Current practices and guidelines for clinical next-generation sequencing oncology testing

Samuel P. Strom
Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA

ABSTRACT

Next-generation sequencing (NGS) has been rapidly integrated into molecular pathology, dramatically increasing the breadth genomic of information available to oncologists and their patients. This review will explore the ways in which this new technology is currently applied to bolster care for patients with solid tumors and hematological malignancies, focusing on practices and guidelines for assessing the technical validity and clinical utility of DNA variants identified during clinical NGS oncology testing.

KEYWORDS

Cancer genomics; next-generation sequencing; molecular diagnostics

Introduction

Massively parallel sequencing of nucleic acids enables DNA and RNA analysis on a grand scale. A natural implementation of this "next-generation sequencing (NGS)" technology is to assess the unique and complex set of genomic alterations that occur in malignant neoplasms, with the goal of improving patient care through personalized diagnosis, prognosis, and therapy.

The most prevalent current implementation of NGS for oncology is mutation detection via targeted panels\textsuperscript{1-5}. These assays use molecular methods such as multiplex polymerase chain reactions (PCR) to isolate clinically relevant segments of the genome, such as mutation hotspots or coding exons of entire genes. These panels range from a few hundred target loci to many thousands. In these assays, raw sequence reads are first aligned to the reference human genome. Variant calling is then performed to identify small mismatches in these alignments which may represent mutations present in the specimen. Variant analysis and interpretation must then be performed to assess the technical validity and clinical utility of each variant (Figure 1).

Before any variant analysis is performed, the data must be checked for overall assay performance and quality. As many surgical pathology specimens have limited tissue amounts and quality, and have been treated with formalin fixation, assay failures occur with some regularity. As there are many different ways to measure assay performance, each assay requires a unique set of parameters established during validation\textsuperscript{6,7}.

Technical validity and clinical utility are the two major issues that must be resolved for every variant identified via NGS somatic variant detection. If a detected variant fails to meet or exceed predetermined validity and utility criteria, it should not be clinically reported as medically relevant. While simple in principle, there exists deep complexity in these areas which merit close consideration.

Terminology

NGS introduces a new suite of vocabulary into the clinical lab, with some specific to oncology testing. The following glossary is provided as reference and to facilitate further discussion in this review.

Genomic analyst (GA)

GA is typically a doctorate or masters level specialist who performs some or all of the initial review of assay quality, technical validity assessment of detected variants, clinical utility assessment of variants, and generating draft reports. Final reports must be signed out by either a board certified MD clinical pathologist or PhD with clinical laboratory board certification.
Base call quality score (Q score)

Q score is a PHRED-scaled probability ranging from 0-20 inversely proportional to the probability that a single sequenced base is correct. For example, a T base call with Q of 20 is considered likely correct with a confidence \( P \)-value of 0.01. Any base call with Q<20 should be considered low quality, and any variant identified where a substantial proportion of reads supporting the variant are of low quality should be considered potentially false positive.

Read depth

Read depth (or coverage, conventionally a number followed by "x") is the number of independent reads with overlapping alignment at a locus of interest. This is often expressed as an average or percentage exceeding a cutoff over a set of intervals (such as exons, genes, or panels). For example, a clinical report might say that a panel average coverage is 1,105x with 98% of targeted bases covered >100x.

Mutations

Mutations are events that result in changes of genomic DNA. Variants are deviations from the human genome reference sequence observed in a specimen. The goal of NGS cancer panel testing is to identify variants which are extremely likely to be caused by somatic mutation. The difference is subtle; the two terms are often used interchangeably.

Variant read number (variant reads)

Variant reads is the number of independent sequence reads supporting the presence of a variant. Due to the high error
rate of NGS at the per-base call level, calls supported by fewer than 5 variant reads are typically considered to be likely false positive calls.

Variant allele frequency (VAF)

VAF is the percentage of sequence reads observed matching a specific DNA variant divided by the overall coverage at that locus. Because NGS provides a near random sample, VAF is thus a surrogate measure of the proportion of DNA molecules in the original specimen carrying the variant. For constitutional genetic testing, VAF is a measure of diploid zygosity: heterozygous loci should be near 50% VAF, homozygous loci should be near 100%, and reference loci should be near 0%. Deviations from these three expected values should be considered suspicious as potential errors due to incorrect base calls or alignment. For somatic testing, contamination from normal cells and tumor heterogeneity combine to cause unpredictable VAFs. In some clinical scenarios, such as testing a patient for therapeutic resistance mutations, the desired sensitivity dictates that variants with low VAF be included in downstream analysis. This all means that in order to maintain an acceptable level of sensitivity, variant filtering for somatic variant must be highly permissive from a technical validity perspective.

Variant quality scores (QUAL)

QUAL are generated during the variant calling step and a requisite component of the Variant Call File (VCF). Generally speaking, QUAL scores are transformed log-scaled (PHRED) values where, for example, a score of 90 supports the variant call with a P-value of $1 \times 10^{-9}$. It is critical to keep in mind that there are significant multiple testing considerations due to the large number of variants that are detected by NGS. The scale for QUAL values vary widely and depend on assay platform, capture method, and variant calling software. As such, it is strongly recommended at this time that each laboratory performing NGS independently assess the performance characteristics of QUAL scores for each assay, preferably by orthogonal testing.

Strand bias (SB)

SB is a measure of how far the observed variant reads deviate from an expectation of equal likelihood of sequencing the plus and minus strands. A high SB score indicates that the variant call may be caused by an artifact of the alignment process rather than a true mutation. Certain targeting methodologies, including those using Illumina TruSight capture techniques selectively amplify one target strand; for these assays, SB is not a meaningful measure of quality.

Technical validity

To date, there are no NGS oncology in vitro diagnostics (IVD) cleared by the United States Food and Drug Administration (FDA). As such, all current clinically validated NGS oncology assays are classified as "laboratory developed" tests (LDT) which can vary highly in content and quality from lab to lab. In the US, LDTs must be performed in laboratories with Clinical Laboratory Improvement Amendments (CLIA) licensure and are subject to state and federal regulations. The New York State Department of Health has issued guidelines for NGS oncology testing, setting a very high bar for clinical validation and ongoing assay performance (Table 1).

For labs without New York approval, there are no matching federal guidelines at this time. As such, it is strongly recommended to prefer laboratories with optional College of American Pathology (CAP) accreditation. The major benefit of CAP accreditation comes from proficiency testing participation, wherein laboratories performing similar assays are given identical reference materials for comparison of results. This proficiency testing gives labs a mechanism and incentives to continually improve their test methodologies by direct comparison with others (Table 1).

Orthogonal confirmatory testing can validate that detected variants is a direct means of assuring technical validity. This requires a laboratory to have access to duplicate sequencing of entire panels using alternate methodologies (e.g. sequencing using both Life Technologies and Illumina platforms for every specimen), or to have a single-locus test of equivalent sensitivity for every targeted interval. As neither approach is simple or inexpensive to implement, many laboratories eschew such confirmatory testing during routine analysis. While PCR-based dideoxy sequencing ("Sanger" sequencing) is a generally accepted method of confirmatory testing for constitutional clinical NGS, its sensitivity is limited to approximately 10%-20% VAF. As cancer NGS testing routinely identifies clinically relevant mutations from 5%-10% VAF, it Sanger is not an ideal method for confirmation for all variants.

Manual review of aligned reads in a genome browser ("variant inspection") can greatly reduce the risk of false positive or incorrect results (Figure 2). While variant calling software has become increasingly sophisticated, it still cannot...
yet perform as well as a GA when it comes to pattern recognition. For cancer NGS, it is best to think of the role of variant calling software as to alert the GA to loci of interest. From there, it is the job of the GA to assess first whether any variant is present and if the variant called is the correct change. In some cases, variant inspection will reject dubious variant calls which have received a high QUAL score; this is the necessary cost of setting high sensitivity. In other cases, the variant observed will differ from that called. Complex insertion/deletion events are especially susceptible to this type of error. Given how common this mutation type is in certain tumors, laboratories must be prepared for this issue to be routine rather than exceptional.

There are many metrics that can be used to assess if a variant call is likely a true positive result. VAF, QUAL, SB, and other metrics may be available depending on the variant callers used (Figure 1). As each LDT NGS assay performs differently, it is necessary to establish filtering criteria anew for all new tests. If hard cutoffs are to be used to filter out low quality variants, a rigorous validation should be performed, including accuracy and limit of detection experiments to determine and test the cutoffs. Alternatively, very inclusive filters can be set to minimize false negative risk. This will increase the burden of confirmatory analysis and/or testing required for each case. Both options can be rate-limiting for small laboratories.

During the validation phase of assay development, it is valuable to test assay performance using reference materials. DNA or false FFPE blocks from positive control cell lines are commercially available for many of the most recurrent mutations. Some reference materials now available contain a variety of mutations with different VAF, genes, and mutation types. These are designed to be run during validation and/or during routine testing to ensure ongoing assay performance. Acrometrix (a division of Thermo Fisher Sciences Inc. Waltham, MA, United States), and Horizon Diagnostics (a division of Horizon Discovery Group plc. Waterbeach, United Kingdom) are two commercial vendors specializing in this area.

### Clinical validity

It is tempting to think of tumors as "gene positive" or "gene negative". This is a convenient way to discuss certain classical oncogenic changes such as fusion genes, but it glosses over a significant layer of complexity we are only now beginning to understand. For example, some mutations in EGFR render a tumor more susceptible to targeted tyrosine kinase inhibitor (TKI) therapy; other mutations confer resistance to the same\textsuperscript{21-24}. Certain KRAS mutations are known to be targetable driver mutations while others are commonly observed but seem to have little to no impact on drug response (Figure 1)\textsuperscript{1,25-26}. These now classic examples are the vanguard of a new way of thinking about cancer genes and mutations.

For each detected variant, there are multiple different types of considerations a GA should weigh.

<table>
<thead>
<tr>
<th>Source</th>
<th>Title</th>
<th>Content summary</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York State Board of Health</td>
<td>“Next Generation” Sequencing (NGS) guidelines for somatic genetic variant detection</td>
<td>Detailed standards for technical validity</td>
<td>11</td>
</tr>
<tr>
<td>ACMG</td>
<td>Standards and guidelines for the interpretation of sequence Variants</td>
<td>Guidelines for clinical validity assessment, particularly for germline/constitutional variants</td>
<td>12</td>
</tr>
<tr>
<td>ACMG</td>
<td>ACMG clinical laboratory standards for next-generation sequencing</td>
<td>Broad summaries of major areas of consideration for clinical validation of all NGS assay</td>
<td>13</td>
</tr>
<tr>
<td>CDC</td>
<td>Assuring the quality of next-generation sequencing in clinical laboratory practice</td>
<td>Detailed recommendations for technical validity assessment/validation of all NGS assays</td>
<td>14</td>
</tr>
<tr>
<td>CDC</td>
<td>Good laboratory practice for clinical next-generation sequencing informatics pipelines</td>
<td>Detailed recommendations for clinical validity assessment for all NGS assays</td>
<td>15</td>
</tr>
<tr>
<td>Quest Diagnostics (reference laboratory)</td>
<td>Annotation of sequence variants in cancer samples processes and pitfalls for routine assays in the clinical laboratory</td>
<td>The framework for a repeatable workflow for clinical validity assessment in use at a high volume testing facility is described</td>
<td>16</td>
</tr>
</tbody>
</table>

ACMG: American College of Medical Genetics and Genomics; CDC: United States Centers for Disease Control and Prevention
Somatic vs. germline

If normal tissue is available and tested, variant subtraction can be used to directly assess which DNA variants arose in somatically in the tumor. As this is not always feasible (consider hematological malignancies, where blood specimens contain neoplastic cells) or practical, indirect measures should be taken to avoid reporting benign germline variants as relevant for oncology. The front-line tool for indirectly filtering out germline variants is population allele frequency. Any variant that is present on >1% of normal human chromosomes is almost certainly not clinically relevant for cancer. VAF databases such as the Exome Aggregation Consortium (ExAC) or dbSNP (which uses 1,000 genome project allele frequencies) should be used to perform this filtering. As they match expectations for heterozygous and homozygous frequencies respectively, variants with VAF at nearly 50% or 100% should be considered potentially germline during analysis. Hard filtering for VAF however is not recommended as true mutations can certainly be observed at low VAF.

Tumor type

Many mutations impact different tumor types in different ways, or have only been studied rigorously in a specific set of tumor types but are recurrently observed in other types. This means that variant annotation and interpretation procedures must take tumor type into account prior to rendering a clinical report. Resources such as COSMIC, The Cancer Genome Atlas (TCGA), and MyCancerGenome.org (MCG) have publically available databases with tumor-type specific mutation information and should be referred to when performing variant analysis. Literature review may be necessary for rare mutations.

VAF and tumor cellularity

If the VAF and percent tumor cell prediction from pathology are highly discordant, it could indicate an invalid or unreportable variant. If the VAF is far higher than expected, it could indicate that the variant is either germline or in a region of loss of heterozygosity (LoH). If the variant is in a gene where LoH is not anticipated (e.g. an oncogene where activating mutations are the mechanism), germline may be
the more likely explanation. A VAF far lower than expected could indicate either false positivity or low level tumor heterogeneity. If the technical validity of a lower than expected VAF variant is very strong, tumor heterogeneity is more likely. If this is the case, the relevance of the mutation may depend on contextual considerations such as mutation type, tumor type, and patient treatment history. For example, the observation of a low relative VAF of a known recurrent acquired resistance mutation is of much stronger clinical validity in a treated patient than the same observation in a treatment naive patient.

It is also important to remember that percent tumor cell prediction, especially surgical pathology/histology, are essentially qualitative assessments never intended to have highly accurate or precise measurements. This is not an indictment of pathologists, rather a statement of fact that the methods used are not high enough resolution to be used in this way. Genomic analysts, who may or may not have any training in anatomic pathology, should be aware that these assessments may be off by as much as 10%-20% when performed by the best practitioners. Flow cytometry assessment of hematological malignancies should be a much more precise and accurate quantitation, though the correlation between flow cytometry results and NGS VAFs is as yet poorly understood.

**Highly recurrent mutations**

The observation of certain mutations, even at borderline technical quality, represents a "smoking gun" and must be reported. Examples such as *BRAF* p.V600E and *KRAS* codon 12/13 mutations should come to mind immediately for experienced cancer genomic analysts. These variants must be reported.

**Clinically significant mutation types**

In-frame insertions/deletions (indels) and loss of function (LoF) mutations are two mutation types which can present as apparently novel variants at the DNA level despite having well-known implications. Examples include in-frame indels in *EGFR* exon 19 or LoF mutations in *TP53*. As practically all genomic analysis software solutions perform comparisons at the DNA (or predicted mRNA) level, this pattern can cause patient variants to be incorrectly classified as of uncertain clinical significance. Until such a times as more sophisticated software tools are available, it is the role of the genomic analyst to determine whether apparently novel variants are in fact of known clinical significance.

This area can become quite complex. For example, acute myeloid leukemia (AML) patients positive for biallelic LoF mutations in *CEBPA* have a favorable prognosis only if one mutation is in the C terminal region and the other is in the N terminal region\(^1\). This means that a specific mutation seen alone should be considered of uncertain significance, whereas that same mutation in conjunction with a second mutation would be clinically relevant to prognosis. This pattern can be challenging to convey in standard databases, requiring either highly specialized software or manual review by experts.

**Variants of uncertain clinical significance (VUS)**

The number of VUS identified in a given specimen closely correlates with the scale of targets analyzed. As more genes or hotspots are added to tests, more VUS will be observed. Whether and how to report VUS is highly laboratory and test dependent: some labs will always report all VUS, some will never report any, and some will report only when specific conditions are met. The ordering oncologist should be aware of VUS reporting policies and utilize the test that best addresses the clinical needs of their patient. Contacting a lab directly is the best way to ascertain VUS reporting policies; if no such policy exists or cannot be readily ascertained, it may be wise to consider a provider with better clarity. However, due to a lack of evidence supporting the clinical utility of VUS, reporting of VUS should not be considered a necessary component of tumor sequencing analysis.

**Integration of cytogenetics information**

There are clinically significant cytogenetic abnormalities or nearly all major malignancies routinely detected by karyotype or FISH. Some of these have a significant impact on clinical interpretation of NGS results. For example, there are multiple mechanisms of *TP53* loss, including point mutation (detectable by NGS) and gene deletion.

**Clinical trials**

The unfortunate reality for many patients is that standard of care treatments available are unlikely to result in a positive outcome. A major goal of precision medicine initiatives is to improve and expand on these treatments through the use of targeted therapies, either small molecules or biologics designed to attack cells expressing specific mutations. While there are now several such drugs available, the modest gains achieved so far are hopefully the vanguard of a much larger
and more effective suite of targeted therapies.

To assist in managing the complexities of drug development, discovery, and evaluation, the National Cancer Institute (NCI) has created the NCI Match program where patients are connected with clinical trials for which their mutation status makes them eligible. This nationwide collaboration among academic labs, pharmaceutical companies, oncologists, and patients hold great promise. From the molecular pathologist perspective, these trials offer a window into which targets may be clinically utile in the near future, pending their results. In the meantime, labs are developing or even offering testing of targets that remain in the trial phase both to prepare for therapies which may come available in the near future and to enable their patients to enroll in a clinical trial.

**Hereditary cancer syndrome variants**

There is significant overlap between the genes known to have clinically relevant somatic alterations and hereditary cancer predisposition. As such, panel testing may reveal germline variants of clinical relevance to the patient and the patient’s family. At this time, such variants are considered incidental, as patients are not typically consented or counseled that such information may be identified during testing. As screening for hereditary cancer syndromes is only recommended in high risk patients (e.g. those with family history or atypical presentation), this is a moral grey area with minimal guidance from professional societies or governmental agencies at this time.

**Conclusions**

Clinical NGS testing in cancer is a new field experiencing rapid change. As a result, the current landscape is extremely varied, with each laboratory deciding how and when to report detected variants. For now, labs must perform extensive assay performance testing and complex validations to ensure a high level of reliability, accuracy, and sensitivity.

It is also important to recognize the current limitations of NGS analysis. Due to the limited number of prospective studies and sample sizes, mutation screening has very limited negative predictive value. This means that at present NGS testing cannot discern between benign and malignant neoplasms from solid or liquid biopsy samples. Rather, as we have seen, it can help guide therapy and management for individuals known to have malignancies.

Genome-wide sequencing, including genome and exome, has enabled major advances in our understanding of the molecular basis of cancer. Some findings, such as the roles of calreticulin (CALR) mutations in myeloproliferative neoplasms and IDH1 mutations in gliomas have led to rapidly adopted, high-impact clinical tests. Thus the current common practice is to use the results of genome-wide sequencing as a resource to inform clinical test development.

While the benefit of these research efforts have been tremendous, the utility of performing the same tests on clinical specimens is less certain. At present, the total number of genes relevant to targeted therapies, diagnostic and prognostic implications, and clinical trials is likely in the hundreds. For any given cancer type, this number may be far lower: in some cases only half a dozen or so. As such, the majority of variants or mutations detected by genome-wide sequencing of neoplastic specimens cannot be interpreted in a clinical setting and are of limited clinical utility. However, several commercial and academic laboratories offer clinical exome or genome sequencing, and have reported some early success. So far though, the vast majority of relevant mutations detected by genome-wide sequencing are within genes on standard panels. For rare cancer types or types with unusually broad mutational spectra, genome-wide sequencing may already have clinical use enough above panels to merit semi-routine implementation.

Genome-wide sequencing may confer additional benefits, such as structural rearrangement detection (copy number changes, loss of heretozygosity, ploidy, etc.) and tumor purity and heterogeneity estimation. At this time, the bioinformatics required to perform these additional analyses are not mature enough for clinical diagnostics. Eventually, the benefits of genome-wide sequencing will increase (along with the decrease in the cost of sequencing) to meet a crossing point where panels are no longer sufficient. Thus it is likely a matter of "when?" - not "if?" - genome-wide sequencing replaces panel testing.

Over time, labs will coalesce around similar methods and approaches. The driving forces behind this will be increasingly sophisticated professional guidelines, more powerful reference materials, and availability of in vitro diagnosis (IVD) tests. Professional societies such as the College of American Pathologists (CAP) and the US Centers for Disease Control and Prevention (CDC) have laid the early groundwork for regulations and guidelines that will likely grow more specific and rigorous in years to come. If the pace at which commercial vendors have set continues, the rate-limiting step of mutation-positive control specimen acquisition will be reduced. And it is likely that the US FDA will continue to approve more NGS IVD tests including cancer panels.
All of these trends will combine to make clinical cancer NGS testing less costly and complex to offer, thus making them more accessible to laboratories and patients. Due to mainly economic concerns, DNA testing of solid tumors and hematological malignancies is typically limited to high level tertiary care centers and/or patients with refractory disease at this time. As the cost of testing comes down and as more targeted therapies are available, its net benefit to population health will rise dramatically.

**Conflict of interest statement**

No potential conflicts of interest are disclosed.

**References**


