**UPDATED AND REVISED** October 2024

**Oncology – Molecular and Cellular Tumor Markers**

**Next Generation Sequencing (NGS) guidelines for somatic genetic variant detection**

The following describes requirements for the development of procedures and the establishment of performance (validation) of assays for the detection of genetic germline variants by Next Generation/massively parallel sequencing (NGS) technologies. These requirements should be used in conjunction with and not in lieu of any other applicable checklist at https://www.wadsworth.org/regulatory/clep/clinical-labs/obtain-permit/test-approval. Overall, clinical validation of NGS assays follows the same basic principles for validating most other complex molecular diagnostic procedures. It is anticipated that these guidelines will evolve as the field matures. Please make sure you use the most up-to-date version of these guidelines.

Items that must specifically be addressed in the validation submission include:

**SOP**

- Summarize the purpose of the test and the indications for testing (intended use).
- Describe the region(s) of the genome targeted by the assay as well as the type(s) of variants the assay is intended to detect, with any associated limitations.
- Provide a step-by-step description of the entire testing process, from sample receipt through library preparation, sequencing, data analysis and interpretation, including detailed steps of all automated processes. If a third party performs any function, including data analysis and/or interpretation, then the protocols must detail their specific involvement in the overall testing process.
- Include the criteria for when confirmatory testing is performed and indicate the specific methods used.
- State the acceptability criteria for all variant types as defined by your validation studies.

**QC**

- Quality control metrics and acceptance criteria must be clearly defined in the SOP.
- Quality of base scoring must meet a minimum of Q20 or equivalent per base for all single nucleotide polymorphism (SNP) and insertion deletion (indel) calls. Quality of base scoring for copy number variants (CNVs) and other structural variants must be clearly defined based on validation studies.
- Establish minimum criteria for depth and uniformity of coverage, i.e., number of reads, across all target areas (genes). A minimum average of 500 unique reads or greater is **strongly** recommended.
- Define the minimum coverage required for a target area below which you cannot confidently define the area's mutation status. The minimum target area coverage required may be different for confidently calling a 'variant detected' vs. having adequate depth of coverage to confidently call the target area free of detectable variants.
- Define the minimum percentage and/or number of variant reads in a background of normal reads required to call a variant 'detected' at your established level of confidence and sensitivity.
- Define maximum allowable strand bias (if applicable).
- New reagent lots require verification/confirmation of the analytical sensitivity and specificity to ensure that low positives will not be missed by new lots of reagents. This applies to all critical reagents and includes depth and

uniformity of coverage to detect possible target area drop out.

- All software updates that affect key parameters, such as base calling, alignment, etc., must be revalidated using data from at least 3-5 previously analyzed runs to verify that all variants are still detected with the same analytical sensitivity and specificity as previously determined. The revalidation process must be clearly described in the SOP.
- Data retention: All FASTQ files (or equivalent) should be maintained for a minimum of 2 years per *Document and Specimen Retention Standard of Practice 8 (DSR S8): Analytic System Records Retention.*
- All QC metrics must be followed and documented over time to verify that there is no decrease (drift) in performance, e.g., Levy-Jennings plots, tracking variant allele fractions (VAFs) of positive controls.

**Controls**

- A **No Template Control** (NTC) must be included and taken through the entire testing process (including sequencing) to verify that there is no contamination across samples and reagents.
- A **positive/sensitivity control** should be included in each run. We suggest this control be a low positive sample (near the sensitivity of the assay) containing multiple known variants, of each kind to be detected by the assay, to verify that low percentage variants can be identified consistently. A defined rotation schedule should be employed if not all variant types can be incorporated in a single control sample and/or if multiple positive controls are used.

**Reports**

- Reports should include all detected somatic variants, whether of known or unknown clinical significance, in a manner clearly identifying each variant's significance with supporting level of evidence as described in the AMP guidelines (http://jmd.amjpathol.org/article/S1525- 1578(16)30223-9/pdf). However, we would not object to the inclusion of a statement such as *"This test is designed to detect x, y and z…in genes a, b and c… However, variants other than the ones listed above may also have been detected. If interested, these can be released upon request"*.
- Variants can only be reported if they are confirmed by an independent reference method or if the assay has been fully validated and confirmation is no longer necessary (see Validation section below for more information on confirmation testing and full validation requirements).
- Reports must clearly define the laboratory's policy for reporting incidental findings (see the most recent list published by American College of Medical Genetics and Genomics). We suggest you include these on your report separately and alert the treating physician to their potential clinical relevance, with a recommendation that the variants should be confirmed by a laboratory permitted to perform germline genetic testing.
- Reports must include a statement(s) that identify the limitations of the assay, including which target area(s) of the assay failed to meet the lab's acceptability criteria.
- Reports must indicate any limitations in the detection of specific types of variants (e.g., CNV), maximum length of indels) and sensitivity for all variant types targeted by the assay.
- If the full list of targeted genes is not included on the report, then reference to where this list can be found must be included on the report.

**Validation**

- Provide a detailed description of all validation studies. All data must include, at minimum, gene/variant nomenclature, allele fractions, coverage, and Q score.
- Performance characteristics must be **established and validated separately** for **each type** of variant the assay is intended to detect, e.g., SNV, indel, CNV, and structural variants, as well as microsatellite instability (MSI) status and tumor mutation burden (TMB).
- Performance characteristics for each sample type (e.g., formalin fixed paraffin-embedded tissue, fresh/frozen tissue, whole blood, bone marrow, fine-needle aspirate) must be established and validated, along with demonstration of quality sequences for all target areas without sample type bias. Areas that consistently fail to meet minimum quality metrics must clearly be defined. **Note:** circulating-tumor DNA (ctDNA) requires separate

full validation from all other specimen types (see Initial Validation bullet below).

- **Minimum** data required to establish key performance characteristics (please include the prevalence, i.e., number of normal and variant reads of each detected variant for all studies; a table and graph (e.g., histogram or box-and-whisker plot) of read depths for each target area is recommended):

  - **Analytical accuracy**: Sequence a minimum of 2 well-characterized reference specimens (e.g. HapMap DNA NA12878, NA19240, or Genome in a Bottle) to determine a robust laboratory specific error rate across all areas targeted by your assay (specificity). This error rate is expected to be < 2%.

  - **Initial validation**: Must include a minimum of 50 unique patient samples comprising specimens of all intended sample and tumor types. These samples must contain a representative distribution of reportable variants across all target areas (including GC-rich and homopolymer sequences) and must be confirmed by an independent reference method. The independent reference method cannot utilize the same technology as the NGS platform unless it is performed in a different lab. Provide a summary table of variants detected in each sample, as well as independent reference method results as obtained from independent reference method testing.

    - ctDNA based assays must meet all validation requirements independently, i.e., must have 50 ctDNA samples for initial validation as well as meet all requirement for full validation independently (see below).

    - **NOTE**: Methylation specific NGS and RNA transcript sequencing requires their own validation. Contact clepval@health.ny.gov to discuss requirements.

- **Full validation**: For reported variants with clinical significance identified during clinical runs that have not yet been fully validated an ongoing validation, i.e., confirmation by an independent reference method, must be performed until such a time that the criteria outlined below have been met.

  - Full validation for SNVs will be considered in 3 separate categories:
    1) Standard sequence (not GC rich or homopolymer)
    2) GC rich (60% or greater)
    3) Homopolymers (≥ 3 bases)

    Confirmation is required within each category until the following number of cases have been reached:

    For Category 1, full validation requires confirmation of 50 SNVs with no more than 5 cases in a particular gene. Once an individual gene under category 1 has been confirmed 5 times that gene no longer requires confirmation before reporting.

    For categories 2 **and** 3 (they are treated separately), full validation requires confirmation of 75 SNVs for each category with no more than 5 confirmations in a particular gene. Once an individual gene under categories 2 and 3 has been confirmed 5 times that area no longer requires confirmation before reporting.

  - Full validation for indels will be considered in 3 separate categories:
    1) Standard sequence (not GC rich or homopolymer)
    2) GC rich (60% or greater)
    3) Homopolymers (≥ 3 bases)

    Confirmation is required within each category until the following number of cases have been reached:

    For Category 1, full validation requires confirmation of 50 indels with no more than 5 cases in a particular gene. Once an individual gene under category 1 has been confirmed 5 times that gene no longer requires confirmation before reporting.

    For categories 2 **and** 3 (they are treated separately), full validation requires confirmation of 75 indels for each category with no more than 5 confirmations in a particular gene. Once an individual gene under categories 2 and 3 has been confirmed 5 times that area no longer requires confirmation before reporting.

- Full validation for CNVs will be considered in 3 separate categories:

  1) Single gene gains and losses
  2) Large gains and losses that cover entire chromosomal arms
  3) Moderate gains and losses that fall in size between the above two categories, typically megabase in size.

  Confirmation is required within each category until the following number of cases have been reached:

  For category 1, full validation requires confirmation of 100 gains/losses with no more than 10 cases at a particular locus. Once an individual locus under category 1 has been confirmed 10 times that area no longer requires confirmation before reporting.

  For categories 2 **and** 3 (they are treated separately), full validation requires confirmation of 50 gains/losses for each category with no more than 5 confirmations at a particular locus. Once an individual locus under categories 2 and 3 has been confirmed 5 times that area no longer requires confirmation before reporting.

- Translocations will require a minimum of 3 confirmations per gene fusion partner **targeted** by the assay. For example, if the assay targets ALK then you need only confirm 3 positive ALK rearrangements to fully validate **any** rearrangements that contain the ALK gene as one of the fusion partners.
- MSI: Confirm 10 microsatellite stable and 10 MSI-H samples.
- TMB: TMB cut-off values must be established and verified separately for different tumor/tissue types.

- **Analytical sensitivity**: Establish the analytical sensitivity of the assay for each type of variant detected by the assay at the lower limit of nucleic acid input. This can initially be established with defined mixtures of cell line nucleic acids (not plasmids) but needs to be verified with 3-5 unique patient samples. You need to establish both the lowest input amount of nucleic acid that still gives reliable results and the lowest VAF that can reliably be detected at that amount of input nucleic acid.

- **Reproducibility studies** (between run) for all intended variant types should be performed at the lower limit of nucleic acid input and VAF. If possible, perform these studies by multiple technologists and sequencers.
  - For SNVs and indels, 4-5 unique positive patient samples must be analyzed in 4-5 separate runs using different barcodes (from the original nucleic acid through sequencing and data analysis). A minimum of 20 replicates each must be analyzed and achieve 95% overall reproducibility.
  - For structural variants a minimum of 3 unique positive patient samples (each targeting different fusion partners) must be analyzed in three separate runs.
  - For CNVs a minimum of 3 unique positive patient samples for gains and 3 unique positive patient samples for losses (all in different genes) must be analyzed in three separate runs.
  - For MSI, a minimum of 3 MSS and 3 MSI-H patient samples must be analyzed in three separate runs.
  - For TMB, a minimum of 3 patient samples in each class (e.g. low, indeterminate, high) must be analyzed in three separate runs.
- The initial validation studies should be done with a single version of all software utilized.