Contains Nonbinding Recommendations

Considerations for Design, Development, and Analytical Validation of Next Generation Sequencing (NGS) – Based In Vitro Diagnostics (IVDs) Intended to Aid in the Diagnosis of Suspected Germline Diseases

Guidance for Stakeholders and Food and Drug Administration Staff

Document issued on April 13, 2018.

The draft of this document was issued on July 8, 2016.

For questions about this document concerning devices regulated by CDRH, contact Zivana Tezak at 301-796-6206 or Adam Berger at 240-402-1592 or by email at OIRPMGroup@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
Center for Biologics Evaluation and Research

GRZAN

Preface

Public Comment

You may submit electronic comments and suggestions at any time for Agency consideration to https://www.regulations.gov. Submit written comments to the Dockets Management Staff, Food and Drug Administration, 5630 Fishers Lane, Room 1061, (HFA-305), Rockville, MD 20852. Identify all comments with the docket number FDA-2016-D-1233. Comments may not be acted upon by the Agency until the document is next revised or updated.

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.

GRZAN

Table of Contents

I. Introduction 1
II. Background 3
III. Scope 4
IV. Regulatory Pathway for NGS-Based Tests Intended to Aid in the Diagnosis of Suspected Germline Diseases 5
V. Elements of an NGS-Based Test Intended to Aid in the Diagnosis of Suspected Germline Diseases 6
VI. Recommendations for Design, Development, and Validation of NGS-Based Tests Intended to Aid in the Diagnosis of Suspected Germline Diseases 7
A. Test Design Considerations 8
1. Indications for Use Statement(s) of the Test 8
2. Specific User Needs for the Test 8
3. Specimen Type 9
4. Interrogated Regions of the Genome 9
5. Performance Needs 10
6. Test Elements and Methods 10
B. Test Performance Characteristics 13
1. Accuracy 13
2. Precision (Reproducibility and Repeatability) 16
3. Limit of Detection (LoD) 17
4. Analytical Specificity 17
C. Test Run Quality Metrics 18
1. Coverage (Read Depth and Completeness) 18
2. Test Run Metrics and Performance Thresholds 19
D. General Recommendations for Performance Evaluation Studies 22
E. Supplemental Procedures 25
F. Variant Annotation and Filtering 25
G. Presentation of Test Performance in Labeling 26
H. Test Reports 28
VII. Modifications 29
VIII. Recommendations for Reviewing Changes to Device Design and Production 31
IX. Additional Resources 33
X. Appendix A 35

Considerations for Design, Development, and Analytical Validation of Next Generation Sequencing (NGS) – Based In Vitro Diagnostics (IVDs) Intended to Aid in the Diagnosis of Suspected Germline Diseases

---

Guidance for Stakeholders and Food and Drug Administration Staff

This guidance represents 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

The field of genomic testing is dynamic, building off of increasing amounts of data and a rapidly evolving technology base. While current regulatory approaches are appropriate for conventional diagnostics that measure a limited number of analytes associated with a disease or condition, the new sequencing technologies used in genomic testing can examine millions of DNA variants at a time, and thus warrant a flexible approach to oversight that is adapted to the novel and evolving nature of these tests. The Agency's intent is to optimize its approach to regulatory oversight for Next Generation Sequencing (NGS) *in vitro* diagnostic (IVD) tests to support the needs of the rapidly evolving novel technologies in genomic medicine while at the same time meeting its critical mandate to ensure that medical devices have a reasonable assurance of safety and effectiveness.

FDA's vision is that NGS-based tests can be developed, validated, and offered for clinical use through a process that leverages appropriate standards, quality systems controls and community assessment of clinical validity to streamline the premarket review process. This guidance document describes one part of FDA's efforts to create a flexible and adaptive regulatory approach to the oversight of next generation sequencing (NGS)-based tests.

As a step toward this vision, FDA is outlining key considerations for designing, developing, and establishing analytical validity of NGS-based tests used for whole exome human DNA sequencing (WES) or targeted human DNA sequencing intended to aid in the diagnosis of symptomatic individuals with suspected germline diseases or other conditions (hereinafter referred to as "NGS-based tests intended to aid in the diagnosis of suspected germline diseases" or "NGS-based tests"). These considerations are based on FDA's overall review experience with NGS-based tests to date, the agency's understanding of the design and analytical validation considerations appropriate for such tests, and extensive interactions with the community.

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. Examples of signs and symptoms of suspected germline diseases could include developmental delay, congenital dysmorphologies, or other clinical features. This guidance does not address tests intended for use in the sequencing of healthy individuals.

FDA intends the recommendations in this guidance to both assist test developers directly, and also to inform the development of consensus standards by experts in the community. FDA has a successful and widely used standards recognition program that facilitates the use of consensus standards to meet premarket submission requirements for devices. Ultimately, if appropriate consensus standards for NGS-based tests intended to aid in the diagnosis of suspected germline diseases are developed by the community and recognized by FDA, test developers could certify conformity to such standards in a premarket submission, which could significantly streamline the review process. FDA is particularly interested in the development of standards that address considerations beyond what is described in this guidance, as standards can be updated regularly by the community of experts as technologies and their applications advance.

Further, FDA believes the recommendations in this guidance and/or standards that address these recommendations may potentially form the basis of special controls that could reasonably assure the safety and effectiveness (which, for IVDs, generally means a reasonable assurance of analytical and clinical validity) of these tests, allowing NGS-based tests intended to aid in the diagnosis of suspected germline diseases to be candidates for classification as Class II devices via the De Novo process. If classified in class II, subsequent NGS-based tests intended to aid in the diagnosis of suspected germline diseases would be reviewed through the 510(k) program, potentially by accredited Third Party organizations.[1] FDA hopes that conformance to robust FDA-recognized standards that include performance

[1] Under section 523 of the FD&C Act, 510(k)s for certain devices may be reviewed by FDA-accredited third parties, who make a recommendation to FDA regarding substantial equivalence.

criteria will ultimately provide sufficient assurance of analytical validity such that, as long as there is also sufficient assurance of clinical validity (e.g. through an FDA-recognized database containing adequate, well-vetted, and publicly accessible valid scientific evidence) and appropriate labeling (see applicable parts of 21 CFR part 801 and 21 CFR 809.10 and Sections VI.G and VI.H below), FDA can consider exempting these tests from premarket review altogether in the future.

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

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.

While most IVDs 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 specific 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.[2] 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."[3] FDA

[2] Public Workshop - Optimizing FDA's Regulatory Oversight of Next Generation Sequencing Diagnostic Tests Public Workshop, Feb. 20, 2015, http://wayback.archive-it.org/7993/20170111165845/http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm427296.htm.

[3] Public Workshop - Standards Based Approach to Analytical Performance Evaluation of Next Generation Sequencing In Vitro Diagnostic Tests, Nov. 12, 2015, http://wayback.archive-it.org/7993/20170111165845/http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm459449.htm.

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"[4] and another workshop on March 2, 2016, entitled "Patient and Medical Professional Perspectives on the Return of Genetic Test Results."[5] Much of the public feedback obtained at the 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 intended to aid in the diagnosis of suspected germline diseases that it believes could be used to help provide a reasonable assurance of the safety and effectiveness of these tests. FDA hopes this guidance document will spur development of such standards.

III. Scope

This guidance document provides recommendations for designing, developing, and validating NGS-based tests intended to aid clinicians in the diagnosis of symptomatic individuals with suspected germline diseases.

This document does not apply to NGS-based tests intended for:

- aid in the diagnosis of microbial infection,
- cell-free DNA testing,
- direct-to-consumer uses,
- fetal testing,
- microbial genome identification and detection of antimicrobial resistance and virulence markers,
- pre-implantation embryo testing,
- risk assessment,
- risk prediction,
- RNA sequencing,
- screening,
- stand-alone diagnostic purposes,
- tumor genome sequencing,
- use as a companion or complementary diagnostic.

[4] Public Workshop - Use of Databases for Establishing the Clinical Relevance of Human Genetic Variants, Nov. 13, 2015, http://wayback.archive-it.org/7993/20170111165844/http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm459450.htm.

[5] Public Workshop - Patient and Medical Professional Perspectives on the Return of Genetic Test Results, Mar. 2, 2016, http://wayback.archive-it.org/7993/20171115050724/https://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm478841.htm.

Tests for these uses may have other performance characteristics and/or risk considerations not addressed by the recommendations presented here, although the principles described in this guidance may be applicable to tests for some of these uses. Sponsors are welcome to discuss opportunities to apply the principles of this guidance to their tests.

This guidance document also discusses FDA's vision for future regulatory oversight of NGS-based tests intended to aid clinicians in the diagnosis of symptomatic individuals with suspected germline diseases

IV. Regulatory Pathway for NGS-Based Tests Intended to Aid in the Diagnosis of Suspected Germline Diseases

To date, FDA has cleared a small number of single-gene, disease-specific, targeted, NGS-based tests (e.g., Illumina MiSeqDx Cystic Fibrosis 139-Variant Assay (k124006) and Illumina MiSeqDx Cystic Fibrosis Clinical Sequencing Assay (k132750)). However, FDA has not classified NGS-based tests with a more general intended use to aid in the diagnosis of suspected germline diseases, and there are currently 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. 360(k)). As a result, such a test is automatically classified into class III by operation of law, and these tests are currently subject to Premarket Approval application (PMA) requirements.[6]

Where there is not an existing predicate device, low-to-moderate risk devices can also come to market through the De Novo classification process.[7] FDA believes there is a reasonable possibility that the risks associated with NGS-based tests intended to aid in the diagnosis of suspected germline disease may be sufficiently mitigated by a combination of general and special controls. Accordingly, such tests may be suitable candidates for the De Novo classification process. An applicant who submits a De Novo request for classification 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 their test. If FDA were to grant a De Novo request for an NGS-based test intended to aid in the diagnosis of suspected germline disease and classify the category of tests in class II, the test would be authorized for marketing and could serve as a predicate for future 510(k) submissions for other NGS-based tests intended to aid in the diagnosis of suspected germline disease. All such tests would be subject to both general and any special controls that might be included in the new classification regulation, such as requirements regarding specific performance data and/or content of the labeling. FDA encourages applicants to engage with the Agency using the Pre-Submission process to

[6] *See* sections 513(f)(1) and 515(a)(2) of the FDA&C Act (21 U.S.C. 360c(f)(1) and 360e(a)(2)).
[7] *See* section 513(f)(2) of the FD&C Act (21 U.S.C. 360c(f)(2)). Further information about the De Novo process can be found on FDA's website (https://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHTransparency/ucm232269.htm).

discuss any anticipated De Novo requests. Information related to FDA's Pre-Submission process is available in FDA's guidance entitled "Requests for Feedback on Medical Device Submissions: The Pre-Submission Program and Meetings with Food and Drug Administration Staff" (https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm311176.pdf).

Further, under section 510(m) of the FD&C Act, 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 safety and effectiveness of the device. As the Agency gains more experience with these devices, FDA believes that it may be possible, in the future, to develop special controls that could provide a reasonable assurance of the safety and effectiveness of NGS-based tests intended to aid in the diagnosis of suspected germline disease, possibly under certain conditions of exemption, without the need for 510(k) premarket review.

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 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 as well as 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)). FDA believes that similar public availability and access to information regarding NGS-based tests, regardless of whether they are FDA reviewed or exempt from 510(k) premarket review, 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 Intended to Aid in the Diagnosis of Suspected 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 evaluation and assertion, 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. Generally, for each test element, test developers should establish metrics and acceptance criteria specific to their particular test.

For the purposes of this guidance, the term "variant evaluation" refers to the process by which evidence is assessed regarding a linkage between a genetic variant and a disease or condition and an assertion is made about that linkage (or lack thereof).

Interpretation of the clinical significance of an identified variant, performed by healthcare providers and laboratory professionals for the sole purpose of diagnosing or treating a specific individual patient, is not considered part of the test, but certain standard operating procedures (SOPs), including but not limited to protocols for variant evaluation, and some software products may be considered test elements. FDA recommends that applicants discuss their particular tests through the Pre-Submission process as early as possible in the development of the test.

VI. Recommendations for Design, Development, and Validation of NGS-Based Tests Intended to Aid in the Diagnosis of Suspected Germline Diseases

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 study type in the form of a minimum and target value that are appropriate and meet the performance needs for the indications for use of the test. After design and development of the test, validation studies should indicate if the predefined performance is met. If the test does not meet any of the predefined performance thresholds, 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.

FDA encourages development of a standard(s) that includes the design, development, and validation activities outlined in this section. For a standard to be recognized by FDA,[8] 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, as appropriate. FDA expects that demonstration of conformity with such standards may be used by developers of NGS-based tests intended to aid in the diagnosis of suspected germline diseases in premarket submissions, and possibly in the future in lieu of premarket review. However, the adequacy

[8] Guidance for Industry and FDA Staff – Recognition and Use of Consensus Standards (https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm077274.htm).

of a declaration of conformity with FDA-recognized standards for analytical validity may depend on the specific indications for use and the type of premarket review, or, potentially, exemption.

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 indications for 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 consistently meets performance metrics appropriate for the indications for use of the test. Each performance metric should be accompanied by either predetermined acceptance intervals or thresholds (as applicable). During the test design phase, developers should establish and justify minimum acceptable and target values for each performance metric that are appropriate and meet the performance needs 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

While this guidance applies to all NGS-based tests intended to aid in the diagnosis of suspected germline diseases, each test developer should define the specific indications for use for a given 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 but are not limited to aiding in the diagnosis of: children with signs and symptoms of developmental delay or intellectual disability, symptomatic patients with undiagnosed diseases, patients undergoing differential diagnosis, etc.

Examples of considerations for target populations would include but are not limited to: 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 users' 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. Test developers should take into consideration the spectrum of test users when designing the test.

An example of a specific user need for the test is a user's need 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 design questions such as the type of collection device required, minimum volume or quantity of sample, or 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 but are not limited to: whole 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.

The types of genes sequenced and/or reported by the test will depend on the specific indications for use, which in turn will influence aspects of test design and definition of test performance. Provide and justify, using an evidence-based assessment, the genes that should be included for a given indication. Known clinically relevant variants for a given indication that fall outside of the region detected by the test should also be specified. If critical interrogated regions do not meet minimum performance specifications (e.g., minimum coverage thresholds), the test should be modified and revalidated to reach minimum applicable thresholds.

Example: A test intended to aid in the diagnosis of suspected genetic disorders in newborns may use WES rather than a targeted 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, precision, limit of detection (LoD)) 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, or 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, and non-reported regions should be documented. Similarly, if part of an interrogated genomic region is difficult to sequence and cannot reach performance thresholds, 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. Test Elements and Methods

a. Test Element Specification

Specify and document all test elements (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 elements based on identified user needs, the indications for use statement, and predefined performance specifications (e.g., accuracy thresholds). Document the limitations of each test element for critical factors (e.g., coverage, multiplexing).

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

Listed below are recommendations for selected elements 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, but are not limited to, controls per sample or per run, or other controls 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 elements needed to compile, install, and run the bioinformatics pipeline.*
- *Specify and document whether software will be run locally or remotely (e.g., cloud-based).*
- *Specify and document which databases (including versions as applicable) 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 by the test, and considering any limitations of the pipeline in variant calling and evaluation. 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 selected elements 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 the minimum DNA volume, quantity, and quality to meet predetermined test performance specifications.*
- *Establish and document whether the test can be run when the sample is extracted DNA from outside sources, and establish and document the acceptance criteria and other 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 reported regions.*
- *Specify and document the composition of molecular barcodes and the procedures for their use, including any required procedures for avoiding barcode collision, mis-identification or mis-sorting.*
- *Specify and document methods for normalizing each input sample included in multiplex pools*

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 a re-run, fill in of certain regions that failed to meet appropriate quality metrics or Sanger confirmation of test results, for example.*

Note that the above list is not comprehensive, and that all steps of the test should be documented, including a justification for specifications.

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.

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 optimize individual steps of an NGS-based test and to verify that test elements are operating as expected and meeting predefined performance thresholds appropriate for the indications for use of the test. 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, including documentation that performance meets predefined thresholds) prior to initiating clinical use of the test. This section recommends a set of performance metrics that should be assessed when analytically validating NGS-based tests intended to aid in the diagnosis of suspected germline diseases.

1. Accuracy

Demonstrate accuracy by measuring positive percent agreement (PPA), negative percent agreement (NPA), and technical positive predictive value (TPPV) Set thresholds for PPA, NPA, and TPPV that assure that the test will meet its predefined performance specifications. Based on different scenarios or the methodology used, additional metrics for evaluation of accuracy may be developed.

Accuracy involves determining the closeness of agreement between a measured value and a true value of a measurand.[9] 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 high confidence sequence of the reference.[10]

[9] Adapted from JCGM 200:2012 (Joint Committee for Guides in Metrology).
[10] See Gargis AS et al., for Next Generation Sequencing: Standardization of Clinical Testing (Nex-StoCT) Workgroup. Assuring the quality of next-generation sequencing in clinical laboratory practice. Nat Biotechnol. 2012 30(11):1033-6.

Minimum acceptable overall and target thresholds for PPA, NPA, and TPPV point estimate and lower bound of the 95% confidence interval (CI) of an NGS-based test should be predefined and reported (see applicable parts of 21 CFR part 801 and 21 CFR 809.10, and Sections VI.G and VI.H below) for each variant type claimed by the test. These thresholds will depend on the indications for use of the test, and variables such as the types of variants detected and whether variants are confirmed using an orthogonal assay. Thresholds should be justified using objective evidence and valid statistical techniques, which should also be reported.

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. In the context of this guidance, a test result is "positive" when it returns a variant, which is a difference in sequence at the same position in a reference sequence (e.g., a current build of the human genome).

b. Negative Percent Agreement

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

NPA is generally defined as the proportion of correct calls by the assay for the absence of an analyte[11] (in this case, the absence of a genetic variant) and reflects the frequency of FPs. More specifically, for a genetic test, NPA is the ability to correctly identify wt bases (i.e., the probability that the assay will not detect a sequence variation when none are present).

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.

TPPV relates to the likelihood that a variant call is a true positive, and reflects the number of false positives per test. This number is particularly important if variant confirmation will not be performed.

[11] See CLSI Harmonized Terminology database at http://htd.clsi.org/ - "Negative percent agreement (NPA) - the percentage of agreement of the test method's ability to obtain negative results in concordance with negative results obtained by the comparative method" (CLSI MM17-A, Verification and Validation of Multiplex Nucleic Acid Assays; Approved Guideline).

d. Calculating Accuracy

Accuracy should be calculated for each variant type, in the context in which the test can detect them, as well as for clinically relevant variants. For variant types, it may be helpful to use the results of studies conducted with well-characterized reference materials or agreed-upon samples with high confidence calls. For clinically relevant variants, the accuracy calculation should use the results of studies conducted with clinical samples pertinent to the test's indications. (See Section VI.D: General Recommendations for Performance Evaluation Studies for further details on evaluating NGS-based test accuracy.)

Accuracy can be calculated by constructing the following type of table:

		Comparator Method		Total
		Positive	Negative	
Test	Positive	A	B	A+B
	Negative	C	D	C+D
	No calls or invalid calls	E	F	E+F
	Total	A+C+E	B+D+F	N

N is the total number of samples in the accuracy study (N=A+B+C+D+E+F). No calls or invalid calls describe the result where a base call is not made, which can be a result of a number of factors such as, but not limited to, the base level performance not meeting predefined thresholds for quality, resulting in insufficient data to make a variant call. Some tests also report equivocal results, which are non-missing, non-erroneous results that are valid from a QC perspective but are neither positive nor negative; i.e., a result that does not provide definitive binary information regarding whether a mutation is present or absent. Note that a no call or invalid result is different from an equivocal call.

The percentage of no calls or invalid calls in the accuracy study should be estimated as (E+F)/N along with a 95% two-sided confidence interval.[12] In addition, the percentage of no calls or invalid calls for positive results from the comparator method should be estimated as E/(A+C+E) and for negative results from the comparator method as F/(B+D+F).

No calls and invalid calls should not be used in PPA, NPA, or TPPV calculations, but should be individually documented as part of the accuracy study results. Minimum acceptable values for the number of 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.

Removal of no calls or invalid calls yields accuracy study data that includes only valid test results, as presented by the following table:

[12] For calculation of 95% confidence intervals, follow CLSI EP12-A2 User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline - Second Edition.

		Comparator Method		Total
		Positive	Negative	
Test	Positive	A	B	A+B
	Negative	C	D	C+D
	Total	A+C	B+D	A+B+C+D

- PPA is calculated by dividing A, the number of known variants detected by the test (true positive test results) by (A+C), the number of known variants tested (true positive plus false negative test results). These calculations should not include no calls or invalid calls, which should be presented separately. PPA yields the false negative (FN) rate (FN = 1 – PPA). Calculate PPA at the variant and specimen level, for each variant type, as well as for clinically relevant variants, as applicable.

- NPA can be calculated by dividing D, the number of true negative results, by (D+B), the number of wt results for variants tested (true negative plus false positive test results). NPA reflects the frequency of false positives (FP) (FP = 1 – NPA). Calculate NPA at the specimen level. Ideally, NPA should be calculated for each variant type, as well as for clinically relevant variants, as applicable. It can be more generally calculated as the number of wt bases correctly identified as wt, or alternatively, expressed as the number of false variants in a pre-defined interval.

- TPPV is calculated by dividing A, the number of true positive test results by (A+B), the number of all positive results detected by the test (true positive plus false positive test results). Calculate TPPV at the interrogated region (variant) and specimen level. TPPV should be calculated for each variant type, as well as for clinically relevant variants, as applicable.

For details, see Section X: Appendix A.

2. Precision (Reproducibility and Repeatability)

Evaluate precision (reproducibility and repeatability) for both variant and wt calls, with each metric separately reported for each condition, interrogated region, and variant type. Conduct testing that accounts for major factors that may contribute to test variability, including but not limited to testing multiple samples, runs, reagent lots, days, and operators. Conduct testing that accounts for other sources of variability as applicable, including but not limited to testing multiple instruments, multiple testing sites, lane replicates, and lanes.

Reproducibility for NGS-based tests involves measuring test variability under a variety of specified conditions (such as when using different operators, different operating conditions, different days of measurement, different instruments, etc.) using the same sample (including samples around the test cutoff), 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, 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 result; rather, test developers should calculate the

percentage of replicates with the same results for each sample. The percentage of no calls or invalid replicates should be also reported.

Thresholds for reproducibility and repeatability of NGS-based test should be predefined and reported for each condition tested and genomic context, separately for each variant type. These thresholds will depend on a number of variables, including the indications for use and the types of variants that will be detected. Thresholds should be justified using objective evidence and valid statistical techniques.

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 calls or no call 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 calls or no calls is obtained among the replicates tested for that concentration. When different variant types may have different LoDs, calculate the LoD for each variant type, in different sequence contexts. Similarly, an upper LoD 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.

Analytical specificity relates to the ability of a test to measure solely the intended analyte. Interference in measurement from endogenous or exogenous substances that, based on the indications for use and test design, may be expected to result in failure to detect an analyte, and could yield 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.

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 the specimen or sample types and the DNA extraction method 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.

Note that for tests such as WES, where many queried regions will be identical across any two specimens, cross-contamination evaluation may require the assessment of known differences between samples.

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, sequencing instrumentation, bioinformatics pipeline related metrics). Set sample and variant level quality control acceptance criteria (e.g., number of reads mapped, coverage, strandedness).

Test run quality metrics are used to determine whether an individual test run, sample or variant call should be accepted, or, when applicable, whether supplemental procedures should be used to further query a variant call. In particular, test run quality metrics should be documented and reported for the clinically relevant parts of the genome (i.e., genes and regions for which there is adequate evidence for causal associations between pathogenic variants and diseases).

A number of test run quality metrics associated with the whole test or specific steps of an NGS-based test may be used, and will, among other factors, depend on the technology used. These metrics are described below.

Variants should not be reported from regions of a test run that fail to meet quality thresholds. Regions that are uninterrogated due to failure to meet quality thresholds should be reported as such. Also see labeling recommendations under Section VI.G. and VI.H.

1. Coverage (Read Depth and Completeness)

Establish, justify, 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. [13]

Establish and document the percentage of the claimed region that is covered / completeness (i.e., percentage of test with sufficient coverage above minimum thresholds).

Average and minimum depth of coverage thresholds should be predefined and reported based on the indications for use of the test and justified using objective evidence and valid statistical techniques, which should also be reported. In general, reducing average coverage increases the percentage of the test with insufficient minimum coverage. If critical interrogated regions do not meet minimum coverage thresholds, 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 instances when interrogated regions do not meet predefined coverage thresholds.

Selection of thresholds should demonstrate adequate test performance for the indications for use statement and predefined user needs. Average and minimum coverage, percentage of completeness, 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), performance (e.g., base-call error rates, number of independent reads), and other information, as relevant.

Tests should meet predefined thresholds across reported variants for expected heterozygous and homozygous germline allele fractions. If a test is reporting variants with allele frequencies outside of the expected heterozygous and homozygous germline ranges (e.g., germline mosaicism), test developers should ensure that the necessary minimum coverage needed to detect those variant allele frequencies is achieved. Supplemental procedures may be necessary to adequately interrogate these variants.

2. Test Run Metrics and Performance Thresholds

FDA recommends establishing test run metrics and performance thresholds for all critical NGS-based test steps. These thresholds will depend on a number of variables, and should be justified using objective evidence and valid statistical techniques. These metrics and their performance thresholds should be 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 and re-validated.

[13] Jennings et al. (2017) define depth of coverage as "…the number of aligned reads that contain a given nucleotide position…". Gargis et al. (2012) define average depth of coverage as "…the average number of overlapping reads within the total sequenced area", and uniformity of coverage as "the distribution of coverage within specific targeted regions in which variant calling will occur".

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, document, and justify thresholds for genomic DNA concentration, volume, and quality.*
- *As applicable, 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, document, and justify a threshold for base quality score (e.g., 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.*

If base quality score is not used, document the applicable method used and justify why it is an appropriate method.

Other metrics of sequence generation may be used, as 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, documented, and justified 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 percentage 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 (variant–level metrics) and sample-level, overall variant calling summary metrics. Appropriate metrics may depend on the bioinformatics pipeline used for variant calling. Thresholds for reporting should be established, justified and documented.

Examples of appropriate metrics include:

- Variant call quality score.
- Number and percentage of reads with the variant reported.

- Allelic read percentages, including percentage 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 percentage of variant reads defined for homozygous and heterozygous calls).

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

- Strand bias.

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 test performance, incorporate the features listed below, as applicable. Provide a detailed justification for deviations from 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 test performance on variants in highly homologous, highly polymorphic, or other difficult regions if these regions are part of the indications for use of the test.

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

- Use specimens, as appropriate, that reflect the actual specimen types to be tested (e.g., whole blood, saliva) and the population the test is indicated for.

- Determine the number and type(s) of samples required to demonstrate that performance thresholds have been met with confidence for relevant metrics and the indications for use. This number will depend on the indications for use of the test, the number and types of variants (if relevant) claimed to be detected and reported by the test, and the critical performance parameters that must be met to support that use.

- Evaluate test performance for different genotypes (i.e., wt, heterozygous, compound heterozygous, homozygous) consistent with the test's indications for use. Additionally, 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.

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

- Include the finalized bioinformatics pipeline for data processing and analysis as a part of the overall end-to-end test validation.

- 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, evaluate and document the impact of barcoding on test performance, and establish and document that the combinations of patients/barcodes in a run provide accurate and reproducible results 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 comparator by FDA, such as bidirectional sequencing or another well-validated method. If available and as appropriate, evaluate accuracy by comparing the sequence generated by your test to a well-characterized "gold standard" reference sequence or a consensus sequence of agreed-upon well-characterized samples.

- To determine the types of variants the test can detect and the contexts in which the test can detect them (regardless of whether the variants are pathogenic), determine accuracy using well-characterized reference materials or agreed-upon samples with high confidence calls (baseline accuracy). Accuracy studies should include representative variants and variant types according to the indications of the test. Studies should account for clinically meaningful regions relevant for the test's indications of use, variants across different genomic contexts, difficult-to-sequence or difficult-to-map regions of the genome, mosaicism, and any other variant types, regions, or contexts included in the indications for use of the test. For example, if the test is intended for detecting and reporting indels, include a distribution of variants in appropriately-sized increments (e.g., no more than five base pairs), presented separately for insertions and deletions.

- Clinical samples containing clinically relevant variants pertinent for the test's indications for use should be used in accuracy studies and should be selected to ensure adequate representation of the clinically relevant sequence variations and genomic contexts claimed to be detected and reported by the test. Evaluate the accuracy of the test using fresh clinical samples, collected and processed in a manner consistent with the test's instructions for use. Fresh clinical samples containing specific clinically relevant variants should be used whenever possible. If fresh clinical

samples are not available, it may be acceptable to substitute or supplement studies with archived clinical samples or human cell line samples that contain specific variants in clinically relevant regions. In some limited circumstances, use of contrived or biosynthetic samples may also be used as an acceptable alternative, if deemed appropriate by the FDA.

- The number and types of specimens will depend on a variety of test-specific factors, and should be determined and justified before conducting studies. Although there is no requirement to test a certain number of loci, or to test samples from a certain number of different patients (this may also depend on the number of variants of different types in a given validation sample), the number and types of samples used in the study should be statistically justified for the test's indications for use and details of the types of samples and description of how the study was performed should be documented and reported.

- When clinical samples, cell lines or biosynthetic materials are not available, *in silico* constructed sequences containing known sequence variants of various claimed types (e.g., SNVs, indels, duplications, repeat expansions, CNVs, structural variants) may be used in addition to biological specimens to assess performance of the bioinformatics pipeline. Those data files should, however, be generated using the same pre-analytical and analytical methodology that is used by the test. *In silico* sequences can be constructed in several different ways, which should be clearly described, justified, and documented. Further information about software validation can be found in the guidance document entitled "General Principles of Software Validation" (https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm085281.htm).

- Calculate PPA, NPA and TPPV separately for each type of variant claimed (e.g., SNVs, indels, structural variants) and sequence context (e.g., highly homologous, highly polymorphic, or other difficult regions) to be assessed by the test. See Section VI.B for more information on calculating accuracy.

- If deemed appropriate by FDA, it may be possible to use certain validation study results across multiple, similar NGS-based tests on the same platform that use, for example, the same preanalytical steps, library preparation method and sequencing method.

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 predefined thresholds. Alternatively, adequate justification for reporting those variants with lower accuracy can be provided based on other means.

F. Variant Annotation and Filtering

Select and justify 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), genome Aggregation Database (gnomAD) 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 (when the performance in different contexts has been well validated and demonstrated to achieve predefined performance characteristics) 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 in Labeling

Test labeling must comply with 21 CFR 809.10 and applicable sections of 21 CFR part 801. The following should be included in the labeling when providing information on test performance:

- Make available and provide unrestricted public access to view the test's indications for use statement, limitations, and summary performance information via a prominent hyperlink on the test developer's 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, duplications, repeat expansions) claimed to be detected and reported by the test.
    - Any limitations of the test (e.g., interrogated targets such as genes, types of sequence variants, allele frequencies, or genomic contexts that the test cannot detect with validated performance, failure to detect certain rearrangements, insertions and deletions larger than a certain size, or in specific genes relevant for test's indications for use).
    - If possible, 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.

- Report variants 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 metrics (e.g., 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 size, separately for deletions and insertions, by polymorphic and non-polymorphic regions), summarized as a mean percentage 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 (e.g., number of runs, days, instruments, reagent lots, sites, operators, specimens/type, etc.).

    - Indicate the average depth of coverage and the percentage of target regions 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 elements 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 elements of the test 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). Record any modifications made to open-source software. Further information about software validation can be found in the guidance document entitled "General Principles of Software Validation" (https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm085281.htm).

- 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 (signs and symptoms), as applicable, and the source of the information (e.g. including whether this information is supported by information in an FDA-recognized database of human genetic variants[14]).

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

[14] See FDA guidance, Use of Public Human Genetic Variant Databases to Support Clinical Validity for Genetic and Genomic-Based In Vitro Diagnostics (https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM509837.pdf).

- 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. As applicable, reported variants can encompass known or predicted effect (including pathogenicity, penetrance), and some may have uncertain significance. 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 variants) 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 evaluation of reported variants.

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

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

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

- 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 validation, the types of studies that need to be performed will depend on the type and the extent of the modification. Some modifications to a test that has received premarket authorization from FDA are subject to FDA premarket review.

As described in section IV of this guidance document, NGS-based tests with general intended uses such as aiding in the diagnosis of suspected germline diseases are currently automatically classified into class III by operation of law, and as such, are currently subject

to PMA requirements.[15] As further described in section IV of this guidance document, FDA believes there is a reasonable possibility that the risks associated with such NGS-based tests may be sufficiently mitigated by a combination of general and special controls, and accordingly, such tests may be suitable candidates for the De Novo classification process. If FDA were to grant a De Novo request for an NGS-based test intended to aid in the diagnosis of suspected germline disease and classify this type of test in class II, the test would be authorized for marketing and could serve as a predicate for future 510(k) submissions for other NGS-based tests intended to aid in the diagnosis of suspected germline disease. The remainder of this section discusses the regulatory standards that apply for modifications to an IVD subject to 510(k) ("510(k) device").

Under FDA regulations, a new 510(k) is required for a modified 510(k) device when the modification constitutes a "major change or modification in the intended use of the device" or when the modification "could significantly affect the safety or effectiveness of the device."[16] Modifications to an IVD that "could significantly affect the safety or effectiveness of the device" include technology, engineering, performance, and material changes to the IVD. The guidance "Deciding When to Submit a 510(k) for a Change to an Existing Device" (https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm514771.pdf) (the 510(k) modifications guidance) provides information to aid IVD manufacturers during the process of deciding whether a modification to a 510(k)-cleared device (or group of devices), or other device subject to 510(k) requirements, such as a device that was granted marketing authorization via the De Novo classification process, requires a new 510(k).

The discussion here is intended to be consistent with and supplement the 510(k) modifications guidance by highlighting certain aspects of the discussion at sections V.D3 and D4 of that guidance for NGS-based tests intended to aid in the diagnosis of suspected germline diseases. If these tests are classified in class II subject to 510(k), FDA would recommend that, before modifying their test, test developers perform an appropriate risk-based assessment to determine whether the modification would significantly change the analytical or clinical performance or safety specifications of their test. If a test developer's risk-based assessment determines that the modification is likely to significantly change analytical or clinical performance or safety specifications, the change to the test likely could significantly affect the safety or effectiveness of the test and submission of a new 510(k) is likely required. If the risk-based assessment determines that the modification is not likely to significantly change these attributes, test developers should perform appropriate analytical or clinical verification and validation activities to more fully characterize the scope and impact of the modification. Should the results of verification and validation using standard methods and performance criteria that have been established for evaluation of the specific device (e.g.,

[15] For general information regarding when to submit a PMA supplement for changes to a PMA-approved device, refer to 21 CFR 814.39 and FDA's guidance "Modifications to Devices Subject to Premarket Approval (PMA) – The PMA Supplement Decision-Making Process" (https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM089360.pdf).

[16] *See* 21 CFR 807.81(a)(3).

protocols and criteria used to support the original 510(k), or a protocol and criteria established in the original 510(k) that described how anticipated changes would be evaluated) indicate that (a) the performance of the modified test system is within the criteria, (b) the performance of the modified test system has not significantly changed relative to previously cleared claims, and (c) there are no other new risks or significant changes to existing risks, then it is likely that the change could not significantly affect safety or effectiveness, and a 510(k) is likely not required.

Some test modifications that could significantly impact performance or safety specifications may be anticipated at the time of an original 510(k) submission. Where practical, test developers may include an appropriate prospective change protocol with specific procedures and acceptance criteria in the original submission. If FDA clears such a change protocol that outlines the specific types of anticipated changes, the procedures that will be followed to implement them, and the acceptance criteria that will be met prior to implementation as part of the 510(k), future modifications that are made in accordance with the specific procedures and acceptance criteria outlined in the change protocol generally would not require a new 510(k). The test developer would be responsible for documenting the modification as necessary but would not be required to submit the documentation to FDA.[17] However, if modifications are made to the procedures or acceptance criteria in the change protocol cleared in the previous 510(k), a new premarket submission may be required.

If a modification expands the indications for use of an NGS-based test beyond aiding in the diagnosis of symptomatic individuals with suspected germline diseases or other conditions, the resulting test is outside the scope of this guidance.

VIII. Recommendations for Reviewing Changes to Device Design and Production

Regardless of whether a change requires premarket review by FDA, FDA's quality system (QS) regulation requires manufacturers of finished medical devices to review and approve changes to device design and production (21 CFR 820.30 and 820.70) and document changes and approvals in the device master record (21 CFR 820.181). Test performance should always be re-evaluated when modifications to the test are made. When reviewing, approving, and documenting such 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 SOP, indicate anticipated modifications and the procedures

[17] Whenever manufacturers change the design or production of their device, they must take certain actions to comply with the quality system (QS) regulation, 21 CFR part 820, unless the device is exempt by regulation from these requirements. The QS regulation requires, among other things, that manufacturers keep records, and these records must be made available to an FDA inspector (see 21 CFR part 820 subpart M; section 704(e) of the FD&C Act, 21 U.S.C. 374(e)).

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, as determined by the results of a risk-based assessment, 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 end-to-end and document performance. If the changed test does not meet performance requirements, redesign may be necessary. Depending on whether changes were made to the software in order to mask or unmask genes for which the test performance was already demonstrated as part of the original validation, the validation status of the software may need to be re-established.

- Include a procedure to account for updates to internal and external genetic variant databases and their potential impact on the evaluation of variants. Document any updates including but not limited to the name, location, or new version of the database.

IX. Additional Resources

- FDA guidance document entitled "Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classifications." (https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm517504.pdf).

- FDA guidance document entitled "Requests for Feedback on Medical Device Submissions: The Pre-Submission Program and Meetings with Food and Drug Administration Staff." (https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm311176.pdf).

- FDA guidance document entitled "Procedures for Class II Device Exemptions from Premarket Notification." (https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm080199.pdf).

- 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). (https://www.palmettogba.com/palmetto/MolDX.nsf/DocsCat/MolDx%20Website~MolDx~Browse%20By%20Topic~Technical%20Assessment~9WRHPN3576?open&navmenu=Browse%5eBy%5eTopic).

- "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.

- Jennings, L.J. et al, "Guidelines for Validation of Next-Generation Sequencing–Based Oncology Panels A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists," The Journal of Molecular Diagnostics (2017), 19(3):341-365

X. Appendix A

Here, we provide a purely illustrative, simplified example, where the numbers and terms used are not indicative of the actual samples or performance that could be used to demonstrate the safety and effectiveness of a test. Test developers should calculate accuracy for each variant type, in the context in which the test can detect them, as well as for clinically relevant variants. Consider an accuracy study of a test with 6 interrogated regions with a single variant in each region, including the following 7 samples:

- 5 samples with variant calls (based on the comparator method)
- 2 samples with wild-type (wt) calls at the variant positions of interest (based on the comparator method).

The description of the accuracy study results can be presented by table A.1 below where colors of the cells represent the truth (with dark gray cells representing regions with true variants of interest based on the comparator method, and light gray cells representing true wt regions based on comparator method), while the wording in the cells represents the test results.

Table A.1

	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6
Sample 1	Variant correct	wt	Variant incorrect	wt	Variant	wt
Sample 2	wt	Variant correct	Variant	wt	Wt	wt
Sample 3	wt	wt	Variant correct	wt	Variant correct	wt
Sample 4	wt	wt	wt	Variant correct	Wt	wt
Sample 5	Variant correct	wt	wt	wt	Wt	Variant correct
Sample 6	wt	wt	wt	wt	Wt	wt
Sample 7	Variant	wt	wt	wt	Wt	wt

In this example:

- Samples 1, 2, 3 and 5 have 2 regions with variants and 4 regions with wt;
- Sample 4 has 1 region with variant and 5 regions with wt;
- Samples 6 and 7 have no regions with variants and 6 regions with wt.

I) Estimate accuracy measures PPA, NPA and TPPV for *each sample separately* across all interrogated regions as it is presented in table A.2. Provide also percentage of "no calls or invalid" results for each sample.

Table A.2

			Comparator method				PPA	NPA	TPPV	%inv.
			Variant	wt	Total					
Sample	Test	Variant correct	A_1	B	A_1+A_2+B		A_1 / (A_1+A_2+C)	D / (B+D)	A_1 / (A_1+A_2+B)	
		Variant incorrect	A_2							
		wt	C	D	C+D					
	Total		A_1+A_2+C	B+D	$A_1+A_2+B+C+D$					

For samples 6 and 7, calculations of PPA are not applicable.

Examples of calculations for sample 1 and 2 are presented in tables A.3 and A.4:

Table A.3

			Comparator method				PPA	NPA	TPPV	%inv.
			Variant	wt	Total					
Sample 1	Test	Variant correct	1	1	3		50.0% (1/2)	75.0%(3/4)	33.3% (1/3)	0% (0/6)
		Variant incorrect	1							
		wt	0	3	3					
	Total		2	4	6					

Table A.4

			Comparator method				PPA	NPA	TPPV	%inv.
			Variant	wt	Total					
Sample 2	Test	Variant correct	1	1	2		50.0% (1/2)	75.0% (3/4)	50.0% (1/2)	0% (0/6)
		Variant incorrect	0							
		wt	1	3	4					
	Total		2	4	6					

Investigate whether PPAs and NPAs are similar across all 7 samples.

II) Estimate accuracy measures PPA, NPA and TPPV for *each variant separately* across all samples as it is presented in table A.5. Provide also percentage of "no calls or invalid results for each variant.

Table A.5

			Comparator method				PPA	NPA	TPPV	%inv.
			Variant	wt	Total					
Variant	Test	Variant correct	A_1	B	A_1+A_2+B		A_1 / (A_1+A_2+C)	D / (B+D)	A_1 / (A_1+A_2+B)	
		Variant incorrect	A_2							
		wt	C	D	C+D					
	Total		A_1+A_2+C	B+D	$A_1+A_2+B+C+D$					

If a clinical action based on the test results of incorrect variant is the same as based on the correct variant test results, calculate in addition
PPA = $(A_1+A_2)/(A_1+A_2+C)$ and
TPPV = $(A_1+A_2)/(A_1+A_2+B)$.

In the example of the accuracy study with 6 interrogated regions:

- Variants 1, 3 and 5 have 2 samples with variants and 5 samples with wt;
- Variants 2, 4 and 6 have 1 sample with variants and 6 samples with wt.

Examples of calculations for variant 1 (region 1) and variant 2 (region 2) are presented in tables A.6 and A.7:

Table A.6

			Comparator method				PPA	NPA	TPPV	%inv.
			Variant	wt	Total					
Variant 1	Test	Variant correct	2	1	3		100% (2/2)	80.0% (4/5)	66.7% (2/3)	0% (0/6)
		Variant incorrect	0							
		wt	0	4	4					
	Total		2	5	7					

Table A.7

			Comparator method				PPA	NPA	TPPV	%inv.
			Variant	wt	Total					
Variant 2	Test	Variant correct	1	0	1		100% (1/1)	100% (6/6)	100% (1/1)	0% (0/6)
		Variant incorrect	0							
		wt	0	6	6					
	Total		1	6	7					

Investigate whether PPAs and NPAs are similar across all 7 regions (variants).

III) Estimate accuracy measures PPA, NPA and TPPV combined across all variants and samples as it is presented in table A.8. Provide also percentage of "no calls or invalid results" combined across all variants and all samples.

Examples of calculations are presented in table A.8:

Table A.8

			Comparator method				PPA	NPA	TPPV	%inv.
			Variant	wt	Total					
Variant	Test	Variant correct	7	3	11		77.8% (7/9)	90.9% (30/33)	63.6% (7/11)	0% (0/42)
		Variant incorrect	1							
		wt	1	30	31					
	Total		9	33	42					

PPA = 77.8% (7/9) with 95% CI: (50.4%; 92.4%);
NPA = 90.9% (30/33) with 95% CI: (79.3%; 96.3%).

TPPV depends on PPA, NPA and also on the percentage of regions of wt among all interrogated regions in the samples in the accuracy study (in the example, this is 78.6% (33/42)); therefore, an interpretation of this metric should take this into consideration.

Note: All sample and variant numbers and PPA, NPA, TPPV values are provided as an extremely simplified illustration of basic calculations and should not be considered examples of acceptable levels for PPA, NPA and TPPV. As described in the section VI.B.1, acceptable numbers will primarily depend on the indications for use of the test and its performance characteristics.