Use of Standards in FDA Regulatory Oversight of Next Generation Sequencing (NGS)-Based In Vitro Diagnostics (IVDs) Used for Diagnosing Germline Diseases

---

Draft Guidance for Stakeholders and Food and Drug Administration Staff

DRAFT GUIDANCE

This draft guidance document is being distributed for comment purposes only.

Document issued on July 8, 2016.

You should submit comments and suggestions regarding this draft document within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to http://www.regulations.gov. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. Identify all comments with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions about this document concerning devices regulated by CDRH, contact Personalized Medicine Staff at 301-796-6206 or PMI@fda.hhs.gov. For questions regarding this document as applied to devices regulated by CBER, contact the Office of Communication, Outreach and Development in CBER at 1-800-835-4709 or 240-402-8010 or by email at ocod@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* Diagnostics and Radiological Health

Center for Biologics Evaluation and Research

Preface

Additional Copies

CDRH

Additional copies are available from the Internet. You may also send an e-mail request to CDRH-Guidance@fda.hhs.gov to receive a copy of the guidance. Please use the document number 16009 to identify the guidance you are requesting.

CBER

Additional copies are available from the Center for Biologics Evaluation and Research (CBER), by written request, Office of Communication, Outreach, and Development (OCOD), 10903 New Hampshire Ave., Bldg. 71, Room 3128, Silver Spring, MD 20993-0002, or by calling 1-800-835-4709 or 240-402-8010, by email, ocod@fda.hhs.gov or from the Internet at http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

Table of Contents

Use of Standards in FDA Regulatory Oversight of Next Generation Sequencing (NGS)-Based In Vitro Diagnostics (IVDs) Used for Diagnosing Germline Diseases

Draft Guidance for Stakeholders and Food and Drug Administration Staff

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff or Office responsible for this guidance as listed on the title page.

I. Introduction

Many advances in precision medicine will depend on the safe and effective use of next generation sequencing (NGS) technology. As part of the Precision Medicine Initiative (PMI), FDA has been focused on optimizing FDA's regulatory oversight for NGS *in vitro* diagnostic (IVD) tests to help accelerate research and the clinical adoption of precision medicine while assuring the safety and effectiveness of these tests. As part of the PMI effort, this draft guidance document provides FDA's proposed approach on the content and possible use of standards in providing oversight for whole exome human DNA sequencing (WES) or targeted human DNA sequencing NGS-based tests intended to aid in the diagnosis of individuals with suspected germline[1] diseases or other conditions (hereinafter referred to as "NGS-based tests for germline diseases" or "NGS-based tests").

This document provides recommendations for designing, developing, and validating NGS-based tests for germline diseases, and also discusses possible use of FDA-recognized standards for regulatory oversight of these tests. These recommendations are based on FDA's understanding of

[1] In this document, the term "germline diseases or other conditions" encompasses those genetic diseases or other conditions arising from inherited or de novo germline variants.

the tools and processes needed to run an NGS-based test and the design and analytical validation considerations appropriate for such tests.

FDA's guidance documents, including this guidance document, 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.

II. Background

As part of the PMI, FDA is committed to implementing a flexible and adaptive regulatory oversight approach that fosters innovation and simultaneously assures that patient test results are accurate and meaningful.

Unlike most IVDs, which are typically intended to detect a limited number of predefined analytes to diagnose pre-specified conditions, NGS-based tests can measure millions of analytes (i.e., bases) related to numerous conditions and have the potential to detect previously unidentified variants. Moreover, NGS-based tests often have broad intended uses, and the types of variants and the nature of the clinical information that will be returned from these tests is often not known until after the test has been run. Crafting the appropriate approach for regulatory oversight for NGS-based tests presents a challenge for FDA and has been considered in several discussion papers containing questions and ideas related to possible approaches. Central to these discussions is whether conformity with appropriately constructed standards for analytical validation of an NGS-based test could be useful in providing more efficient regulatory oversight.

On February 20, 2015, FDA held a public workshop entitled, "Optimizing FDA's Regulatory Oversight of Next Generation Sequencing Diagnostic Tests" to discuss and receive feedback from community stakeholders on possible regulatory approaches for tests for human genetics or genomics using NGS technology. To build on the feedback received, FDA held a second public workshop on November 12, 2015 entitled, "Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests."[2] Much of the public feedback obtained at both workshops suggested that conformity with standards for analytical validation of NGS-based tests would be a reasonable approach to allow for the differences in development and validation of these tests and could accommodate the expected rapid evolution of NGS technology. A number of stakeholder comments at the November 12, 2015 workshop suggested a need for standards covering test design and performance evaluation for NGS-based tests. FDA is unaware of any existing, comprehensive standards for analytical validation applicable to NGS-based tests for germline diseases that it believes could be used to help provide a reasonable assurance of the safety and effectiveness of these tests.

[2] FDA also held a public workshop on the use of genetic databases on November 13, 2015 entitled "Use of Databases for Establishing the Clinical Relevance of Human Genetic Variants" and another workshop on March 2, 2016 entitled "Patient and Medical Professional Perspectives on the Return of Genetic Test Results."

III. Scope

This guidance document, when finalized, will provide recommendations for designing, developing, and validating NGS-based tests for germline diseases that FDA believes are appropriate for use in providing a reasonable assurance of the analytical validity of such tests. Upon finalization of this guidance, test developers will be able to follow these recommendations when preparing a premarket submission. The recommendations in this draft guidance document are applicable for NGS-based tests for germline diseases, whether results are intended to be provided directly to patients or through healthcare professionals; however, for direct-to-consumer NGS-based tests for germline diseases additional recommendations and controls would be needed.

This draft guidance document also outlines considerations for possibly classifying certain NGS-based tests for germline diseases in class II and potentially exempting them from premarket notification requirements. Over the longer-term-, FDA will consider how these recommendations may form the basis for standards that FDA could recognize or whether FDA could establish special controls and/or conditions for premarket notification (510(k)) exemption.

The considerations and recommendations in this draft guidance are limited to targeted and WES NGS-based tests intended to aid in the diagnosis of individuals with suspected germline diseases or other conditions. A further discussion of the elements of NGS-based tests for germline diseases can be found in Section V below. This document does not apply to NGS-based tests intended for stand-alone diagnostic purposes. Additionally, this document is not intended to apply to NGS-based tests intended for screening, microbial genome testing, risk prediction, cell-free DNA testing, fetal testing, pre-implantation embryo testing, tumor genome sequencing, RNA sequencing, or use as companion diagnostics, as these may have other analytical characteristics not addressed by the recommendations presented here. FDA intends to provide recommendations and discuss pathways for additional intended uses of NGS-based tests in future guidance documents. In the interim, the public may contact FDA with questions about these issues.

IV. Classification and Premarket Review of NGS-Based Tests for Germline Diseases

To date, FDA has cleared a small number of single-gene, disease-specific, targeted, NGS-based tests.[3] However, FDA has not previously classified NGS-based tests with a broad intended use for suspected germline diseases. An NGS-based test for germline disease is a medical device of a new type that FDA has not previously classified. As a result, it is automatically classified into class III by operation of law. There are no legally marketed devices of the same type that could serve as a predicate device for review of such an NGS-based test in a premarket notification under section 510(k) of the Federal Food, Drug and Cosmetic Act (FD&C Act) (21 U.S.C.

[3] *See*, e.g., Illumina MiSeqDx Cystic Fibrosis 139-Variant Assay (k124006) and Illumina MiSeqDx Cystic Fibrosis Clinical Sequencing Assay (k132750).

360(k)).[4] Thus, these tests are at present subject to FDA approval of a premarket approval application (PMA).

A. Possible Classification of NGS-Based Tests for Germline Diseases in Class II

An applicant may submit a *de novo* request for classification of a new device type when that device is class III by operation of section 513(f)(1) of the FD&C Act (21 U.S.C. 360c(f)(1)), there is not a legally marketed predicate device on which to base substantial equivalence in a 510(k), and the applicant believes that the test is appropriate for classification in class I or class II.[5] The applicant should provide information in the premarket submission to demonstrate that general controls or general controls and special controls are sufficient to provide a reasonable assurance of safety and effectiveness for that test. If FDA grants the *de novo* request and classifies the test as class II, the test may then be marketed, serve as a predicate for future 510(k) submissions, and would be subject to both general and special controls.

Additionally, if FDA believes there is a reasonable possibility that the safety and effectiveness of the test can be reasonably assured by general controls or a combination of general and special controls, FDA may identify such a test as a suitable candidate for the *de novo* process. Because FDA believes there is a reasonable possibility that the risks associated with the use of NGS-based tests for germline diseases (e.g., those related to the consequences of a false positive or negative result provided to a patient) may be sufficiently mitigated by a combination of general and special controls, and that the safety and effectiveness of this type of test may be reasonably assured by such controls, FDA believes that an NGS-based test for germline disease can be a suitable candidate for the *de novo* classification process. FDA encourages applicants to engage with the Agency using the Pre-Submission process to discuss any anticipated *de novo* requests for NGS-based tests for germline diseases.

B. Possible Exemption of NGS-Based Tests for Germline Diseases from Premarket Notification Requirements

FDA may exempt a class II device from the premarket notification requirements of section 510(k) of the FD&C Act on its own initiative or upon petition of an interested person, if FDA determines that premarket notification is not necessary to provide a reasonable assurance of the

[4] *See* section 513(f)(1) of the FD&C Act (21 U.S.C. 360c(f)(1)).
[5] *See* section 513(f)(2) of the FD&C Act (21 U.S.C. 360c(f)(2)). The Food and Drug Administration Modernization Act (FDAMA) of 1997 provided FDA with the authority to evaluate automatic class III designations for possible classification in class I or II through the *de novo* classification process for devices that were found to be not substantially equivalent (NSE) to a legally-marketed predicate device through 510(k). The Food and Drug Administration Safety and Innovation Act (FDASIA) of 2012 amended section 513(f)(2) of the FD&C Act (21 U.S.C. 360c(f)(2)) to provide that sponsors may submit a *de novo* without having to first submit a 510(k) and receive an NSE decision. Further information about the *de novo* process can be found on FDA's website.

safety and effectiveness of the device.[6] There are a number of factors FDA may consider to determine whether a 510(k) is necessary to provide a reasonable assurance of the safety and effectiveness of a class II device. These factors are discussed in the January 21, 1998, Federal Register notice (63 FR 3142) and subsequently in the guidance the Agency issued on February 19, 1998, entitled "Procedures for Class II Device Exemptions from Premarket Notification, Guidance for Industry and CDRH Staff." FDA believes that these factors may not be appropriate for assessing the need for a 510(k) to provide a reasonable assurance of the safety and effectiveness for NGS-based tests for germline diseases. Because of the unique features of NGS-based tests for germline diseases, FDA believes instead that special controls and/or conditions of exemption, where appropriate, could be developed for these types of tests that could provide the same reasonable assurance without a 510(k). Accordingly, we propose that this guidance, when finalized, will supersede the aforementioned guidance as it applies to NGS-based tests for germline diseases.

Should FDA allow an exemption from the requirement of premarket notification, the device would not be exempt from any other statutory or regulatory requirements, unless such exemption is explicitly provided by order or regulation. Furthermore, this would not alter any "limitations of exemption" that apply to a 510(k)-exempt type of device, and 510(k) clearance would still be required prior to marketing such a test. All 510(k)-exempt devices are subject to the limitations of exemption found at 21 CFR parts 862 to 892 at section .9 of each part, which limit exemptions to devices with the same indications and technological characteristics or ones with reasonably foreseeable differences.

If FDA were to classify NGS-based tests for germline diseases in class II (e.g., in response to a de novo request), FDA would consider exempting such class II NGS-based tests from premarket notification requirements. In determining whether a 510(k) would be necessary to provide a reasonable assurance of the safety and effectiveness of the test, FDA would rely upon the recommendations below, in addition to other considerations, including assurance of the clinical validity of the test.

Conformity with an FDA-Recognized Standard for Supporting or Assuring Analytical Validity

FDA believes that the recommendations in Section VI below can help assure the analytical validity of an NGS-based test for germline diseases. FDA may also consider recognizing standards developed by the scientific community or by standards development organizations (SDOs) that have criteria similar to the recommendations provided in Section VI. Conformity with such recognized standards may be appropriate to support or provide a reasonable assurance of analytical validity.[7] Alternatively, these recommendations may form the basis of special

[6] Section 510(m)(2) of the FD&C Act (21 U.S.C. 360(m)(2)) requires that, before granting an exemption from 510(k), FDA must publish a *Federal Register* (FR) notice of its intent to exempt the class II device type or of a petition to exempt a class II device type if one is submitted. This FR notice will provide a 30-day period for public comment. After consideration of public comment, within 120 days of publishing this FR notice, FDA will publish an order in the FR of its final determination regarding the exemption of the device type.

[7] FDA has not yet determined how conformity with standards for NGS-based tests should be demonstrated and plans to discuss this in future guidance documents.

controls. Note that these recommendations address analytical validity, and whether used as special controls or standards, conformity will not provide support for clinical validity, which is also required for a demonstration of reasonable assurance of the safety and effectiveness of the test.

Public Availability of and Access to Performance Information

FDA has long believed that making its reviews of cleared and approved products available is important so that all interested persons (e.g., healthcare providers and patients), can see, for these products, the performance that FDA has cleared or approved. To that end, for all IVDs that have received clearance or *de novo* classification from FDA since November 2003, FDA has published a Decision Summary containing a review of the analytical and clinical validity data and other information submitted by the applicant to support the submission and FDA's justification for clearing the IVD; FDA is also required to publish Summaries of Safety and Effectiveness Data for approved PMAs under section 520(h) of the FD&C Act (21 U.S.C. 360j(h)).[8] FDA believes that similar public availability and access for information regarding NGS-based tests, regardless of whether they are FDA reviewed or exempt from 510(k), is important so that patients and healthcare providers can have access to information about the capabilities and limitations of these tests in order to make fully informed medical decisions.

V. Elements of an NGS-Based Test for Germline Diseases

NGS-based tests for clinical use typically include reagents, consumables, instruments, and software. The determination of which reagents, consumables, instruments, and software are suitable for achieving the intended purpose for a particular indication is dictated by the particular attributes necessary for proper and consistent functioning. For this reason, any two NGS-based tests may differ in their design and workflows.

NGS-based tests may encompass the following steps: (a) specimen collection, processing, and storage, (b) DNA extraction, (c) DNA processing and library preparation, (d) generation of sequence reads and base calling, (e) sequence alignment/mapping, (f) variant calling, (g) variant annotation and filtering, (h) variant classification/interpretation, and (i) generation of test report. Certain of these may not always be considered to be part of the test, depending on the design of the specific test. Manual variant interpretation, performed by healthcare providers and laboratory professionals, is not considered part of the test, but certain standard operating procedures (SOPs), decision matrices, and some software products may be considered test components. FDA recommends that applicants discuss their particular tests through a Pre-Submission as early as possible in the development of the test.

[8] No Decision Summaries or Summaries of Safety and Effectiveness Data are posted for those devices for which the applicant failed to demonstrate substantial equivalence or a reasonable assurance of the safety and effectiveness of the test.

VI. Recommendations for Design, Development, and Validation of NGS-based Tests for Germline Diseases

FDA believes that one approach for supporting the analytical validation of NGS-based tests may be through conformity with one or more FDA-recognized standards (if available) or special controls. This approach should allow a test developer to design, develop, and validate an NGS-based test with a range of design and performance characteristics consistent with the intended use discussed in this guidance.[9]

FDA believes that for a standard to be recognized by FDA it should include, among other things, a description of the design activities that should be carried out and the performance characteristics that should be validated, as well as specific methodology, materials, and performance thresholds, where appropriate and justifiable. FDA expects that demonstration of conformity with such standards may be used by developers of NGS-based tests for germline diseases in premarket submissions, and possibly in the future in lieu of premarket review. However, the adequacy of a declaration of conformity with FDA-recognized standards for analytical validity may depend on the specific intended use and the type of premarket review, or, potentially, exemption.

For a standard to be recognized by FDA, the standard should include, at a minimum, the design, development, and validation activities outlined in this section. These are a combination of test design activities, performance metrics, and thresholds that FDA believes can help demonstrate a reasonable assurance that an NGS-based test for germline diseases is analytically valid.

The recommendations below relate to how a test is designed, developed, and validated. As a general principle, test developers should first define the indications for use statement of their test, as this determines how the test should perform. When defining appropriate test performance, developers should prospectively determine the types of studies that should be conducted (e.g., accuracy) as well as the thresholds that should be met for each in the form of a minimum and target value. After design and development of the test, validation studies will indicate if the predefined performance is met. If the test does not meet any one of the predefined performance specifications, the test should be modified and revalidated. The cycle of design, development, and validation should continue until the test meets the predefined performance specifications. Throughout this process, test developers should document all activities, decisions, and outcomes, along with the justification for each of these activities.

The descriptions provided below only apply to the type of test described in this guidance document.

[9] The test's labeling (e.g., test report, information about the test's performance, and other written information accompanying the test) must comply with all applicable labeling requirements, such that it is truthful and not misleading and bear adequate directions for use. *See, e.g.,* sections 502(a) and (f) of the FD&C Act (21 U.S.C. 352(a) and 352(f)).

A. Test Design Considerations

A test's conformity with an FDA-recognized standard can help demonstrate that an NGS-based test developer has performed the activities necessary to identify the intended clinical use of the test and to design the test for that use. A design and development standard or standards should address the competence of the test designer to perform and record the activities discussed below in order to yield a test that has the intended characteristics and consistently delivers results within predetermined acceptance intervals or thresholds. During the test design phase, developers should establish and justify minimum acceptable and target values for each performance metric appropriate for the indications for use of the test. Standards can provide additional explanation, examples, formats, and other information.

1. Indications for Use Statement(s) of the Test

Prospectively define and document the specific clinical need that is driving the development of the test. This will usually include specifying the disease or other condition of interest, the clinical use of the test, and the population that the test is intended to be used for (i.e., the target population). It may also be informative to define and document the clinical setting (if other than a general one), in which the test is to be offered.

Examples of common clinical uses under the broad indications for use statement considered here include: aid in diagnosing children with signs and symptoms of developmental delay or intellectual disability, patients with undiagnosed diseases, patients with hereditary cancer syndromes, etc.

Examples of target populations include: patients with signs and symptoms of a specific disease or other condition, patients in a particular age range, patients of the same sex. Examples of considerations for target populations would include the fraction of the affected population for which the test is expected to provide results, or the prevalence of the specific disease or other condition targeted by a test, if applicable.

2. Specific User Needs for the Test

Prospectively determine and document, through consultation, professional experience, professional guidelines, and other relevant sources, specific test features that are needed to assure development of a test that meets user needs.

The specific user needs will define critical factors to address during test design. It may be helpful to prioritize user needs so that the most critical ones receive the greatest design attention.

An example of a specific user need for the test includes: when a user has to process large numbers of samples within a limited turn-around time. This user need will help determine which NGS platform should be used as part of the test, how multiplexing is performed, and, potentially, how other aspects of the test are designed.

3. Specimen Type

Specify and document the acceptable specimen types to be used for the test.

Specimen types accepted for testing will raise questions in design such as the type of collection device required, minimum volume or quantity of sample, any collection conditions that must be adhered to for sample stability between collection and use. Multiple specimen and collection types may be appropriate for a test, but each type should be validated for use in producing DNA of the appropriate quality and quantity and for overall test performance. Appropriate specimen types may depend on the use of the test.

Examples of specimen types include: whole blood, ethylenediamine tetraacetic acid (EDTA)-preserved blood, buccal swab.

4. Interrogated Regions of the Genome

Specify and document the region(s) of the genome, including genes and variants, that will be interrogated by the test. If necessary, pre-specify what will be reported in the event only a portion of sequenced targets are requested by the ordering clinician.

The types of genes sequenced and/or reported will depend on the specific indications for use, which in turn will influence aspects of test design and definition of test performance.

Example: A test intended to diagnose suspected genetic disorders in newborns may use WES rather than a more restricted panel of genes with well-defined clinical significance. In such a case, the test may be configured to report only a subset of genes from WES that may be related to suspected disease(s) or other condition(s) based on a patient's phenotype, clinical presentation, and previous available test results for the patient. For instance, a test might only report results from genes known to be related to cardiac disorders when such disorders are suspected based on clinical presentation.

5. Performance Needs

To demonstrate performance needs, consider the following:

- *Define and document a minimum set of metrics (e.g., accuracy) that should be evaluated for an adequately analytically validated test.*
- *Define and document appropriate performance thresholds for those metrics based on the test's indications for use statement and predefined user needs.*
- *Define and document the degree to which interrogated regions that do not meet test run quality metrics (e.g., depth of coverage; see Section VI.C) can be included in the test.*
- *Identify and document the use of secondary procedures (e.g., familial testing, orthogonal confirmation of results), as their use may affect performance needs.*
- *Document possible limitations to test performance.*

Example: If specimens are limited (e.g., specimen volume is small, it will not be possible to collect additional specimens from a patient) or if results will not be confirmed by an orthogonal method, the minimum accuracy of the test should be higher in regions from which results will be reported. Similarly, if an interrogated genomic region is difficult to sequence, this should be reported as a test limitation, and may inform the inclusion of confirmation by an orthogonal method during test design or may necessitate higher coverage in that region.

6. Components and Methods

a. Component Specification

Specify and document all test components (e.g., instrumentation, software, consumables, reagents), including those for procedures (e.g., materials for library preparation) and general laboratory equipment used for the test (e.g., automated liquid handlers). For each step of an NGS-based test, set technical specifications (e.g., throughput of a sequencing platform) for test components based on identified user needs, the indications for use statement, and predefined performance. Document the limitations of each component for critical factors (e.g., coverage, multiplexing).

Specifications should be determined and documented for each component of an NGS-based test. These specifications are generally driven by user needs, the indications for use statement, and the performance specifications. In some cases, test design issues may feed back into the indications for use statement or predefined performance specifications. For instance, there may be a need to modify the indications for use statement to fit limitations imposed by the availability of a specific sequencing platform.

Listed below are recommendations for select components or steps of an NGS-based test:

i. *Sequencing Platform*

Specify the sequencing platform that will be used.

The particular sequencing platform should have specific performance characteristics that align with user needs and the indications for use statement of the test.

ii. *Controls and Reference Materials*

Specify controls and reference materials for achieving confidence in the test.
These should include per sample, per run, etc., as needed, in order to establish the quality of performance. They can also include gene and disease specific controls for detecting common pathogenic variants used to diagnose well-defined diseases or other conditions, pan-disorder positive controls (most common pathogenic variants), and other appropriate controls and reference materials.

iii. *Bioinformatics*

- *Describe and document data processing and analysis, including all procedures for variant calling, filtering, and annotation.*
- *Specify and document all software to be used, including the source (e.g., developed in-house, third party), and any modifications.*
- *Document software versions and traceability, reference sequence assembly, and components needed to compile, install, and run bioinformatics pipeline.*
- *Specify and document whether software will be run locally or remotely (e.g., cloud-based).*
- *Specify and document which databases will be used (if any), and whether these are internal or third-party.*

The bioinformatics pipeline should be selected based on the type of sequencing and the types of variants that will be reported, and considering any limitations of the pipeline in variant calling and interpretation. Inclusion in the test design of third party bioinformatics tools should be done by documenting and validating bioinformatics software performance in the context of the end-to-end NGS-based test.

b. Methods

Develop and document procedures and methods for running the test. Document in detail methods for each step of the test (e.g., DNA extraction, multiplexing). Develop and document procedures for using instruments, consumables, reagents, and supporting methods. Identify and document limitations, if any, for each step, including the potential impact on other steps. Identify and document, as applicable, the type of sequencing that will be used (e.g., single-end/pair-end/mate-pair sequencing).

Specifications should be determined and documented for each method required for the NGS-based test. These specifications are generally defined by user and performance specifications.

Below is a list of recommendations for select components of an NGS-based test:

i. *Sample Preparation and Input*

- *Establish and document specific methods for specimen handling, preservation, processing, storage, and rejection criteria, as applicable.*
- *Specify and document methods that will be used for determining DNA quantity and quality.*
- *Establish and document whether the test can be run when the sample is extracted DNA from outside sources, and establish and document the requirements for such outside samples.*

ii. *Multiplexing*

- *Specify and document the number of samples that may be multiplexed in a single test without negatively affecting quality scores or coverage in important interrogated regions.*
- *Specify and document the composition of barcodes and the procedures for their use, including any required procedures for avoiding barcode collision, mis-identification or mis-sorting.*

iii. *Library Preparation and Target Enrichment*

- *Establish and document specific methods for library preparation and target enrichment (e.g., amplicon-based, capture-based), as applicable.*
- *Specify and document performance metrics (e.g., on-target sequencing, uniformity, library complexity) and the threshold that will be used for accepting the method.*

iv. *Follow-up Procedures*

Define and document the procedures to be used when a test run fails (e.g., due to failure to meet one or more of its test run quality metrics).

Such procedures may include fill in of certain regions that failed to meet appropriate quality metrics or Sanger confirmation of test results, for example.

B. Test Performance Characteristics

Analytical test validation involves measuring a test's analytical performance over a set of predefined metrics to demonstrate whether the performance is adequate for its indications for use and meets predefined performance specifications. This typically involves evaluating whether the test successfully identifies or measures, within defined statistical bounds, the presence or absence of a variant that will provide information on a disease or other condition in a patient. For sequencing outside of specific targeted regions, the ability to routinely detect the "wild type" sequence may be sufficient to establish accuracy in these areas. Once all methods are finalized and documented, and the end-to-end performance of the test is validated for the test's indications for use, test performance should be continuously monitored during clinical use. It is generally important, as part of test design and development, to validate individual steps of an NGS-based test and to verify that components are operating as expected. The complete NGS-based test should be analytically validated in its entirety (i.e., validation experiments should be conducted starting with specimen processing and ending with variant calls, and performance should be documented) prior to initiating clinical use of the test.

This section recommends a set of performance metrics that should be accounted for when analytically validating NGS-based tests for germline diseases. Note that for some of the metrics listed below, FDA provides recommendations for minimum performance thresholds.

1. Accuracy

Demonstrate accuracy by measuring positive percent agreement (PPA), negative percent agreement (NPA), technical positive predictive value (TPPV), and the rate of "no calls" or "invalid calls". Set thresholds for PPA, NPA, and TPPV that assure that the test will meet its predefined performance specifications.[10]

FDA recommends that PPA, NPA and TPPV be set at no less than a point estimate of 99.9% with a lower bound of the 95% confidence interval (CI) of 99.0% for all variant types reported by the test.

Accuracy involves determining the closeness of agreement between a measured value and a true value of a measure. For NGS-based tests, accuracy represents the degree of concordance (or agreement) of results between a sequence obtained from the test and the same sequence determined by a valid comparator method identified as appropriate by FDA, or between a reference sample run on an NGS-based test and the known sequence of the reference. The minimum acceptable overall and target accuracy of an NGS-based test may vary depending on the type of variations and on whether variants are confirmed using an orthogonal assay.

a. Positive Percent Agreement

Calculate and document PPA as the number of known variants detected by the test (true "positives" (TP)) divided by the number of known variants tested (TP plus false negatives (FNs)). Calculate and document PPA for each variant type.

PPA is the ability of the test to correctly identify variants that are present in a sample. PPA reflects the frequency of FNs.

b. Negative Percent Agreement

Calculate and document NPAs as the number of true "negative" (TN) results divided by the number of wild type results for variants tested (TN plus FP) for each variant type that is being reported.

NPA is the ability of the test to correctly identify wild-type (wt) bases (i.e., the probability that the test will not call a variant that is not present). NPA reflects the frequency of FPs.

c. Technical Positive Predictive Value

Calculate and document TPPV by dividing the number of TPs from the test by the total number of positive results (TP plus FP) obtained by the test.

[10] Based on different scenarios or the methodology used, additional metrics for evaluation of accuracy may be developed.

TPPV relates to the likelihood that a variant call is a TP.

d. "No Calls" and "Invalid Calls"

Determine and document the rate of "no call" and "invalid call" results in the accuracy study.

Do not use "no calls" or "invalid calls" in PPA, NPA, or TPPV calculations.

Minimum acceptable values for "no calls" or "invalid calls" will depend on indications for use and test design. For example, a test for which results should be generated with a short turnaround time may require that the rate of "no calls" or "invalid calls" be minimal.

2. Precision (Reproducibility and Repeatability)

Evaluate precision (reproducibility and repeatability) for both variant and wild type calls, with each metric separately reported for each condition, interrogated region, and variant type. Test important factors that may contribute to test variability, including multiple samples, runs, reagent lots, and operators. Test other sources of variability as applicable, including multiple instruments, multiple testing sites, lane replicates, and lanes.

FDA recommends thresholds for reproducibility and repeatability that meet or exceed 95.0% for the lower bound of the 95% CI, calculated by conditions tested and genomic context, separately for each variant type.

Reproducibility for NGS-based tests involves measuring test variability under a variety of specified conditions (such as when using different operators, different operating conditions (if applicable), different days of measurement, or different components (if applicable)) using the same sample, and accounting for major sources of variability in the test. Repeatability involves measuring test result variability when using the same operators, the same measuring system (e.g., the same instrument and components), the same operating conditions and the same location, and replicating measurements on the same or similar objects over a short period of time. These studies do not require a gold standard sequence for comparison; rather, test developers should compare their replicates and calculate pair-wise positive agreement or pair-wise negative agreement.

3. Limit of Detection (LoD)

Establish and document the minimum and maximum amount of DNA (e.g., acceptable input range) that will enable the test to provide expected results in 95% of test runs with an acceptable level of invalid or "no calls" results (i.e., without a loss of accuracy). Establish and document the lower LoD for each variant type included in the test's indications for use. If testing specimens with mixed content (e.g., mosaic specimens), establish and document the ability of the test to detect different allele ratios and determine the lower LoD of variants based on dilution assays, performed by mixing two pure clinical samples or creating blends from cell lines that represent a range of percentages.

The LoD for an NGS-based test should be evaluated under different routine clinical laboratory conditions and in a defined specimen type. In general, the (lower) LoD is calculated as the lowest concentration of analyte at which at least 95% of positive calls and an acceptable level of invalid or no calls is obtained among the replicates tested for that concentration. When different variant types may have different LoDs, calculate the LoD for representative variants. Similarly, an upper limit of detection should be established and documented.

4. Analytical Specificity

Establish and document analytical specificity using the metrics listed below. Establish and document whether, using proposed methods, potential interfering and cross-reacting substances or cross-contamination affects the test performance. If interfering, cross-reacting substances, or cross-contamination affect test performance, revise methods or performance specifications to exclude their effect.

a. Interference

Identify and document any interfering substances (including matrix effects) that might reduce the ability to amplify or sequence. Select substances for interference experiments that are relevant to specimen or sample types covered by the test's indications for use.

b. Cross-Reactivity

Assess and document the potential for cross-reactivity of known cross-reactive alleles and homologous regions (e.g., pseudogenes), based on the targets that will be interrogated by the test.

c. Cross-Contamination

Develop, validate, and document methods to detect carryover or cross-contamination between patient specimens or samples.

Analytical specificity relates to the ability of a test to measure solely the intended analyte. Interference in measurement from endogenous or exogenous substances that may be expected based on the indications for use and test design may result in failure to detect an analyte, yielding false negative results. Cross-reactivity (e.g., from homologous regions, pseudogenes and other type of cross-reactive sequences) may result in erroneous detection of an incorrect analyte, yielding false positive results. Cross-contamination of patient specimens introduces incorrect sequences into the test which can lead to false positive and false negative results.

C. Test Run Quality Metrics

Establish and document minimum acceptable thresholds for coverage, base quality, and other test run quality metrics relevant to the specific design and test processes (e.g., input DNA quality, library complexity, bioinformatics pipeline related metrics).

Test run quality metrics are used to determine whether an individual test run or variant call should be accepted, or, when applicable, whether supplemental procedures should be used to further query a variant call. A number of test run quality metrics associated with the whole test or specific steps or components of an NGS-based test may be used. These metrics are described below.

1. Coverage (Read Depth and Completeness)

Establish and document minimum performance thresholds for average and minimum depths of coverage, uniformity of coverage, and the percentage of bases in the target region(s) above the minimum depth of coverage for the test.

For detecting germline heterozygous variants using a targeted panel, set a threshold of 20X or greater for minimum coverage depth and 300X for average coverage depth at 100% of the bases for targeted panels and at least 97% of the bases for WES.

If critical interrogated regions do not meet minimum coverage thresholds, revise methods to enable the test to reach minimum coverage thresholds or revise test claims to limit the types of results reported.

Supplemental procedures (*see* Section VI.E below) may have to be incorporated into the testing scheme to address interrogated region coverage problems.

Selection of thresholds should demonstrate adequate test performance for the indications for use statement and predefined user needs. Minimum coverage and related metrics will vary based on the details of a test's indications for use, design (e.g., instrumentation), procedures (e.g., testing of familial trios vs. testing of patients only), and performance (e.g., base-call error rates, number of independent reads). For instance, higher coverage thresholds should be considered for detecting variants from mixed or mosaic specimens (e.g., germline mosaicism). FDA does not intend to recommend specific thresholds for coverage metrics in most instances. However, FDA believes that, for any test, thresholds should not be set below the levels specified below.

2. Test Run Metrics and Performance Thresholds

FDA recommends establishing test run metrics and performance thresholds for all critical NGS-based test steps. These metrics and their performance thresholds are assessed in test validation. If validation results indicate that the metrics are not appropriate for the test, or that the performance thresholds cannot be met, the test design should be modified.

The following is a list of factors for establishing test run metrics and performance thresholds for test elements:

a. Specimen Quality

Establish and document criteria for accepting or rejecting specimens.

b. DNA Quality and Processing

- *Establish and document thresholds for genomic DNA concentration, volume, and quality.*
- *Establish and document methods for the evaluation of quantity and concentration of DNA (e.g., fluorometric methods).*
- *As applicable, establish and document the acceptable DNA size range and/or mode of range after shearing, establish and document performance thresholds for library yield, and establish and document target enrichment method.*

These methods and thresholds will influence the selection of the appropriate DNA extraction method.

c. Sequence Generation/Base-Calling

- *Establish and document a threshold for base quality score (Q score) for sequencing reads.*
- *Establish and document thresholds for median base quality by cycle and percentage of bases above a predetermined quality threshold.*
- *If applicable, establish and document a threshold for percentage of trimmed bases.*

FDA recommends a base quality score of at least 30. Other methods for evaluating base quality may also be appropriate. If Q score is not used, document the method used and why it is an appropriate method.

Other metrics of sequence generation may be used, if appropriate. Examples of these are:

- Cluster density and cluster passing filter rate.

- Reads (e.g., number of reads); percentage of unique reads (before removal of duplicates); percentage of duplicate reads (which reflects the number of reads that start at the same position and is an indicator of library complexity).

If such other metrics are used, thresholds should be established and documented for each metric.

d. Mapping or Assembly Metrics

Establish and document appropriate metrics and their associated thresholds for mapping quality.

Examples of possible metrics include:

- Percentage of reads mapped to the reference genome.

- Percentage of reads mapped to the target region.

- Mapping quality scores and percent of reads correctly mapped.

- Percentage of target covered, percentage of reads mapped to off target/decoy sequences, and percentage of reads not mapped to any human sequence.

- Depth of coverage (*see* Section VI.C.1 above).

- Non-specific mapping such as misaligned or clipped reads due to large indels, non-specific mapping due to sequence homology, and mapping errors assessed using a pan-ethnic reference sequence.

If upon test validation, critical bases/positions do not meet mapping quality thresholds, the test design and/or the metrics and thresholds should be evaluated for appropriateness for the indications for use, including user needs. Alternatives such as supplementing the NGS-based test with a second method for such regions, or specifying the regions not reported (and modifying the test's indications for use statement and limitations accordingly), may be acceptable when a small number of bases/positions are known to map poorly.

e. Variant Calling Metrics

Establish and document the appropriate metrics and their associated thresholds for variant call quality.

Variant calling metrics include single variant metrics and overall variant calling summary metrics. Appropriate metrics may depend on the bioinformatics pipeline used for variant calling.

Examples of appropriate metrics include:

- Variant call quality score.

- Number and percentage of reads with the variant reported.

- Allelic read percentages, including percent of different variant types (e.g., heterozygous calls, indels, nonsense variants), and portion and ratios of base substitutions (transition/transversion (ti/tv)).

- Variant allele frequency (e.g., expected call frequency thresholds/minimum percent of variant reads defined for homozygous and heterozygous calls).

- Percent of novel variants, concordance rates with reference variant/sequence.

- Strand bias.

- Percentage of claimed region covered / percent completeness (i.e., percent of test with sufficient coverage above minimum threshold).

Systematic error profiles and suppression may need to be considered or incorporated in pipeline development. If this is the case, establish and document the method for profiling and suppression.

D. General Recommendations for Performance Evaluation Studies

When evaluating a test design and configuration, incorporate the features listed below into performance evaluation studies, as applicable. Provide a detailed justification if there are any deviations from, or deletions or additions to, these recommendations.

- Perform validation studies on genomic regions, variant types, and sequence contexts representative of the test's indications for use, including clinically relevant targets. Establish performance of variants in highly homologous, highly polymorphic, or other difficult regions if these regions are part of the indications for use of the test. Account for variant prevalence when selecting specific variants to include in accuracy studies. For indels, include a distribution of variants in increments of no more than five base pairs, for both insertions and deletions.

- Assess test limits, such as insertions or deletions larger than a certain size and rearrangements, and identify types of sequence variations that the test cannot detect with the intended accuracy and precision.

- Use specimens that reflect the actual specimen types (e.g., whole blood, saliva) and population that the test developer has established as acceptable for clinical testing. If necessary, supplement clinical specimens with well-characterized samples containing known sequence variants (e.g., from cell lines). Conduct commutability studies if inferring performance based on validation using plasmids or other synthetic constructs.

- Include specimens and DNA samples representing different variant genotypes (i.e., wild-type, heterozygous, compound heterozygous, homozygous) consistent with the test's indications for use statement.

- Include DNA preparation, specimen and reagent acquisition, handling and storage (where applicable) when evaluating end-to-end test performance.

- Evaluate test performance for different allele ratios if specimens or DNA samples with mixed content (e.g., mosaicism) are being claimed in the indications for use of the test. This may be performed by mixing two pure clinical samples or creating blends from cell lines covering a range of allele fractions.

  Determine the number of specimens required to demonstrate that performance thresholds have been met with confidence for relevant metrics. This number will depend on the indications for use of the test and the critical performance parameters that must be met (e.g., how many types of variants the test is expected to detect, and the number of variants of different types in a given validation sample) to support that use.

- Include the finalized bioinformatics pipeline for data processing and analysis as a part of the overall beginning-to-end test validation. The performance of the bioinformatics pipeline can be established and documented by analyzing data files containing known sequence variants of various claimed types (e.g., single-nucleotide variants, small indels, large CNVs, structural variants). Those data files should, however, be generated using the test's pre-analytical and analytical methodology

- If applicable, validate sample pooling methods, including minimum and maximum number of multiplexed samples, to ensure that individual sample identity is maintained. If barcoding is used for multiplexing, establish and document that there is no crosstalk between samples with distinct barcodes and that the combinations of patients/barcodes in a run provide accurate and reproducible results for all amplicons regardless of which barcode is used for each sample, and when the maximum number of samples is multiplexed.

When evaluating NGS-based test accuracy:

- Evaluate and document accuracy by comparison to a method identified as appropriate by FDA, such as bidirectional sequencing or another well-validated method. As an alternative comparator method, supplement accuracy evaluation using a comparison of the sequence generated by the test to a consensus sequence of agreed-upon well-characterized samples, if such samples are appropriate.

- Calculate PPA, NPA and TPPV separately for each type of variant claimed (e.g., single nucleotide variants, indels, structural variants) and sequence context (e.g., highly homologous regions) to be assessed by the test.

When documenting the results of validation studies:

- Present results as a mean and associated 95% two-sided CI. Present results in a tabular format, with results documented separately for each variant, variant type tested, and sequence context. Where relevant (e.g., for insertions, deletions), document results by size distribution.

- Present results separately for each specimen type used for validation, and indicate the type of specimen used (e.g., clinical specimen, cell line).

- For reproducibility studies, document results for each variant or variant type. Indicate the number of replicates tested for each variant and the conditions that were tested (e.g., number of runs, days, instruments, reagent lots, operators).

- When presenting the results of reproducibility and repeatability studies, indicate the failed quality control rate, and list all "no calls" or "invalid calls." Data from runs that do not meet coverage depth, coverage uniformity, and other technical metrics are typically considered quality control failures.

E. Supplemental Procedures

Include any applicable supplemental procedures (e.g., orthogonal confirmation, fill-in, trio testing) whose reflex use will be directed in the test's instructions in design, development and validation activities and documentation. If supplemental procedures are not performed, document the types of results that will not be reported by the test.

Supplemental procedures refer to those procedures that are not part of the core process for generating variant calls from input specimens or DNA, although they may be considered part of an NGS-based test. Supplemental procedures, such as fill-in or orthogonal confirmation, should be implemented when variants or interrogated regions of the genome that are critical parts of the indications for use of the test cannot meet predefined test run quality metrics or performance thresholds. In these cases, supplemental procedures may be established to assure that the test can reliably report on variants in those regions. Furthermore, for some rare undiagnosed diseases, sequencing trios or additional familial testing is recommended, and test results may be inconclusive without the appropriate parental or familial testing.

For example, there may be a need to perform confirmatory testing for critical variant types where lower bound of the 95% CI for accuracy falls below 99.0%. Alternatively, adequate justification for reporting those variants with lower accuracy can be provided based on other means.

F. Variant Annotation and Filtering

Select filtering algorithms appropriate for the indications for use of the test, establish and document filtering thresholds, and document how and when filtering will be used. Document any filtering criteria that are applied and describe their purpose, e.g., eliminating from consideration variants of low allele frequency, difficult-to-sequence regions or variants that are hard to call or analyze, filtering out specific type of variant, etc. When using databases to aid in annotation and filtering (e.g., estimating allele frequency from large control cohorts such as those found in the Exome Aggregation Consortium (ExAC) or 1000 Genomes databases), verify that the indicated population of the test is included in the dataset, and record the version of the database used. Include a process to identify and incorporate changes in external sources of data into the annotation and filtering procedures.

Filtering algorithms to identify and prioritize candidate causal variants or genes from exome or genome sequencing can include selecting variants based on population frequency, prioritization based on impact on gene and gene production function and/or phenotypic data, probabilistic methods, or shared genomic segments (e.g., regions of identity by descent and co-segregation of variants with phenotype in family studies).

G. Presentation of Test Performance

The following should be included when providing information on test performance:

- Make available and provide public access to the test's indications for use statement, limitations, and summary performance information via a prominent hyperlink on the company website.

- Include the following in the indications for use statement of the test:

    - Type(s) of sequence variations (e.g., single nucleotide variants, multiple nucleotide variants, insertions, deletions) detected as a part of the test.
    - Any limitations of the test (e.g., interrogated targets such as genes or types of sequence variants that the test cannot detect with validated performance, failure to detect insertions and deletions larger than a certain size).
    - The fraction of the affected population for which the test is likely to provide relevant results, for example, if the test only detects a subset of all variants that are causative of a particular disease or condition.

- Identify region(s) of the genome in which sequence meeting pre-specified performance specifications can be generated by the NGS-based test.

- List types of variants that the test will report using a widely accepted nomenclature.

- For targeted NGS panels, list the gene(s) included on the panel using a widely accepted nomenclature.

- For WES based tests, describe how known, clinically relevant regions of the exome are defined, and the relevant coverage for those regions.

- In the summary performance information, include:

    - Results for test accuracy and precision/reproducibility presented in a tabular format, across the regions queried by the test, by variant type and size (e.g., sizes that include distribution of results by 5 and 10 bps, separately for deletions and insertions, by polymorphic and non-polymorphic regions), summarized as a mean percent agreement and disagreement with the reference sequences and 95% CI, separately for positive and negative results, and broken down by whether results were generated from clinical specimens, contrived samples, cell lines, or reference sample sets.

    - For results of reproducibility studies, list the number of replicates for each variant/variant type, and conditions tested (i.e., number of runs, days, instruments, reagent lots, sites, operators, specimens/type, etc.).

    - For targeted panels, indicate the average depth of coverage and the percentage of target region covered at the minimum depth of coverage.

    - For WES, indicate the average depth of coverage and the percentage of target region covered at the minimum depth of coverage.

- Provide information about the probability of test failure based on performance data (e.g., failed quality control). Describe scenarios in which a test can fail (e.g., low sample volume, low DNA concentration), any control material included or recommended with the test, and follow-up actions to be taken when a test fails.

- Describe any additional procedures, methods, and practices incorporated into the directions for use, including confirmatory testing that should be conducted. Indicate whether parental or familial testing is a required part of the test.

The following information on test design should be provided:

- Specify the components of the test, including the sequencing platform and associated technology (e.g., long reads) and ancillary reagents, instrumentation, and equipment.

- Describe all steps of the test design, development, and validation (e.g., DNA extraction, library preparation, variant calling) and the procedures and components associated with each step.

- Provide details about the specimen type (e.g., saliva, whole blood), matrix (e.g., preservatives, anticoagulants) and minimum and maximum volume appropriate for testing. Specify specimen collection, pre-processing (e.g., nucleic extraction methods), storage and any additional pre-analytical specimen preparation steps, as applicable.

- Indicate the minimum yield and quality of DNA appropriate to obtain test accuracy.

- Indicate methods for processing DNA for sequencing (e.g., amplification, capture) and ways to assess the yield and quality of the final processed material.

- Indicate the level of multiplexing, if applicable.

- Specify all software components, whether developed in-house or obtained from a third party. Indicate the name and version and provide descriptions of all software components, including for sequencing instruments and post-sequencing data analysis and processing (i.e., bioinformatics pipeline). Indicate whether software is run locally or on a remote service (e.g., cloud-based), and record any modifications made to open-source software.

- Indicate databases and versions used for data analysis and describe how new versions of existing database(s) or a new database will be incorporated into the test and validated. Indicate whether sequence is aligned against the full human reference assembly or the targeted sequences, and document accession and version numbers for the full human reference assembly used for alignment.

- Describe criteria used for annotation and filtering of variants.

H. Test Reports

Include the following information in test reports consistent with 21 CFR 809.10 compliant labeling (as applicable):

- The relationship between reported variants and the clinical presentation of the patient.

- A description of genomic and chromosomal regions detected by the test. For panels, all targeted genes should be indicated.

- A summary of the results of performance studies performed in accordance with Section VI.D.

- A prominently-placed list of pathogenic or actionable variants on the first page of a test report. If variants of unknown significance will be reported, clearly separate these from pathogenic or actionable variants in the test report, and include a statement that their clinical relevance is not known. Indicate which classes of variants (e.g., benign polymorphisms) are not included in the test report. Also include the following information:

    - Report variants using a widely accepted nomenclature.

    - Provide a description of the clinical evidence supporting the interpretation reported variants.

    - Provide a summary of genes related to patient's phenotype, and any databases relied upon for variant interpretation, if relevant.

    - Indicate whether additional information, such as test results from family members, is needed to definitively interpret the variant.

- Indicate test limitations, including interrogated regions that failed sequencing, any interfering substances, and limitations to variant interpretation.

- Specify risk mitigation elements, including rationale for and description of any additional procedures, methods, and practices incorporated into the directions for use or recommended as a follow-up that mitigate risks associated with the test.

- Throughout the report, use clear, consistent language that can be easily understood.

VII. Modifications

Modifications to an NGS-based test can vary greatly in type, scope and impact. They may range from new reagent supplier and software updates to new platforms, changes in chemistry, or the addition of new sequencing targets. While these changes necessitate analytical validation, the types of studies that need to be performed will depend on the type and the extent of the

modification. At present, under FDA regulations, a modification to a cleared or approved test may require a new submission to FDA.

In order to remain within the scope of this guidance, modifications to targeted and WES NGS-based tests should stay within the intended use of aiding in the diagnosis of individuals with suspected germline diseases or other conditions.

Always re-evaluate test performance when modifications to the test are made. When making modifications, FDA recommends the following:

- Document all modifications to a test, including the protocol. This should include software updates and other modifications to the bioinformatics pipeline.

- Prepare a detailed SOP for revalidation after anticipated test modifications, including those to software. In this protocol, indicate anticipated modifications and the procedures that will be followed to implement them, including the types of validation studies that will be performed, and the performance metrics and thresholds that must be achieved introducing the modification.

- Conduct revalidation using a sufficient number of well-characterized samples to provide assurance of stated test performance. Sample numbers and types should be documented and justification provided for sample numbers and types selected.

- Document the types of validation studies that will be conducted after a modification and document the test's post-modification performance.

- Where appropriate, revalidate the test end-to-end, not simply the modification, and document performance. If available, existing well-characterized data files of sequences representative of the test's indications for use, containing known variants, may be used when modifications are made solely to the bioinformatics pipeline. Minor modifications to the pipeline can be validated by comparing results from the new pipeline to the existing test pipeline. Always document performance.

- If multiple modifications are made to a test over time, assess each modification separately as well as in aggregate, and document performance.

- When adding new genes to an existing panel, evaluate test performance for the original genes on the panel and document performance. If the changed test does not meet performance requirements, redesign may be necessary. Unmasking of genes in a panel for reporting is not considered a modification if performance for those genes was already demonstrated as part of the original test validation.

- Include a procedure to account for updates to internal and external databases and their potential impact on the clinical interpretation of variants. Document any updates including name, location, and new version of the database.

VIII. Additional Resources

- FDA guidance document entitled "Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classifications."

- FDA guidance document entitled "Requests for Feedback on Medical Device Submissions: The Pre-Submission Program and Meetings with Food and Drug Administration Staff."

- FDA guidance document entitled "Procedures for Class II Device Exemptions from Premarket Notification."

- Gargis A.S. et al, "Assuring the Quality of Next-Generation Sequencing in Clinical Laboratory Practice," Nat Biotechnol. 2012 30(11):1033-6.

- "Molecular Pathology Checklist," College of American Pathologists (April 21, 2014).

- Rehm H.L. et al., "ACMG Clinical Laboratory Standards for Next-Generation Sequencing," Genet Med. 2013 15(9):733-47.

- Schrijver I. et al., "Methods-Based Proficiency Testing In Molecular Genetic Pathology," The Journal of Molecular Diagnostics (2014), 16(3):283-7.

- Analytical Performance Specifications for Comprehensive Genomic Profiling (M00118, V1).

- "CLSI MM09-A2, Nucleic Acid Sequencing Methods in Diagnostic Laboratory Medicine; Approved Guideline – Second Edition," Clinical and Laboratory Standards Institute (February 2014).

- Aziz N. et al., "College of American Pathologists' Laboratory Standards for Next-Generation Sequencing Clinical Tests," Arch Pathol Lab Med (2015),139:481-93.

- "Next Generation Sequencing (NGS) Guidelines for Somatic Genetic Variant Detection," New York State Department of Health (March 2015).

- "Guidelines for Validation Submissions of Next Generation Sequencing (NGS) Assays under the NYS Testing Category of Genetic Testing – Molecular," New York State Department of Health (July 2015).

- Matthijs G. et al, "Guidelines for Diagnostic Next-Generation Sequencing," European Journal of Human Genetics (2016) 24, 2–5.