**Introduction**

Next-generation sequencing (NGS) has expanded beyond research applications to deliver clinically actionable test results that can effectively inform medical decision making. This new area of clinical testing lacks uniform practices for quality assurance (QA) and quality control (QC) that are essential to ensure the analytic validity of test results. Guidance is also essential to the research community to ensure that novel technologies are able to meet or exceed current standards of practice. A national workgroup was convened to take a science- and quality management-based approach to define metrics and processes not currently described that are needed to establish analytic validity and to meet regulatory and professional standards.

**Methods**

- Establish a national workgroup
  - Forty-one participants with expertise in NGS and clinical lab testing
  - Laboratory directors, clinicians, platform developers, researchers, software developers, bioinformatics experts, and policy makers
- Two day meeting (April 2011) with continued consultation with workgroup members
  - Areas of focus
    - Emphasis on heritable sequence variations
    - Identification and application of quality metrics that are platform independent
    - Focus upon test validation, QC, reference materials (RMs), proficiency testing (PT) and alternate assessment (AA)

**Results: Establishing test performance and QC for NGS**

**Test Validation**

- For NGS, validation is the process of establishing the analytical performance specifications and determining QC practices needed for the future patient testing within the laboratory in which the test is performed.
- The complexity of NGS requires consideration for both platform and test specific validation processes.

<table>
<thead>
<tr>
<th>Performance Characteristics</th>
<th>Workgroup established definitions for NGS applications</th>
<th>Workgroup established metrics and processes for evaluation of NGS analytic performance</th>
</tr>
</thead>
</table>
| Accuracy                    | The closeness of agreement between a measured value and the true value, which for NGS is the accepted reference sequence. | • Coverage - The number of independent overlapping base calls made at a given position  
  • Depth of coverage  
  • Average coverage  
  • Uniformity or distribution of coverage  
  • Quality scores - The confidence in a base or variant call |
| Precision                   | The degree to which repeated measurements give the same result (repeatability and reproducibility). | • Monitor performance for:  
  - Library variability: independent library preparations  
  - Intra-run variability: same sample, same library, same run  
  - Inter-run variability: same sample, same library, different runs  
  - Inter-operator variability |
| Analytic Sensitivity        | The likelihood that the assay will detect a sequence variation, if present. | Depth of coverage must be sufficient to minimize a loss of sensitivity and specificity. The depth of coverage achieved with NGS will vary across the genome and therefore should be established across all regions of the sequence targeted for the clinical application. Analysis of RMs possessing comparable types of sequence variations across the targeted region by an orthogonal technique can provide a useful comparator. |
| Analytic Specificity        | The probability that the assay will not detect a sequence variation, if not present. | |
| Reportable Range            | The regions of the genome for which the NGS technology can accurately produce sequence information (e.g. multiple genes, exome, large genomic regions). | Define areas of difficulty (e.g. repeat regions, insertions and deletions, allele dropouts) near the regions of interest. Biases introduced by capture-based or enrichment methods should be identified. |
| Reference Range             | Establishment of reportable sequence variations expected to occur in the target population that the assay can detect. | Materials containing the type of sequence variation(s) appropriately distributed within the target sequence may establish the capacity of the test to detect similar disease-associated mutations. |

**QC/QA Considerations**

- Monitoring established performance specifications
  - Coverage
  - Quality Scores
  - Allelic read percentage
  - Mapping Quality
- Use reference materials to monitor assay performance and for PT/AA
  - Reference materials are needed that contain the range and distribution of sequence variations comparable to those which the assay is designed to detect.
  - Method-based PT may be a component of an inter-laboratory comparison because of the size of the genome interrogated and the number of potential sequence variants targeted.
  - PT/AA Materials useful for NGS:
    - DNA from a well characterized cell line or patient sample (PT)
    - Electronic data (PT)
    - Inter-laboratory sample or electronic data exchange (AA)
- Use informative DNA with sequence variations comparable to those which the assay is designed to detect.

**Conclusions**

- Recommendations were made for defining performance characteristics for NGS
- Validation should establish a depth of coverage sufficient to minimize a loss of sensitivity and specificity
- Method-based PT challenges should take into account that different laboratories will interrogate different regions of the genome
- Reference materials and sequences used for test validation, QA/QC, and PT/AA should include the types of sequence variants targeted by the test and appropriately distributed across the targeted region of the genome. This is in addition to use of actual patient samples that would be expected to have only a subset of the targeted sequence variants.
- Contact Information: Amy Gargis, AGargis@cdc.gov

---

**What are the steps of NGS that require test validation and QA/QC?**

**Platform Validation**

To establish an acceptable error rate in the sequencing of a reference material for each NGS technology and informatics analysis tool.

**Test Validation**

To establish an acceptable error rate for detection of specific targeted sequence variants for each test's technology and informatics data analysis.

---

**Example Alignment**

```
<table>
<thead>
<tr>
<th>Reference Sequence</th>
<th>A/C</th>
<th>Reference Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>TACGCTAGCTAGCTAGCTA</td>
<td>A/C</td>
<td>Allelic Read Percentage: C = 53%, A = 47%</td>
</tr>
<tr>
<td>TACGCTAGCTAGCTAGCTA</td>
<td>A/C</td>
<td>Allelic Read Percentage: C = 53%, A = 47%</td>
</tr>
</tbody>
</table>
```

**Depth of Coverage**

- Inter-laboratory sample or electronic data exchange (AA)
- Uniformity or distribution of coverage
- Intra-run variability: same sample, same library, different runs
- Inter-operator variability

---

**Office of Surveillance, Epidemiology, and Laboratory Services**

**Division of Laboratory Science and Standards**

E-mail: cdchfs@cdc.gov | Web: www.cdc.gov

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.