



Korean Society for Genetic Diagnostics Guidelines for Validation of Next-Generation Sequencing-Based Somatic Variant Detection in Hematologic Malignancies

Heyjin Kim , M.D.^{1,2*}, Jae Won Yun , M.D.^{1*}, Seung-Tae Lee , M.D.³, Hee-Jin Kim , M.D.¹, Sun-Hee Kim , M.D.¹, Jong Won Kim , M.D.¹, and The Korean Society for Genetic Diagnostics Clinical Guidelines Committee

¹Department of Laboratory Medicine & Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea; ²Department of Laboratory Medicine, Korea Cancer Center Hospital, Seoul, Korea; ³Department of Laboratory Medicine, Yonsei University College of Medicine, Seoul, Korea

Next-generation sequencing (NGS) is currently used in the clinical setting for targeted therapies and diagnosis of hematologic malignancies. Accurate detection of somatic variants is challenging because of tumor purity, heterogeneity, and the complexity of genetic alterations, with various issues ranging from high detection design to test implementation. This article presents guidelines developed through consensus among a panel of experts from the Korean Society for Genetic Diagnostics. They are based on experiences with the validation processes of NGS-based somatic panels for hematologic malignancies, with reference to previous international recommendations. These guidelines describe basic parameters with emphasis on the design of a validation protocol for NGS-based somatic panels to be used in practice. In addition, they suggest thresholds of key metrics, including minimum coverage, mean coverage with uniformity index, and minimum variant allele frequency, for the initial diagnosis of hematologic malignancies.

Received: December 5, 2018
Revision received: March 31, 2019
Accepted: June 13, 2019

Corresponding author: Jong-Won Kim, M.D.
Department of Laboratory Medicine & Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Korea
Tel: +82-2-3410-2705
Fax: +82-2-3410-2719
E-mail: kimjw@skku.edu

*These authors contributed equally to this study.



© Korean Society for Laboratory Medicine

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Key Words: Next-generation sequencing, Validation, Somatic variant, Guidelines, Panels, Hematologic malignancies

INTRODUCTION

Next-generation sequencing (NGS) has facilitated rapid growth in the development of targeted therapies once it was adopted by research institutions and clinical laboratories to elucidate the mutational profiles of cancers [1, 2]. Several clinical trials targeting specific variants have been performed worldwide, and many new candidate genes have been suggested as specific markers for particular diseases through NGS-based tests [1, 3]. Researches using NGS on the number of somatic and epigenetic vari-

ants have increased understanding of the pathophysiology of cancers [1, 4-6]. Moreover, confirmation of gene variants is also a major criterion for the diagnosis of hematologic malignancies [7]. The importance of the accurate detection of specific gene variants has been emphasized for precision medicine in a number of cancers including hematologic malignancies [2]. The need to identify mutational profiles for use in precision medicine was sufficient for the incorporation of NGS tests into clinical laboratories; however, the setup and validation of these tests for variant detection in malignancies is challenging in clinical labo-

ratories for several reasons. First, owing to tumor heterogeneity and mutational complexity, robust validation is needed to guarantee test accuracy, especially for low variant allele frequency (VAF) variations [8]. Second, a large amount of data needs to be carefully analyzed and interpreted to meet the QC metrics thresholds. Because of the complex workflow of NGS testing, in both wet and dry laboratories, experts in clinical genetics and bioinformatics as well as technicians proficient in molecular genetic testing are necessary. We suggest practical guidelines for validating NGS-based somatic panels for the diagnosis of hematologic malignancies.

GENERAL CONSIDERATIONS FOR TEST DEVELOPMENT

The clinical purpose of the NGS-based somatic panels must first be determined. This could include molecular diagnosis, detecting therapeutic targets, or monitoring minimal residual disease (MRD). Recently, many driver mutations have been identified in mutational profiles and used to develop NGS gene panels for hematologic malignancies [3-6, 8, 9, 11]. Currently, targeted NGS-based somatic panels for hematologic malignancies may be laboratory-developed or commercially available. The genes in the designed panel should be selected after considering the clinical relevance and characteristics of the target genes. Careful consideration is required in the initial gene selection as revalidation would be necessary even if a small subset of genes is changed or added.

Platform selection

The platform used to perform NGS tests should be chosen considering all aspects, including cost, user accessibility, turnaround time, test performance, data quality, expected errors, available bioinformatics tools, and commercial gene panels of interest. Several platforms have been developed for clinical diagnosis. Currently, the two main platforms used in Korean clinical laboratories are MiSeq or NextSeq (Illumina, San Diego, CA, USA) and Ion Torrent (Thermo Fisher Scientific, Waltham, MA, USA). Illumina platforms use reversible terminator-based sequencing with optical detection of fluorescently labeled nucleotides [12]. Ion Torrent platforms use non-optical semiconductor sequencing with unmodified nucleotides [12]. These platforms use hybrid capture or amplicon-based methods in the target enrichment process. Capture-based methods usually provide even coverage of target sequences with good reproducibility but require higher amounts of DNA and a longer run time; they are usually error-

prone in GC-rich regions [12-14]. In contrast, the advantages of amplicon-based methods are that they require a shorter run time and lower amounts of DNA; however, primer dimers or non-specific amplification products can be generated [12-14]. Once the characteristics of each platform are understood, platforms that suit the actual conditions of a particular clinical laboratory should be selected.

Designing or choosing a gene panel

The panel and targeted genes should be determined based on clinical purpose and disease category. The extent of a disease category among hematologic malignancies needs to be determined; genes related to only myeloid neoplasms or all categories of hematologic malignancies are included in the panel. Targeted regions should be determined based on the locations and characteristics of clinically significant variants for diagnosis and therapeutic decisions. According to the WHO classification, *JAK2*, *CALR*, *MPL*, and *CSF3R* are crucial genes for the diagnosis of myeloproliferative neoplasms [15]. In addition to balanced translocations/inversions, gene variants, such as *NPM1*, *CEBPA*, *RUNX1*, *FLT3*, *IDH1/2*, *ASXL1*, and *KIT*, are also important diagnostic and prognostic markers in acute myeloid leukemia [11, 15]. Moreover, *MYD88* and *BRAF* should be included in the diagnosis of a specific type of lymphoma [11]. The reportable range should also be determined considering the specific characteristics of the sequence and type of variants. The areas of targeted regions below the minimum coverage need to be excluded from the reportable range and documented if the excluded regions are clinically important [16-18]. When using a commercial panel, it is necessary to verify the anticipated test performance of the target regions in each laboratory.

VALIDATION OF A GENE PANEL FOR HEMATOLOGIC MALIGNANCIES

A two-step approach is recommended for validation (Fig. 1). The first step of validation (Step 1), known as pilot tests, is necessary for optimization and checking for possible errors during the entire NGS testing process. After confirming >95% concordance rate with known variants and meeting the QC metrics thresholds, the second step of validation (Step 2) is conducted with established thresholds of essential parameters, such as depth of coverage and VAF, for each type of variant. When using commercial panels, laboratories can perform ongoing validation instead of Step 2. Ongoing validation is described in a separate section.

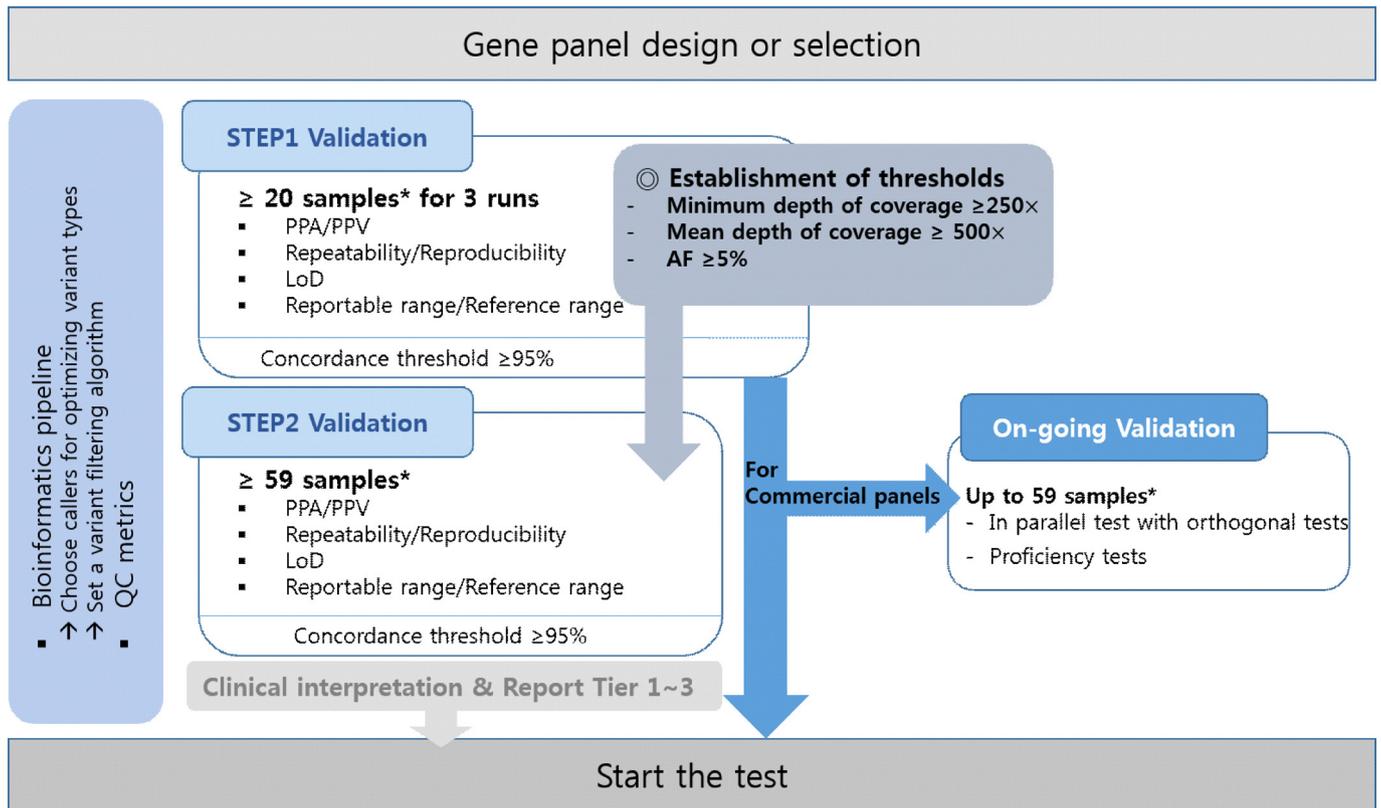


Fig. 1. Overview of validation process for somatic variants in hematologic malignancies using NGS testing.

*Samples include patient samples, validated cell lines, and/or commercial controls.

Abbreviations: NGS, next-generation sequencing; LoD, limit of detection; PPA, positive percentage agreement; PPV, positive predictive value; AF, allele frequency.

General considerations for validation

Samples

Sample types should be determined prior to validation; each sample type to be used in practice should be included in the validation process. Whole blood (WB) or bone marrow (BM) samples are the most commonly used for hematologic malignancies. However, other samples, including formalin-fixed paraffin embedded (FFPE) tissues, various body fluids, cell-free DNA, and skin tissue (skin fibroblasts for germline DNA), can also be used [19]. To examine a particular type of sample, the validation step should be performed for multiple samples of the same type. Adequate purity, volume, and proper storage conditions (e.g., time and temperature) of each sample are necessary for optimal testing. Approximate VAFs of known variants (from previous Sanger testing or NGS testing) in each sample would be useful for validation design. Fresh WB and BM samples are usually considered best for practical NGS testing because of the relatively high quality and quantity of DNA and RNA from neoplastic cells. We discuss the type of samples, mainly focusing on WB and BM. In several cases, reference materials (RMs) or

commercial controls can also be included during validation. Pooling of up to three samples with different variants is also viable and can be regarded as three samples. If the samples originated from different patients, those with the same variant can be used in <10% of the total samples for validation.

Type of variants

Frequent variants with clinical significance should be included as positive samples during validation. In a panel of myeloid malignancies, c.1849G>T (p.V617F) in the *JAK2* gene or an internal tandem duplication in the *FLT3* gene are included in the validation with high priority [20]. The variant type should be determined before implementing validation. We mainly describe single-nucleotide variants (SNVs) and insertions and/or deletions (indels), as the accurate detection of the other types of variants is challenging in clinical diagnosis.

Number of genes in the panel required for validation

There is no consensus regarding how many genes should be included in the panel undergoing validation. Validating many re-

gions would be good for reliability; however, it is nearly impossible in many clinical laboratories. Clinically relevant variants should be included as a priority, as previously described in the section of designing or choosing a gene panel. Samples with two or more known variants, commercial controls, or cancer cell lines could be efficiently used to reduce the number of samples required for validation [16, 17].

Step 1 validation: pilot test of a custom panel or verification of a commercial panel

Step 1 (pilot tests or verification) can provide an overview of initial validation before Step 2. To optimize a test or to verify the performance of a commercial panel, the entire process should be evaluated by Step 1 validation [13, 16, 21]. Unexpected problems are often identified and corrected during this step. We recommend that at least 20 samples and at least three runs be included and performed to evaluate the NGS testing performance parameters, such as accuracy, precision, and limit of detection (LoD), for each type of sample and variant (Fig. 2). Mixing a number of samples with known variant burden is a means of validat-

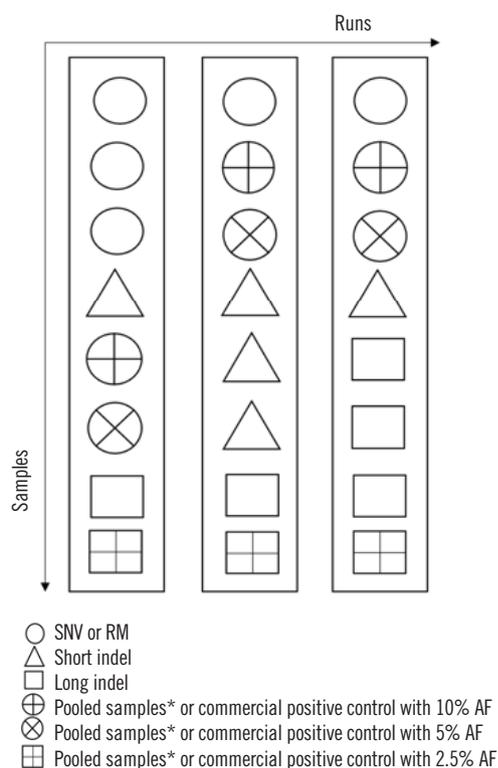


Fig. 2. An example of Step 1 validation (pilot test).

*Pooled samples can comprise one short deletion, one short insertion, and one long insertion in different regions of the patient samples. Abbreviations: AF, allele frequency; RM, reference material; SNV, single-nucleotide variant; indel, insertion and/or deletion.

ing multiple variations with fewer samples. Desired VAF thresholds, depth of coverage for each position, and mean depth of coverage should be established in Step 1 validation at given QC metrics thresholds [3, 9]. We recommend that 5% VAF, a minimum coverage of 250 reads for each position, and a mean coverage of 500 reads should be the threshold for a gene panel for hematologic malignancies [13, 16]. These analytical goals applicable to Steps 1 and 2 (including ongoing validation) require validation with minimum 20 and 59 samples, respectively. Different VAF thresholds can be adopted for different variants. Uncovering common variants, such as an indel in the *CEBPA* gene, may be challenging with a low depth of coverage and low VAF in accordance with highly GC-rich regions. In this case, systematic errors should be discriminated and documented, which may be reviewed when designing a panel [14, 16]. A concordance rate >95% should be met for the detection of known variants through NGS and Sanger sequencing before Step 2 or before beginning to test patients using the panels.

NGS performance characteristics

Positive percentage agreement (PPA) and positive predictive value (PPV)

The accuracy parameters, such as PPA and PPV, for each type of variant should be established during validation [1]. PPA is the ability of NGS tests to obtain positive results measured in concordance with positive results obtained by an orthogonal test (e.g., Sanger sequencing). PPV is the proportion of the number of positive NGS test results that have the target condition, as determined by the orthogonal test. RMs could constitute a good source of PPA/PPV, providing true positive variants [13, 16]. In addition to RM-based PPA/PPV, PPA/PPV can be derived from the results of known true variants, including hotspots and non-hotspots, from patient samples. The formulas for calculating PPA and PPV are as follows:

$$\text{PPA (\%)} = 100 \times \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

$$\text{PPV (\%)} = 100 \times \frac{\text{True positives}}{\text{True positives} + \text{False positives}}$$

Precision (repeatability and reproducibility)

Repeatability (within-run precision) and reproducibility (between-run precision) should be evaluated. We recommend that at least three samples be used for each type of variant and triplicates for both within-run and between-run precisions should be arranged during validation design [16, 21, 22]. Using RMs or commercial positive controls for each type of variant may serve as an alternative option.

LoD

The LoD should be evaluated for each type of variant through dilution studies of pure patient samples with an RM at variable percentages (25%, 10%, 5%, 2.5%, or 1%) based on which the minimum VAF can be determined according to the purpose of the panel [17]. We recommend a 5% VAF for each type of variant as adequate for the diagnosis of hematologic malignancies; however, 1% VAF should be used for the MRD panel [23]. As it might be difficult to obtain 5% VAF in certain circumstances, such as long indels, GC-rich regions, and repetitive regions, if these occur in a clinically significant region, they should be documented in the clinical report. When using patient samples, at least three samples for each type of variant should be tested in three independent experiments [24]. Serial dilution of pure patient samples with RM, pooled patients' samples, cancer cell lines, and/or commercial controls comprising all types of variants at a specific VAF can be used for LoD determination [13, 17, 21].

Reportable range and reference range

The reportable range can be determined based on two factors. First, the range should meet the QC metrics thresholds. Second, the range is determined considering target regions with clinical significance. Limitations should be described for specific regions showing lower than minimal depth of coverage. The reference range can be described as the range of normal sequence variation occurring in the general population [17]. The reference range should be included in the report detailing what types of variants were reported. Recent guidelines for somatic variant interpretation and reporting classify the variants into four tiers [20]. Based on these guidelines, we recommend that tier 1–3 variants should be reported in hematologic malignancies. Additionally, the criteria for confirming detected variants using orthogonal tests should be included in the report. Although providing relatively low to medium sensitivity compared with NGS, confirmatory tests, such as Sanger sequencing, pyrosequencing, quantitative PCR, and multiplex ligation-dependent probe amplification, according to variant type, can be described in the report for clinicians.

Ongoing validation

Ongoing validation would be applicable for commercial NGS-based somatic panels that meet the QC metrics in Step 1. Positive samples based on a panel testing service could be used for other orthogonal methods. We do not recommend parallel testing using samples with negative results. It is difficult to confirm

true negative results by Sanger sequencing, which has lower analytical sensitivity. Although a previous study has reported that orthogonal validation is not required [25], confirmation between NGS and Sanger sequencing or other orthogonal methods regarding positive results in a specific gene region should be performed using >59 samples, which represent an adequate proportion of each type of variant. For practical reasons, pooled samples with different variants from different patients are also available (see the sample section in general considerations for validation). Additionally, laboratories using commercial panels should participate in proficiency tests (PT) with reference laboratories. Reference laboratories would qualify by undergoing external quality assessment such as domestic and international PT programs for the relevant disease panels. Reference laboratories should be able to provide positive samples, which would help reduce the burden of collecting positive samples and the cost of validation and/or quality controls, especially at relatively small institutions.

Step 2 validation

Basic parameters, including PPA/PPV, precision (repeatability and reproducibility), LoD, and reference range/reportable range should be validated again (see the performance characteristics section in Step 1) with the desired VAF thresholds, depth of coverage for each position, and mean depth of coverage established in Step 1 at given QC metrics thresholds. Although these parameters are preliminarily validated in Step 1, this step would strengthen the validation with additional samples. Based on previous recommendations of validation for somatic variants, in the next step of validation, a minimum of 59 samples is suggested as an adequate number of samples (after statistical calculation), as it represents <5% analytical sensitivity with 95% confidence and $\leq 1.9\%$ false-positive rate [16]. Generally, validation needs to be performed using as many samples as known variants among patients. However, it is difficult to collect samples that have clinically relevant variants for certain gene panels. For practical reasons, pooled samples with different variants from different patients are also available (see the sample section under general considerations for validation). An example for the design of Step 2 validation is shown in Fig. 3.

Considerations for bioinformatics pipelines

Bioinformatics pipelines or data analysis pipelines include the following steps: read alignment, variant calling, variant annotation and reporting, and generation of QC matrices. The general rules are as follows: for test performance, different computa-

tional approaches and validation processes are needed for different classes of sequence variants, namely SNVs, indels, copy number variations (CNVs), and structural variations (SVs). Additionally, software tools should be selected considering the design, purpose, and characteristics of the test. A review of various software or pipelines is beyond the scope of these guidelines. General considerations for selecting software and validation are discussed in this section.

SNVs

For SNV detection, it is necessary to select software specific for somatic SNVs. The software algorithms for constitutional genome analysis can miss variants with VAFs falling outside the expected range for heterozygous and homozygous variants [26]. In cases where low VAF is expected, software optimized for somatic variants is recommended, as it would be optimized for cancer samples with varying levels of tumor purity and heterogeneity [27].

Indels

Indel variants have various size and sequence complexities; thus, accurate alignment, calling, and annotation are technically challenging. During the alignment step, algorithms, including local realignment, should be considered to minimize base pair mismatches. Indels <20 bp can be accurately called by algorithms using probabilistic modeling [16]. However, for accurate

detection of medium to large indels (e.g., FLT3-ITD insertion), additional indel-calling algorithms, such as split-read analysis, are required. Various SV callers using split-read analysis are available for detecting long indels; however, most are not fully validated for high-coverage NGS-based somatic panels. Thus, validation of long indel detection with clinical samples is an additional requirement. Furthermore, software characteristics must be considered before installation; for example, some SV callers, such as DELLY and LUMPY, are ideal for detecting duplications or insertions, while Pindel and LUMPY seem to be inferior to SvABA for detecting deletions <300 bp [28].

CNV

Somatic CNV detection in cancer samples is challenging. This limitation makes CNV testing an optional or supportive method. The calling algorithm for CNV is different from those for SNVs or indels in that it is generally inferred based on the read depth data. The most challenging issues for accurate CNV calling are as follows: first, tumor purity and heterogeneity must be taken into account to solve the dilution of CNV signals [27]. However, it is difficult to calculate the purity values, especially in panel sequencing data. Second, batch effect and read-depth normalization should be handled carefully. In contrast, if the CNV calling algorithm is optimized for WB, BM, or fresh tissue samples, it should generally not be applied with additional optimization al-

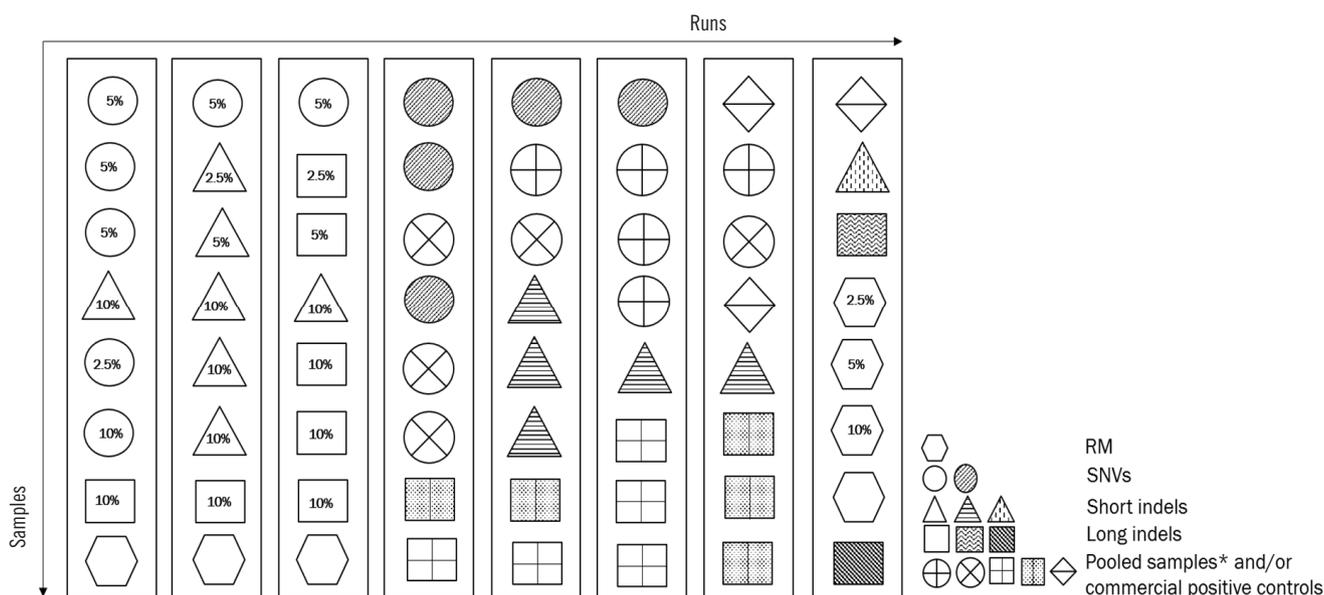


Fig. 3. An example of Step 2 validation (formal validation) with 64 samples (including separately bar-coded samples from the same patient/RMs/commercial positive controls). For LoD validation, samples with various VAF were guaranteed by dilution with RM.

*Pooled samples can be comprised of one short deletion, one short insertion, and one long insertion in different regions of the patient samples. Abbreviations: indels, insertions and/or deletions; RM, reference material; VAF, variant allele frequency; SNV, single nucleotide variant; LoD, limit of detection.

gorithms for FFPE samples. DNA from FFPE could have different characteristics because of DNA degradation [29], which could cause biased CNV results (both false positive or negative signals).

SVs

There are several limitations for calling SVs. The SV breakpoints are mostly located in non-coding DNA regions, introns, or highly repetitive regions [27]. Therefore, the target regions tend to be too broad and significantly less uniform. These limitations are related to reduced test efficiency and accuracy. In addition, although several tools for SV calling have been introduced, many have yet to be validated or optimized for high-coverage panel sequencing. Validation and parameter optimization are required for each panel; without rigorous validation, including LoD, SV analysis should be used as supportive or optional information.

RNA SVs

RNA SV analyses using NGS have a different test scope from conventional RNA tests; novel fusions or extremely rare fusions that are important for diagnosis and treatment can be detected and analyzed [30]. Many tools for chimeric transcript detection have been introduced. However, each tool has different performance and algorithm [31, 32]. After reviewing the characteristics of each software, the performance should be validated with true positive and negative samples. To improve the sensitivity and specificity, several software are used for calling, and the results can be integrated and analyzed considering the sensitivity and specificity of the software [33, 34].

QC metrics

Bioinformatics QC metrics include base and mapping quality scores, on-target reads, duplicated read rate, uniformity of base coverage, mean depth of the on-target regions, target base coverage $<250\times$, numbers and types of variants from the reference, and the transition:transversion ratio in the exome and the genome. These parameters may be modified based on the platform and bioinformatics algorithms; other metrics can also be added. Acceptance criteria are needed for the metrics; most importantly, it is needed to evaluate whether essential target regions, such as mutational hotspots, are fully covered with $>250\times$ coverage (Table 1).

DISCLAIMERS

Disclaimers after validation can be divided into three categories:

Table 1. Example of QC metrics

Description	Criteria for acceptance
Hotspot exons not fully covered with $>250\times$	<5 in 150 hotspot exons
On-target reads (%) = $\frac{\text{On-target reads}}{\text{Total aligned reads}} \times 100$	$>90\%$
Duplicated read rate	$<50\%$
Uniformity of base coverage (%): the proportion of sequences that have >0.2 -fold the mean coverage	$>90\%$
Mean depth of the on-target regions	$>500\times$
Target base coverage $>250\times$	$>90\%$ ($250\times$)

type of sample, specific regions in the panel with technical issues, and interpretations. First, the type(s) of sample(s) used for validation should be described in the validation and the test report document. Additionally, poor conditions related to DNA quality issues should be reported when the tests cannot be performed or when the test results do not meet the established QC parameters. Moreover, high GC content regions, pseudogenes, and repetitive sequences should be documented in the report [13, 16, 22]. Lastly, analytical sensitivity, such as LoD and reportable ranges, should be specified in the report. The policy for reporting incidental findings of potential pathogenic germline variants should also be described in the disclaimers [1, 35].

CONCLUSIONS

Recently, NGS has been widely used in clinical laboratories; targeted gene panels for hematologic malignancies are being adopted for various purposes. Workflow complexity and the extensive cost and number of samples are main hurdles in the application of somatic panels in the clinical setting. The final goal of NGS testing is reporting reliable results over a minimum quality threshold without Sanger sequencing validation. With the rapid development of technology and bioinformatics tools, we expect this demanding workload of NGS testing validation to be simplified in the near future. The issues related to incidental findings and types of variants other than SNVs and indels need to be discussed in future guidelines. The present guidelines provide general considerations in setting up and validating clinical NGS-based somatic panels in hematologic malignancies. Thus, detailed issues regarding target enrichment method, sequencing platform, and bioinformatics software are beyond our scope and need to be reviewed in a separate paper.

Authors' Disclosures of Potential Conflicts of Interest

No potential conflicts of interest relevant to this article were reported.

Acknowledgements

This work was supported by a grant for the development of clinical diagnostic guidelines/recommendations/statement from the Korean Society for Genetic Diagnostics, 2018.

ORCID

Heyjin Kim	https://orcid.org/0000-0002-2213-2579
Jae Won Yun	https://orcid.org/0000-0002-9029-8036
Seung-Tae Lee	https://orcid.org/0000-0003-1047-1415
Hee-Jin Kim	https://orcid.org/0000-0003-3741-4613
Sun-Hee Kim	https://orcid.org/0000-0002-7542-5551
Jong Won Kim	https://orcid.org/0000-0002-0708-9242

REFERENCES

1. Watson IR, Takahashi K, Futreal PA, Chin L. Emerging patterns of somatic mutations in cancer. *Nat Rev Genet* 2013;14:703-18.
2. Coombs CC, Tallman MS, Levine RL. Molecular therapy for acute myeloid leukaemia. *Nat Rev Clin Oncol* 2016;13:305-18.
3. Galanina N, Bejar R, Choi M, Goodman A, Wieduwilt M, Mulroney C, et al. Comprehensive genomic profiling reveals diverse but actionable molecular portfolios across hematologic malignancies: implications for next generation clinical trials. *Cancers (Basel)* 2018;11.
4. Vainchenker W and Kralovics R. Genetic basis and molecular pathophysiology of classical myeloproliferative neoplasms. *Blood* 2017;129:667-79.
5. Grinfeld J, Nangalia J, Green AR. Molecular determinants of pathogenesis and clinical phenotype in myeloproliferative neoplasms. *Haematologica* 2017;102:7-17.
6. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 2012;481:506-10.
7. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 2016;127:2391-405.
8. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multi-region sequencing. *N Engl J Med* 2012;366:883-92.
9. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med* 2016;374:2209-21.
10. Goodman AM, Choi M, Wieduwilt M, Mulroney C, Costello C, Frampton G, et al. Next generation sequencing reveals potentially actionable alterations in the majority of patients with lymphoid malignancies. *JCO Precis Oncol* 2017;1:1-13.
11. Taylor J, Xiao W, Abdel-Wahab O. Diagnosis and classification of hematologic malignancies on the basis of genetics. *Blood* 2017;130:410-23.
12. Ballester LY, Luthra R, Kanagal-Shamanna R, Singh RR. Advances in clinical next-generation sequencing: target enrichment and sequencing technologies. *Expert Rev Mol Diagn* 2016;16:357-72.
13. Kanagal-Shamanna R, Singh RR, Routbort MJ, Patel KP, Medeiros LJ, Luthra R. Principles of analytical validation of next-generation sequencing based mutational analysis for hematologic neoplasms in a CLIA-certified laboratory. *Expert Rev Mol Diagn* 2016;16:461-72.
14. Yohe S and Thyagarajan B. Review of clinical next-generation sequencing. *Arch Pathol Lab Med* 2017;141:1544-57.
15. Swerdlow SH, Campo E, et al. WHO classification of tumours of haematopoietic and lymphoid tissues. revised 4th ed. Lyon: International Agency for Research on Cancer, 2017:16-27.
16. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer 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. *J Mol Diagn* 2017;19:341-65.
17. Gargis AS, Kalman L, Berry MW, Bick DP, Dimmock DP, Hambuch T, et al. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat Biotechnol* 2012;30:1033-6.
18. Matthijs G, Souche E, Alders M, Corveleyn A, Eck S, Feenstra I, et al. Guidelines for diagnostic next-generation sequencing. *Eur J Hum Genet* 2016;24:1515.
19. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 2017;129:424-47.
20. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer: A joint consensus recommendation of the association for molecular pathology, American Society of Clinical Oncology, and College of American pathologists. *J Mol Diagn* 2017;19:4-23.
21. Ellard S, Lindsay H, Camm N, Watson C, Abbs S, Mattocks C, et al. Practice guidelines for targeted next generation sequencing analysis and interpretation. Association for clinical genetic science (ACGS). https://www.acgs.uk.com/media/10789/bpg_for_targeted_next_generation_sequencing_-_approved_dec_2015.pdf.
22. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med* 2013;15:733-47.
23. Ben Lassoued A, Nivaggioni V, Gabert J. Minimal residual disease testing in hematologic malignancies and solid cancer. *Expert Rev Mol Diagn* 2014;14:699-712.
24. Cuomo AW, Zucker HA, Dreslin S. Oncology-molecular and cellular tumor markers "Next Generation" sequencing (NGS) guidelines for somatic genetic variant detection. <https://www.wadsworth.org/regulatory/lep/clinical-labs/obtain-permit/test-approval> (Updated on Jan 2018).
25. Beck TF, Mullikin JC, NISC Comparative Sequencing Program, Biesecker LG. Systematic evaluation of Sanger validation of next-generation sequencing variants. *Clin Chem* 2016;62:647-54.
26. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013;43:11.10.1-33.
27. Shin HT, Choi YL, Yun JW, Kim NKD, Kim SY, Jeon HJ, et al. Prevalence and detection of low-allele-fraction variants in clinical cancer samples. *Nat Commun* 2017;8:1377.
28. Wala JA, Bandopadhyay P, Greenwald NF, O'Rourke R, Sharpe T, Ste-

- wart C, et al. SvABA: genome-wide detection of structural variants and indels by local assembly. *Genome Res* 2018;28:581-91.
29. Scheinin I, Sie D, Bengtsson H, van de Wiel MA, Olshen AB, van Thuijl HF, et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res* 2014;24:2022-32.
 30. Levin JZ, Berger MF, Adiconis X, Rogov P, Melnikov A, Fennell T, et al. Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts. *Genome Biol* 2009;10:R115.
 31. Kumar S, Razaq SK, Vo AD, Gautam M, Li H. Identifying fusion transcripts using next generation sequencing. *Wiley Interdiscip Rev RNA* 2016;7:811-23.
 32. Liu S, Tsai WH, Ding Y, Chen R, Fang Z, Huo Z, et al. Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. *Nucleic Acids Res* 2016;44:e47.
 33. Zeng X, Lin W, Guo M, Zou Q. A comprehensive overview and evaluation of circular RNA detection tools. *PLoS Comput Biol* 2017;13:e1005420.
 34. Liu S, Tsai WH, Ding Y, Chen R, Fang Z, Huo Z, et al. Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. *Nucleic Acids Res* 2016;44:e47.
 35. Singh RR, Luthra R, Routbort MJ, Patel KP, Medeiros LJ. Implementation of next generation sequencing in clinical molecular diagnostic laboratories: advantages, challenges and potential. *Expert Rev Precis Med Drug Dev* 2016;1:109-20.