









## GUIDELINES

# Recommendations related to the analytical equivalence assessment of gene panel testing

Sumimasa Nagai<sup>1</sup>  | Hiroshi Nishihara<sup>2</sup> | Takayoshi Suzuki<sup>3</sup>  | Kazuto Nishio<sup>4</sup>  |  
 Hiroya Taniguchi<sup>5</sup>  | Katsuya Tsuchihara<sup>6</sup>  | Kohei Nakamura<sup>2</sup>  | Reika Takamatsu<sup>2</sup> |  
 Toshihide Ueno<sup>7</sup> | Hiroyuki Aburatani<sup>8</sup> | Takashi Kohno<sup>9</sup>  | Shinji Kohsaka<sup>7</sup> 

<sup>1</sup>Department of Medical Development, Institute for Advancement of Clinical and Translational Science, Kyoto University Hospital, Kyoto, Japan

<sup>2</sup>Genomics Unit, Keio Cancer Center, Keio University School of Medicine, Tokyo, Japan

<sup>3</sup>Division of Molecular Target and Gene Therapy Products, National Institute of Health Sciences, Kawasaki, Japan

<sup>4</sup>Department of Genome Biology, Kindai University Faculty of Medicine, Osakasayama, Japan

<sup>5</sup>Department of Clinical Oncology, Aichi Cancer Center Hospital, Nagoya, Japan

<sup>6</sup>Division of Translational Informatics, Exploratory Oncology Research and Clinical Trial Center, National Cancer Center, Kashiwa, Japan

<sup>7</sup>Division of Cellular Signaling, National Cancer Center Research Institute, Tokyo, Japan

<sup>8</sup>Genome Science and Medicine Laboratory, Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan

<sup>9</sup>Division of Genome Biology, National Cancer Center Research Institute, Tokyo, Japan

## Correspondence

Sumimasa Nagai, Department of Medical Development, Institute for Advancement of Clinical and Translational Science, Kyoto University Hospital, 54 Shogoin-Kawaharacho, Sakyo-ku, Kyoto-shi, Kyoto 606-8507, Japan.

Emails: [sunagai@kuhp.kyoto-u.ac.jp](mailto:sunagai@kuhp.kyoto-u.ac.jp); [sunagai-ky@umin.ac.jp](mailto:sunagai-ky@umin.ac.jp)

## Funding information

Japan Society for the Promotion of Science, Grant/Award Number: 18K06779; Japan Agency for Medical Research and Development, Grant/Award Number: JP21mk0102145

## Abstract

Advances in cancer genome care over the past few years have included the development of gene panel testing for various biomarkers. This article summarizes issues and provides recommendations related to analytical performance evaluations for new oncology gene panels. The scope of these recommendations includes comprehensive genomic profiling assays related to gene panel testing that uses histological or serum specimens to detect gene mutations. As a research project of the Japan Agency for Medical Research and Development Research on Regulatory Science of Pharmaceuticals and Medical Devices, we convened the working group committee that consisted of more than 30 experts from academia, industry, and government. We have discussed the points that should be considered to allow maximal simplification of assessments using clinical specimens in evaluating accuracy and limit of detection in equivalence and analytical performance for 3 years. We provide recommendations specific to each type of gene mutation as well as to reference standards or specimens used for evaluations. In addition, in order to facilitate the discussion on the analytical performance of gene panel tests by multidisciplinary tumor boards of hospitals, the present recommendations also describe the items that companies are expected to

These guidelines are the English translation of the original guidelines in Japanese issued by a working group of the Japan Agency for Medical Research and Development Research on Regulatory Science of Pharmaceuticals and Medical Devices "Investigation for Standardization of Companion Diagnostic System by Gene Panel Testing" (Principal Investigator: Sumimasa Nagai), which are available at websites of the Japanese Society of Medical Oncology (<https://www.jsmo.or.jp/news/jsmo/20220307.html>), Japan Society of Clinical Oncology (<http://www.jsco.or.jp/jpn/index/page/id/2571>), and the Japanese Cancer Association (<https://www.jca.gr.jp/researcher/topics/2022/20220307.html>).

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Cancer Science* published by John Wiley & Sons Australia, Ltd on behalf of Japanese Cancer Association.

provide information on in their packaging inserts and reports, and the items that are expected to be discussed by multidisciplinary tumor boards. Our working group document will be important for participants in multidisciplinary tumor boards, including medical oncologists and genome scientists, and developers of gene panels not only in Japan but also in other countries.

#### KEYWORDS

companion diagnostic, comprehensive genomic profiling, liquid biopsy, next-generation sequencing, pharmaceuticals and medical devices agency

## 1 | BACKGROUND

FoundationOne CDx and the OncoGuide NCC Oncopanel System were approved in December 2018 and have been covered by national health insurance in Japan since June 2019. Since then, cancer genome care, including companion diagnostics (CDx) by gene panel testing and comprehensive genomic profiling (CGP), has been advancing steadily. In addition to the gene panels that use tumor tissue specimens, FoundationOne Liquid CDx was approved in March 2021 and has been covered by national health insurance since August 2021 as a gene panel that uses plasma specimens.<sup>1</sup>

When new CDx are developed for biomarkers for which CDx approved in Japan already exist, equivalence or accuracy testing using clinical or synthetic specimens is needed to evaluate analytical equivalence with approved CDx or established standard testing method.<sup>2,3</sup> However, in reality, the feasibility of assessing analytical equivalence of new CDx is limited due to the fact that clinical specimens are valuable, consensus on the methods of utilizing synthetic specimens has yet to be established, and obtaining measurement results of approved CDx as a control method is sometimes difficult.

Furthermore, when CGP reveals a gene mutation for which a CDx exists and a multidisciplinary tumor board, called an expert panel, of a hospital recommends administration of a pharmaceutical drug related to the genetic mutation, administering the pharmaceutical drug without performing another test by CDx is considered acceptable.<sup>4</sup> However, published documents and information on evaluations of analytical equivalence between approved CGPs or between an approved CGP and a CDx, which assist expert panel discussions, are currently limited.

Given the above, we have been investigating the standardization of methods of analytical performance assessment of gene panel testing by the Japan Agency for Medical Research and Development (AMED) Research on Regulatory Science of Pharmaceuticals and Medical Devices "Investigation for Standardization of Companion Diagnostic System by Gene Panel Testing" (Principal Investigator: Sumimasa Nagai).<sup>1</sup> Whether to use clinical specimens is a central point of debate in evaluating accuracy and limit of detection (LOD) in equivalence and analytical performance assessments. Moreover, these assessments are particularly important subjects of discussion in expert panels on CGP. The points that should be considered to allow maximal simplification of assessments using clinical specimens

are thus summarized in the present recommendations. Conducting assessments of analytical performance using identical reference standards is beneficial in that it allows external comparisons between panels. This is implemented not only for CDx assessments but also for assessment of analytical performance as a CGP, which is one of the most important parameters selected through the discussions in this working group meeting.

However, in ensuring that measurement with clinical specimens is possible, it is not appropriate even in CGP to make an application for approval of a gene panel by concluding all analytical performance assessments with reference standards alone. Clinical specimen eligibility testing, which evaluates whether assessment is possible with clinical specimens derived from multiple tissues, should be conducted. In addition, validation of the test using clinical specimens (not limited to specimens derived from clinical trials), that is, concordance rate assessment with some control method, is necessary for an application for approval at present. Whether or not such validation of clinical specimens should continue to be required from a scientific point of view is described in the present recommendations. Furthermore, although it is not in the scope of discussion in the present recommendations, continuing to exert efforts toward precision management through postmarketing external evaluations is desirable.

In addition, in order to facilitate the discussion on the analytical performance of gene panel tests by the expert panel, the present recommendations also describe the items that companies are expected to provide information on in their packaging inserts and reports, and the items that are expected to be discussed by the expert panel.

## 2 | SCOPE

The present recommendations concern applications for approval of CGP related to gene panel testing that uses tissue or plasma specimens to detect gene mutations. These mutations include, among others, single or multiple nucleotide variants (SNV/MNV), insertions or deletions (Ins/Del), copy number variations (CNV), or fusion genes. However, concerning gene panel testing using cell-free DNA extracted from plasma specimens, only the items mentioned in the main text are applicable.

At present, these recommendations mostly concern gene panel testing using specimens from patients with solid cancers but may be

used as a reference for gene panel testing of hematopoietic malignancies as well. The recommendations were based on discussions of current knowledge and are subject to change according to new scientific findings or developments. Furthermore, the Pharmaceuticals and Medical Devices Agency should be consulted for the development of specific new products based on the ideas mentioned in the present recommendations as necessary.

### 3 | SINGLE OR MULTIPLE NUCLEOTIDE VARIANTS, INS/DEL

#### 3.1 | General considerations

The reference standards of Horizon, AcroMetrix, and other companies are available, and the Japanese Promotion Council for Laboratory Testing “Fundamental Premises on Quality and Precision Assurance of Oncogene Panel Testing (Ver. 2.0)”<sup>5</sup> and Medical Device Innovation Consortium “SRS Report: Somatic Variant Reference Samples for NGS”<sup>6</sup> are good references. Among these, in the working group meeting, the Horizon HD701, HD827, and HD Tru-Q were used for investigation and revealed that the variant allele frequency (VAF) published by the distributor were generally found to be appropriate (Figures S1 and S2). We also confirmed that mutations in homopolymeric regions and regions with high GC content are included on the list of detectable mutations published by the distributor of reference standards are indeed detectable and that the products are therefore usable (Figure S1). The use of data published in papers or published by the distributors of reference standards is possible and contributes to expert panel discussions.

Notably, for reference standards such as HD827 that contain multiple mutations in the same reads, many mutations may be eliminated by oncogene panel software because the software filters may eliminate the mutations as alignment errors when there are many mutations in close proximity.

It is acceptable to indicate LOD in terms of VAF and hit rate (a value obtained by dividing the number of times mutations were detected by the number of valid measurements) and describe them in package inserts. Lowering the cut-off value to assess LOD by hit rate only will produce good hit rates but increase the error rate. As such, it is important to take into account the results of measurement on background noise in addition to analytical specificity and error rate in making the decision. Information on LOD assessment methods as well as “Detects SNV with VAF  $\geq$  XX%” as precautions for use should be provided in package inserts. In measuring background noise, in addition to measurements in normal samples, measurement targeting nonpositive genetic mutations in commercially available reference standards is considered. Furthermore, understanding the strengths and weaknesses of various methods of detecting gene mutations in assessing LOD is important. For example, VAF of relatively large Ins/Del tends to be underestimated by next-generation sequencing (NGS; capture method) relative to that by digital PCR.

Regarding the range of detectable SNV/MNV and Ins/Del mutations, “New Regulations for Companion Diagnostic Systems (Draft)”<sup>2</sup> recommends as a general rule that, in the case of variations in a given mutation, positive specimens should cover  $\geq$ 90% of mutations reported in the target population (on the basis of detection frequency). Relatively large Ins/Del (e.g., EGFR V769\_D770insASV or EGFR  $\Delta$ E746-A750), such as Ins/Del in homopolymeric regions, mutations in regions with high GC content (e.g., GNA11 Q209L or AKT1 E17K), and pseudogenes require more careful analytical performance assessment (Table S1). As described above, assuming that most mutations reported in the target population are covered in bait and primer design, grouping them into mutations that require more careful analytical performance assessment and those that do not makes possible the assessment of the detection target of the gene panel by extrapolating from mutations that can be measured with reference standards.

It is important to note that, although MNV mutations are rare, they could be difficult to evaluate with some software, or corresponding reference standards might not be available.

Exon skipping mutations may be treated as SNV/MNV or Ins/Del in the DNA panel. However, they may be contained in reference standards for fusion in RNA panels; thus, it is appropriate to handle them within the scope of fusion.

With regards to reference standards, the aforementioned Medical Device Innovation Consortium is about to launch the large-scale creation of new reference standards with cell lines in which tumor and normal cells are paired, and the FDA's Sequencing and Quality Control Consortium is also attempting to create reference standards that cover low VAF.<sup>7</sup> As these examples demonstrate, new reference standards could become available in the future.

#### 3.2 | Summary

##### 3.2.1 | If aiming for approval as a comprehensive genomic profiling assay

It is acceptable to calculate LOD using commercially available reference standards. These LOD will most likely be indicated in terms of VAF and hit rate, similar to LOD given on package inserts of approved products. Furthermore, it is also acceptable to evaluate accuracy testing through comparisons with reference values of reference standards (e.g., VAF measured by digital PCR or published as part of quality control testing by the distributor).

##### 3.2.2 | If pharmaceutical drugs are to be administered based on results of comprehensive genomic profiling without confirmation by companion diagnostics after expert panel discussion

Based on the results of accuracy testing and LOD assessment at the time of application for regulatory approval as a CGP, an expert

panel discussion of the equivalence with approved CDx should be considered. For gene mutations not covered by reference standards used for accuracy testing and LOD assessment, the gene mutations should be grouped according to their requirements for more careful analytical performance assessment (e.g., relatively large Ins/Del), and a per-group analysis should be carried out. The analytical performance for gene mutations not covered by the reference standards should be investigated based on analytical performance results for gene mutations that are covered by the reference standards.

### 3.2.3 | Future outlook

From a scientific perspective, for application for CDx approval, gene mutations covered by reference standards can be assessed through accuracy testing by comparing them to reference values of reference standards as a substitute for equivalence testing with approved CDx.

For gene mutations not covered by reference standards, the mutations should be grouped according to their requirements for more careful analytical performance assessment (e.g., relatively large Ins/Del), and a per-group analysis should be undertaken. For application for approval as a CDx, if it is possible to consider the analytical performance for gene mutations not covered by reference standards based on the results of analytical performance for mutations that are, omitting equivalence testing and substituting accuracy testing with reference values is possible.

## 4 | COPY NUMBER VARIATIONS

### 4.1 | General considerations

Our working group discussed *MYC-N* and *MET* copy number analysis by Horizon HD753 (precision managed by droplet digital PCR) and *EGFR* and *MET* copy number analysis (digital PCR results are available for use as the control), which used NIST Reference Material 8366.<sup>8</sup> We confirmed that no problems with the copy numbers published by the distributor existed; thus, the data can be used for analytical performance assessment of oncogene panels (Figure S3). NIST Reference Material 8366 was created from six types of human cell lines, and ratios with reference genes measured by digital PCR are also provided. Of the six cell line types, Hs746T has a *MET* copy number of 33.4 according to digital PCR; however, the estimated copy number of *MET* according to the oncogene panel using NGS (capture method) is lower (Figure S3). Notably, the upper copy number limit estimated by oncogene panels using the capture method is generally lower than that estimated by PCR for such reference genes with high copy numbers because of probe depletion. As for reference standards for analyzing *ERBB2* copy numbers, NIST Reference Material 2373 genomic DNA (results of digital and quantitative PCR are available as controls)<sup>9</sup> is available for import and use. During discussions

in our working group meeting, we confirmed that generally no problems with the copy numbers published by the distributor exist and that they thus can be used for analytical performance assessment of oncogene panels (Figure S4).

It is acceptable to indicate the LOD of CNV in terms of copy number or percentage of tumor and hit rate. In approved products, whether the percentage of tumor is based on the count in the pathological specimen or calculated from VAF is unclear; however, this is unavoidable because indicating the percentage of tumor as a validated value is difficult. As in package inserts or review reports of approved products, detection criteria should be indicated for CNV.

Copy numbers are relative numerical values calculated as ratios. Obtaining information on how pipeline calculations are performed is often difficult, as it is confidential information owned by companies. However, at the very least, information on whether the copy numbers indicated on reports are values corrected by ploidy or percentage of tumor is necessary for expert panel discussions; therefore, this information should be disclosed on package inserts or in reports. As for the results of analysis of CNV, a gene panel that outputs results simply as normal or abnormal is not adequate. In particular, when aneuploidy or severe nuclear abnormality exists, such corrections are limited, and an expert panel discussion is required.

Technically, it is possible to set LOD for CNV using cell lines from the ATCC; however, for ATCC cell lines that are not available in DNA form, it is important to note that the DNA will be extracted after successive passages and that reference values will differ depending on the assessment method used. In fact, there was a discrepancy between copy numbers estimated by Qiagen's real-time PCR published by the ATCC and those estimated by digital PCR in our working group meeting. Furthermore, this discrepancy was larger in cell lines with redundant copies (Figure S5). Therefore, it is possible to use results of real-time PCR or digital PCR after successive passages as reference values for assessing the analytical performance of gene panel CNV instead of copy numbers published by the ATCC. However, caution is warranted in externally comparing differences between panels, as measured values of copy numbers cannot be compared directly. For genes from ATCC cell lines that are not distributed in DNA form and are considered to have no CNV, using PCR or another method to verify that the copy number is actually two is necessary. Given the above, storage of specific lots by a third-party institution could be considered.

Approved CDx for *ERBB2* exist. However, for CNV in new genes for which diagnostics or reference values are not available, digital PCR and FISH are control methods assumed for accuracy testing. As optimizing FISH probes is time-consuming and laborious, digital PCR is more practical given its versatility for numerous genes. A set of ATCC euploidy cells will most likely be used for setting reference values for digital PCR.

In some cases, CNV that is positive by immunohistochemistry (IHC) or FISH in pathological specimens cannot be detected by panel testing because of dilution by tumor heterogeneity; thus, discrepancies between IHC or FISH and panel testing results are inevitable. Furthermore, the threshold for copy number described in the criteria

for verification tests and the effects of pharmaceutical drugs described in their package inserts are important for CDx.

It is difficult with the currently available knowledge to classify CNVs into those that require more careful analytical performance assessment and those that do not like SNV/MNV and Ins/Del.

## 4.2 | Summary

### 4.2.1 | If aiming for approval as a comprehensive genomic profiling assay

Limit of detection for *MET*, *EGFR*, *MYC-N*, and *ERBB2* can be assessed using commercially available reference standards. Limit of detection using reference standards is likely indicated in terms of copy number or percentage of tumor and hit rate, similar to that on package inserts of approved products. In addition, accuracy testing may be evaluated by comparison with reference values of reference standards, with the exception of reference standards with high copy numbers, such as Hs746T.

### 4.2.2 | If pharmaceutical drugs are to be administered based on results of comprehensive genomic profiling without confirmation by companion diagnostics after expert panel discussion

It is desirable to consider equivalence with approved CDx based on the results of accuracy testing and LOD assessment carried out at the time of application for regulatory approval as a CGP.

### 4.2.3 | Future outlook

Scientifically, when applying for approval as a CDx for genes covered by reference standards, LOD assessment using commercially available reference standards can be undertaken using the same precautions noted in "If aiming for approval as a CGP". When applying for CDx approval when the detection target and the CNV of the panel testing are different and where the approved CDx is IHC (e.g., *ERBB2*), assessing concordance rate with approved CDx as an equivalence assessment or accuracy testing is essential. Therefore, it is necessary to evaluate the concordance rate with the control method in accuracy testing (which also serves as equivalence assessment) when applying for approval as a CDx. An approved CDx should be the control for accuracy testing if it is available. If it is unavailable (i.e., the clinical specimen with assessment results from approved CDx cannot be obtained) or does not exist, the use of digital PCR with precision management is practical. For approved CDx, as a general rule, an original rather than a generic CDx should be used for the control method. However, in the event of the emergence of a CNV for which only panel testing or digital PCR are approved CDx, when applying for approval

for CDx of genes covered by reference standards, accuracy testing using reference standards should be carried out by comparing the reference values as an alternative equivalence assessment. However, when applying for approval as CDx for genes that are not covered by reference standards, assessment for LOD and accuracy testing (which doubles as equivalence assessment) using clinical specimens is considered necessary.

## 5 | FUSION

### 5.1 | General considerations

As indicated in the package inserts of approved products, LOD for fusion can be reported in terms of VAF or percentage of tumor or as number of reads and hit rate. In analytical performance assessment of fusion, determination of positive/negative status according to a given threshold value, such as with SNV/MNV or Ins/Del, and the range of detectable translocation partners are important.

For RNA panels, it is theoretically possible to undertake assessments by creating fusion gene RNA with various translocation partners of synthetic RNA, either synthesized from nucleic acid or refined from forced expression in cells. For DNA panels, it is not possible to exhaustively assess fusion genes with various translocation partners using available reference standards or clinical samples; thus, it is virtually impossible to evaluate whether a fusion gene with various translocation partners is detectable.

Horizon's HD753 for *SLC34A2/ROS1* and *CCDC6/RET* fusion is the only reference standard currently available that is related to fusion for DNA panels; it has also been used for analytical performance assessment of approved products. Our working group meeting has confirmed that it is detectable.

Seraseq FFPE Fusion RNA v4 Reference Material can be used for RNA panels but not DNA panels because fusion genes do not exist at the genome level. Unique read numbers assessed by copy numbers evaluated by digital PCR and testing using NGS (Archer FusionPlex) are published as reference values for this reference standard; however, these are not correlated with each other, whereas measurement results from oncogene panels correlate better with unique read numbers measured by NGS (Archer FusionPlex) according to the analysis in our working group meeting (Figure S6). *MET* exon 14 skipping is also included. As the dilution series of this reference standard is not commercially available, the dilution series must be created by the tester to calculate LOD.

The limited availability of reference standards for fusion is a barrier for analytical performance assessments, particularly for DNA panels. Hence, future international development and organization of reference standards is desirable.

Capture and amplicon methods also differ with regards to fusion. In the amplicon method, fusion is detected by PCR analysis of each gene in the fusion; as such, limiting the number of translocation partner is necessary. The capture method can be used to identify translocation partners without limiting their number; however, detection

is sometimes reinforced by capturing the intron of the translocation partner (e.g., *NTRK3* introns are not detected by FoundationOne; however, *ETV6/NTRK3* can be detected by detecting introns of the translocation partner). Therefore, consideration of the testing methods of each gene panel is necessary.

Detecting fusion in genes with repetitive and analogous sequences (e.g., *ROS1*) is difficult using a DNA panel, and there are limitations to equivalence assessment.

It is useful to publish a list of fusion gene data reported to the Catalogue of Somatic Mutations in Cancer (Table S2) as a product of our working group meeting, as such a list does not currently exist. This list compiles both in-frame and out-of-frame fusions. Detection frequency is also listed based on publicly available data.

Whether the various panels can actually detect the fusion genes in the above list cannot be tested exhaustively due to limits on clinical specimens and reference standards. However, estimation based on probe position and other parameters is possible. To provide an expert panel information on which fusions theoretically are or are not detectable on the basis of designed probe position is important, and providing these data by means of documents from distributors and other sources is recommended. Clearly stating the limitations according to panel design, principle, and methodology is desirable.

As stated above, it is not practical to exhaustively study the fusions concerned by means of clinical specimens, synthetic specimens, and reference standards. There are fusions for which the only approved CDx are DNA panels, and limitations to the analytical performance of DNA panels for fusions exist. Hence, it is important to note that there are some limitations in assessing equivalence with approved CDx for fusion.

## 5.2 | Summary

### 5.2.1 | If aiming for approval as a comprehensive genomic profiling assay

Calculating LOD using commercially available reference standards is acceptable. Limit of detection using reference standards will most likely be indicated in terms of VAF or percentage of tumor, number of reads, and hit rate, similar to that in package inserts of approved products. Assessment of accuracy testing by comparison to reference values is also acceptable.

### 5.2.2 | If pharmaceutical drugs are to be administered based on results of comprehensive genomic profiling without confirmation by companion diagnostics after expert panel discussion

It is desirable to consider equivalence with approved CDx in an expert panel based on the results of accuracy testing and LOD assessment undertaken at the time of application for regulatory approval as a CGP. In doing so, considering whether or not fusions not

covered by reference standards used in accuracy testing and LOD assessment are theoretically detectable based on probe design and other parameters is recommended.

### 5.2.3 | Future outlook

Scientifically, when applying for approval as a CDx for genes covered by reference standards, as an alternative method to evaluate equivalence with an approved CDx, accuracy testing by comparing with reference values of reference standards can be utilized. The concerns for such accuracy testing are the same as those in "If aiming for approval as a CGP." Furthermore, investigating the theoretical detectability of fusion based on probe design and other parameters is required. When applying for approval as a CDx for genes not covered by reference standards, evaluating for equivalence with approved CDx using clinical specimens has some significance if approved CDx other than the DNA panels, such as IHC or FISH, are available.

## 6 | TUMOR MUTATIONAL BURDEN

### 6.1 | General considerations

The analytical performance assessment listed in table 2 of the *Journal for ImmunoTherapy of Cancer* (JITC) article on tumor mutational burden (TMB) by Friends of Cancer Research<sup>10</sup> serves as a good reference. In particular, analytical performance of individual SNV/MNV or Ins/Del mutations is assessed separately; thus, evaluating the validity of the searched region by concordance rate with whole exome sequence (WES) in *in silico* analyses using data published by The Cancer Genome Atlas (TCGA) or in published articles is logical for assessing analytical performance of TMB. However, as these data are intended for validation, data used for designing gene panels should be excluded. A software tool published by the Friends of Cancer Research project is also a good reference for *in silico* analysis.<sup>11</sup>

In either analyses using clinical specimens or *in silico* analysis, using WES as the control is problematic because WES itself is not approved as a diagnostic device, and its quality should thus be supported by published papers. Therefore, comparing the results side-by-side would be meaningful if *in silico* analysis is carried out for various gene panels using common datasets from those published in the TCGA or in articles.

As for commercially available reference standards, SeraCare products used in the Friends of Cancer Research project (Seraseq TMB Genomic DNA Mix/FFPE Reference Material) are available.

As stated in the JITC paper by the Friends of Cancer Research, TMB values calculated by various gene panels vary even within the same individuals. Thus, it is possible that various gene panels provide threshold values different from FoundationOne CDx's TMB  $\geq 10$  mutations per megabase (mut/Mb) for administering pembrolizumab, which has already been approved in the United States. To resolve these issues, it is



necessary either to compare TMB values provided by FoundationOne CDx in the same specimen or to test the threshold value of TMB that correlates with pembrolizumab efficacy in new clinical trials using various gene panels. However, neither of these options is very feasible. Thresholds and analytical performance at and near the threshold are particularly important for CDx. As stated in the document published by the Friends of Cancer Research on their website,<sup>12</sup> a clear difference is observed in the response rate to anti-programmed cell death (PD)-1/PD ligand-1 antibodies between TMB high/low groups when TMB 10 mut/Mb is defined as the minimum threshold. Therefore, at present, the consensus is that 10 mut/Mb is the threshold, but the clinical significance of rigorously evaluating equivalence between gene panels using a specific TMB threshold is not necessarily high.

As it contributes to expert panel discussions, information about the methods of calculating TMB (whether it contains actionable or germline mutations) should be provided.

## 6.2 | Summary

### 6.2.1 | If aiming for approval as a comprehensive genomic profiling assay

For accuracy testing, evaluating concordance rate with WES through *in silico* analysis using data published in TCGA or articles is acceptable. In doing so, including various samples of TMB values in the 0–40 mut/Mb range and of various cancer types is desirable. For TMB calculation by WES, the type of mutation, sequence quality control, threshold for mutation detection, and region used for the calculation should be determined by scientifically valid methods. The use of commercially available reference standards is also possible.

### 6.2.2 | If pharmaceutical drugs are to be administered based on results of comprehensive genomic profiling without confirmation by companion diagnostics after expert panel discussion

Based on the results of accuracy testing undertaken at the time of application for regulatory approval as a CGP, discussing the equivalence with approved CDx by an expert panel is desirable. In that event, focusing on the analytical performance near the cut-off mut/Mb value for approved CDx is required.

### 6.2.3 | Future outlook

When the analytical performance near the cut-off mut/Mb value for approved CDx can be considered guaranteed based on results of analytical performance assessment carried out as described in “If aiming for approval as a CGP,” the results of accuracy testing indicated in that section can be used as a substitute for equivalence test with the approved CDx for the application for approval as a CDx.

For evaluation of LOD to detect values equal to or above the cut-off value, indicating the LOD as a percentage of tumor and hit rate, as noted in the package insert for the product approved in the United States, is acceptable.

## 7 | MICROSATELLITE INSTABILITY

### 7.1 | General considerations

Horizon MSI FFPE DNA Reference Standard (microsatellite instability [MSI]-high: HD830; microsatellite stable: HD831), testing of which has been undertaken by distributors on multiple biomarkers, including five locations identical to the Promega panel, is usable as a reference standard for gene panels that target these genes. However, it is difficult to assess analytical performance using this reference standard in gene panels that detect repetitions in regions that have not been tested by this reference standard.

As it contributes to discussions by an expert panel, information about the methods of calculating MSI should be provided.

### 7.2 | Summary

#### 7.2.1 | If aiming for approval as a comprehensive genomic profiling assay

Using dilution series of commercially available reference standards for gene panels to detect biomarkers that have been tested by distributors is acceptable. Limit of detection using reference standards should be indicated in terms of hit rate. For accuracy testing, assessing by comparing the reference values of reference standards is acceptable. If biomarkers that have been tested by the distributor are not the target for detection, assessing equivalence with the approved CDx using clinical specimens or a standard method is required.

#### 7.2.2 | If pharmaceutical drugs are to be administered based on results of comprehensive genomic profiling without confirmation by companion diagnostics after expert panel discussion

Based on the results of accuracy testing and LOD assessment carried out at the time of application for regulatory approval as a CGP, discussion of the equivalence with approved CDx by an expert panel is desirable.

#### 7.2.3 | Future outlook

For applications for approval as CDx, assessing equivalence with an approved CDx using clinical specimens or a standard method in gene

panels that target biomarkers that have been tested by the distributor has some benefits and is desirable. In doing so, as a general rule, the original approved CDx, not a generic, should be the control.

## 8 | PLASMA SPECIMENS

### 8.1 | General considerations

While designing gene panels that use tumor tissue specimens to minimize false positives has been required, efforts have been made to increase sensitivity in liquid biopsy; this difference makes the assessment of LOD, limit of blank, and error rates by measuring reference standards and their dilution sequences particularly important. However, as a scientific note on analytical performance assessment, it is safe to consider that there are no major differences between liquid biopsy and gene panels using tissue specimens. It should be noted that discussions of liquid biopsy will differ from discussions of gene panel testing for tissue specimens in terms of evaluating the concordance rate between the tissue and plasma specimens.

As mentioned in the Japanese Society of Medical Oncology/Japan Society of Clinical Oncology/Japanese Cancer Association's Three-Society Genomic Medicine Promotion Joint Taskforce report,<sup>13</sup> liquid biopsy must be differentiated from clonal hematopoiesis of indeterminate potential, and it may be difficult to assess CNV and fusion.

Regarding liquid biopsy, for example, the availability of reference standards obtained from Horizon, Thermo Fisher Scientific, and SeraCare has been published in *JCO Precision Oncology*, as obtained from the Foundation for the National Institutes of Health Biomarkers Consortium.<sup>14</sup> Some of these reference standards are listed in the Medical Device Innovation Consortium SRS Report: Somatic Variant Reference Samples for NGS.

Biomarkers for which CDx have been approved for liquid biopsy are limited. Assessing concordance rate with measured results of tissue specimens paired with liquid biopsy has biological limitations such as tumor heterogeneity, and paired tissue specimens can often be obtained only from patients enrolled in clinical trials; thus, LOD and accuracy testing using reference standards is considered more practical.

With regards to assessment of analytical performance with liquid biopsies, the US Blood Profiling Atlas Consortium's published paper<sup>15</sup> is a good reference, and the template protocol is available as supplemental material.

### 8.2 | Summary

#### 8.2.1 | If aiming for approval as a comprehensive genomic profiling assay

As with gene panel testing with tissue specimens, evaluating analytical performance using reference standards is acceptable.

#### 8.2.2 | If pharmaceutical drugs are to be administered based on results of comprehensive genomic profiling without confirmation by companion diagnostics after expert panel discussion

Consideration by an expert panel is recommended based on the results of accuracy testing and LOD assessment carried out at the time of application for regulatory approval as a CGP and on whether the clinical trial outcomes of the corresponding pharmaceutical drug were based on results of testing using tissue specimens or liquid biopsy.

#### 8.2.3 | Future outlook

If only approved CDx for tissue specimens exist, even if accuracy testing with an approved CDx as the control is feasible, prioritizing the results of analytical performance evaluations using reference standards is scientifically preferable when applying for approval as a CDx of gene mutations covered by reference standards. The same details as those described in "Future outlook" for gene panel testing on tissue specimens are expected when applying for approval as a CDx of gene mutations not covered by reference standards.

## 9 | CONCLUSION

We convened the working group committee that consisted of more than 30 experts from academia, industry, and government. We have discussed the points that should be considered to allow maximal simplification of assessments using clinical specimens in evaluating accuracy and LOD in equivalence and analytical performance of oncology gene panels for 3 years. This article provides recommendations specific to each type of gene mutation as well as to reference standards used for evaluations. In addition, detailed results of analyses regarding reference standards in our working group are described as supplementary figures in this article. The context and wording of the document were carefully discussed and finalized in our working group committee based on public comments as well as committee members' opinions from scientific and regulatory viewpoints. We believe that our study makes a significant contribution to the literature and will be of interest to global readers because these recommendations will streamline and simplify evaluation protocols for gene panels and promote the development and approval of not only new CGP but also new CDx.

### ACKNOWLEDGMENTS

We would like to thank Hiroyuki Mano, Japan Agency for Medical Research and Development (AMED), the Pharmaceuticals and Medical Devices Agency, Ministry of Health, Labour and Welfare, Japan Pharmaceutical Manufacturers' Association, Pharmaceutical Research and Manufacturers of America, European Federation of Pharmaceutical Industries and Associations, Japan Association




of Clinical Reagents Industries, Japan Analytical Instruments Manufacturers' Association, and Japan Bio Measurement and Analysis Consortium for their fruitful discussion, advice, and participation in our working group meetings. We would like to thank Editage ([www.editage.com](http://www.editage.com)) for English language editing.

## CONFLICTS OF INTEREST

Sumimasa Nagai has received lecture fees from Konica Minolta. Hiroshi Nishihara has received remuneration as an advisor from General Insurance Co., Ltd and Seiwa Memorial Hospital, and has received research funds from Mitsubishi Space Software Inc., Kizawa Memorial Hospital, AccuRna Inc., and Sanofi Co. Kazuto Nishio has received lecture fees from Boehringer Ingelheim Japan, AstraZeneca, Chugai, and Novartis Pharma and has received research funds from Nippon Boehringer Ingelheim, Korea Otsuka Pharmaceutical, Eli Lilly Japan, Ignyta, West Japan Oncology Group, Thoracic Oncology Research Group, North East Japan Study Group, Clinical Research Support Center Kyushu, Nichirei Biosciences, and Osakaminami Hospital. Hiroyuki Aburatani has received remuneration as an advisor from Konica Minolta, has received research funds from Konica Minolta and Chugai Pharmaceutical, and belongs to an endowed chair funded by Chugai Pharmaceutical. Shinji Kohsaka has received research funds from Konica Minolta. Hiroshi Nishihara, Kazuto Nishio, Katsuya Tsuchihara, Hiroyuki Aburatani, and Takashi Kohno are current associate editors of *Cancer Science*. The other authors have no conflict of interest.

## ORCID

Sumimasa Nagai  <https://orcid.org/0000-0001-8401-6074>  
 Takayoshi Suzuki  <https://orcid.org/0000-0002-2261-5263>  
 Kazuto Nishio  <https://orcid.org/0000-0002-8275-0846>  
 Hiroya Taniguchi  <https://orcid.org/0000-0003-1407-6682>  
 Katsuya Tsuchihara  <https://orcid.org/0000-0001-7507-2349>  
 Kohei Nakamura  <https://orcid.org/0000-0002-5162-7132>  
 Takashi Kohno  <https://orcid.org/0000-0002-5371-706X>  
 Shinji Kohsaka  <https://orcid.org/0000-0001-8651-6136>

## REFERENCES

1. Working group of the Japan Agency for Medical Research and Development Research on regulatory science of pharmaceuticals and medical devices "Investigation for Standardization of Companion Diagnostic System by Gene Panel Testing". Recommendations related to the analytical equivalence assessment of gene panel testing. Accessed April 22, 2022. <https://www.jsmo.or.jp/news/jsmo/20220307.html>, <http://www.jsco.or.jp/jpn/index/page/id/2571>, <https://www.jca.gr.jp/researcher/topics/2022/20220307.html>
2. Pharmaceuticals and Medical Devices Agency Companion Diagnostics WG. New Regulations for Companion Diagnostic Systems (Draft). August 9, 2019.
3. Ministry of Health, Labour and Welfare. Technical guidance on pharmaceutical drugs related to companion diagnostic drugs. December 26, 2013, PFSB/ELD Notification.

4. Ministry of Health, Labour and Welfare. Concerns on National Health Insurance coverage of gene panel testing. May 31, 2019.
5. Japanese Promotion Council for Laboratory Testing. Fundamental Premises on Quality and Precision Assurance of Oncogene Panel Testing (Ver. 2.0). Accessed March 31, 2022. <https://www.jpclt.org/message/#ctDNA>
6. Medical Device Innovation Consortium. SRS Report: Somatic Variant Reference Samples for NGS. Accessed March 31, 2022. <https://mdic.org/resource/srs-landscape-analysis-report/>
7. Jones W, Gong B, Novoradovskaya N, et al. A verified genomic reference sample for assessing performance of cancer panels detecting small variants of low allele frequency. *Genome Biol.* 2021;22(1):111. doi:10.1186/s13059-021-02316-z
8. He HJ, Das B, Cleveland MH, et al. Development and interlaboratory evaluation of a NIST reference material RM 8366 for EGFR and MET gene copy number measurements. *Clin Chem Lab Med.* 2019;57(8):1142-1152. doi:10.1515/cclm-2018-1306
9. Lih CJ, Si H, Das B, et al. Certified DNA reference materials to compare HER2 gene amplification measurements using next-generation sequencing methods. *J Mol Diagn.* 2016;18:753-761. doi:10.1016/j.jmoldx.2016.05.008
10. Merino DM, McShane LM, Fabrizio D, et al. Establishing guidelines to harmonize tumor mutational burden (TMB): in silico assessment of variation in TMB quantification across diagnostic platforms: phase I of the friends of cancer research TMB harmonization project. *J Immunother Cancer.* 2020;8:e000147. doi:10.1136/jitc-2019-000147
11. Vega DM, Yee LM, McShane LM, et al. Aligning tumor mutational burden (TMB) quantification across diagnostic platforms: phase 2 of the friends of cancer research TMB harmonization project. *Ann Oncol.* 2021;32(12):1626-1636. doi:10.1016/j.annonc.2021.09.016
12. Friends of Cancer Research. Tissue agnostic TMB clinical cut-off harmonization initiative. Accessed March 31, 2022. [https://www.focr.org/sites/default/files/Tissue-Agnostic-TMB\\_Summary.pdf](https://www.focr.org/sites/default/files/Tissue-Agnostic-TMB_Summary.pdf)
13. Sunami K, Bando H, Yatabe Y, et al. Appropriate use of cancer comprehensive genome profiling assay using circulating tumor DNA. *Cancer Sci.* 2021;112(9):3911-3917. doi:10.1111/cas.15022
14. Williams PM, Forbes T, P Lund S, et al. Validation of ctDNA quality control materials through a precompetitive collaboration of the Foundation for the National Institutes of Health. *JCO Precis Oncol.* 2021;5:910-920. doi:10.1200/PO.20.00528
15. Godsey JH, Silvestro A, Barrett JC, et al. Generic protocols for the analytical validation of next-generation sequencing-based ctDNA assays: a joint consensus recommendation of the bloodPAC's analytical variables working group. *Clin Chem.* 2020;66(9):1156-1166. doi:10.1093/clinchem/hvaa164

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Nagai S, Nishihara H, Suzuki T, et al. Recommendations related to the analytical equivalence assessment of gene panel testing. *Cancer Sci.* 2022;113:3282-3290. doi: [10.1111/cas.15513](https://doi.org/10.1111/cas.15513)