9) Describe how the cut-off point (often the distinction between positive and negative, or the medical decision limit) will initially be set, and how it will be verified, if appropriate. If a cut-off is specified for each of multiple alleles, genotypes or mutations, describe the performance characteristics of each cutoff as it relates to its respective allele, genotype or mutation. The description of how each cut-off is determined should include the statistical method used [e.g., receiver operating characteristic (ROC) curve].

10) Diagnostic devices that assay the presence of a particular pattern (e.g., single nucleotide polymorphism (SNP) set, haplotype pattern), should ideally be validated in a prospective clinical trial. An example of such a device would be a test using a defined SNP set to discriminate between patients who may or may not experience an adverse event associated with a particular drug. Since it is statistically problematic to validate discrimination patterns in the same study in which they were defined, the simplest way to address this is to validate the pattern with an independent data set. Determination of the statistical significance of a retrospectively determined feature pattern may not be possible or minimally would call for careful use of complex statistical procedures, such as bootstrapping, or an explicit cross-validation scheme. Given that it can be easy to obtain a low misclassification rate for a retrospectively determined feature pattern even on random data, you should provide a valid procedure for obtaining the statistical significance of such a pattern. The simplest approach statistically is to evaluate the pattern on an independent data set from a prospective clinical trial, if that is feasible.

11) Account for all individuals and samples. Perform appropriate data audits and verification before submitting to FDA. Give specific reasons for excluding any patient or test result after enrollment.

12) Perform studies using appropriate methods for quality control. Describe the materials and methods used to assess quality control.