Reference Samples to Compare Next-Generation Sequencing Test Performance for Oncology Therapeutics and Diagnostics

John D. Pfeifer, MD, PhD,[1] *Robert Loberg, PhD,*[2] *Catherine Lofton-Day, PhD,*[3] *and Barbara A. Zehnbauer, PhD*[4]

From the [1]Department of Pathology, Washington University School of Medicine, St Louis, MO, USA; [2]Clinical Biomarkers and Diagnostics and [3]In Vitro Diagnostics, Amgen, Thousand Oaks, CA, USA; and [4]Department of Pathology, Emory University School of Medicine, Atlanta, GA, USA.

ABSTRACT

Objectives: Diversity of laboratory-developed tests (LDTs) using next-generation sequencing (NGS) raises concerns about their accuracy for selection of targeted therapies. A working group developed a pilot study of traceable reference samples to measure NGS LDT performance among a cohort of clinical laboratories.

Methods: Human cell lines were engineered via CRISPR/Cas9 and prepared as formalin-fixed, paraffin-embedded cell pellets ("wet" samples) to assess the entire NGS test cycle. In silico mutagenized NGS sequence files ("dry" samples) were used to assess the bioinformatics component of the NGS test cycle. Single and multinucleotide variants (n = 36) of *KRAS* and *NRAS* were tested at 5% or 15% variant allele fraction to determine eligibility for therapy with the EGFR inhibitor panitumumab in the setting of metastatic colorectal cancer.

Results: Twenty-one (21/21) laboratories tested wet samples; 19 of 21 analyzed dry samples. Of the laboratories that tested both the wet and dry samples, 7 (37%) of 19 laboratories correctly reported all variants, 3 (16%) of 19 had fewer than five errors, and 9 (47%) of 19 had five or more errors. Most errors were false negatives.

Conclusions: Genetically engineered cell lines and mutagenized sequence files are complementary reference samples for evaluating NGS test performance among clinical laboratories using LDTs. Variable accuracy in detection of genetic variants among some LDTs may identify different patient populations for targeted therapy.

INTRODUCTION

Massively parallel sequencing approaches, referred to as next-generation sequencing (NGS), permit the simultaneous evaluation of many genes and a range of sequence variations.[1] NGS can be performed to detect all four major classes of sequence variants, including single-nucleotide variants (SNVs), small insertions and deletions (indels, defined as less than 1 kb in length), copy number variants (CNVs), and structural variants (SVs) such as translocations and inversions. The clinical utility of NGS testing methods resides in the flexibility of assay design, including library preparation methods (amplification or hybrid capture), different sequencing instrument platforms, and genomic target regions spanning panels of mutation hotspots in a small number of genes to panels representing hundreds of genes, the whole exome, or the whole genome.

KEY POINTS

- Engineered cell lines and in silico mutagenized sequence files are complementary reference materials that can be used to assess the accuracy of clinical next-generation sequencing (NGS) test results.
- The accuracy of detection of genetic variants differed among the laboratory-developed tests (LDTs) performed by different laboratories.
- The varied accuracy suggests that different LDTs may identify different subsets of oncology patients as candidates for targeted therapy.

KEY WORDS

Precision medicine; Companion diagnostic; Next-generation sequencing; Reference materials; Test performance

Am J Clin Pathol April 2022;157:628-638
HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

Received: June 22, 2021
Accepted: August 24, 2021
Advance publication: December 6, 2021

Corresponding author: John D. Pfeifer, MD, PhD; jdpfeifer@wustl.edu.

Funding: This work was supported by contracts from Amgen that financed the work of Tapestry Networks, CAP, Horizon Discovery Group, P&V Licensing, and of Dr Zehnbauer (consultant).

Disclosure: Dr Pfeifer is the cofounder of P&V Licensing, which provides in silico mutagenized NGS sequence files to CAP, including for the SPOT/Dx pilot.

Much of clinical NGS testing is performed using customized, in-house assay designs (known as laboratory-developed tests [LDTs]) unique to an individual laboratory; while many LDTs are designed around gene panels specific to the laboratory, other LDTs use commercially available kits (which may undergo some limited in-house customization). Mirroring the variability of clinical NGS assay design, many different bioinformatics pipelines are in routine clinical use, including pipelines integrated into the manufactured sequencing platforms, pipelines that may be licensed from commercial vendors, and informatics pipelines developed by individual laboratories.[2,3]

Multiple professional guidelines have been defined by regulatory agencies[4,5] and professional organizations[2,6-11] to assist laboratories in the design, development, and validation of NGS assays. Proficiency testing (PT) programs for clinical NGS-based oncology assays (eg, those offered by the College of American Pathologists [CAP] that focus on common, clinically relevant variants for more than 30 genes) have shown that NGS approaches have high interlaboratory agreement,[12,13] are highly concordant with other molecular diagnostic methods in clinical use for specific genetic variants,[14] and may show superior performance to non-NGS methods.[15] Similar results have been demonstrated by laboratory comparisons outside the setting of formal proficiency testing.[16,17] Studies that focused specifically on the bioinformatics components of NGS tests have likewise demonstrated concordance among various laboratories and sequencing platform types.[18-20]

Nonetheless, the diversity of LDTs used in clinical diagnostic laboratories has generated questions from many stakeholders about how to compare assay performance between LDTs with the same intended use. Comprehensive interlaboratory comparisons of NGS LDT performance are hampered by the limited availability of well-defined reference samples, particularly for rare genetic variants. Similarly, new test development and assay validation require well-designed and manufactured reference samples. Residual specimens from appropriately consented patients may be available to clinical laboratories but are limited in supply, specific to a patient's disease status, not sustainable nor scalable, and not a reproducible standard. In contrast, ideal reference samples include a variety of variant types in many different genes, multiplexed variants, different variant allele fractions (VAFs), and a range of variant types (eg, SNVs, indels, CNVs, and SVs). Different strategies for developing reference samples have been employed,[21] including tumor-derived as well as engineered human cell lines[22,23] and so-called spike-in standards of synthetic or cloned DNA fragments.[13,24-28] Complementary approaches that use data files harboring sequence variants inserted via computerized methods into NGS sequence files (in silico reference samples) may be used to separate the evaluation of bioinformatics data interpretation of an NGS test from the sequencing chemistry portion.[19,29-31]

Within this landscape, Tapestry Networks, a professional services firm,[32] convened senior health care leaders to address emerging challenges in oncology diagnostics and therapeutics. The Sustainable Predictive Oncology Therapeutics and Diagnostics (SPOT/Dx) working group was formed by Tapestry Networks[33] and comprised many stakeholders, including patient advocacy groups, professional oncology organizations, payers, regulatory agencies, NGS laboratories, and others (Supplementary Material S1; all supplemental materials can be found at *American Journal of Clinical Pathology* online). The Diagnostic Quality Assurance Pilot,[34] hereafter the Pilot, was conceived by the SPOT/Dx working group to address a perceived gap in the standardization of personalized medicine laboratory testing for targeted therapies in cancer treatment. The central goal was to model a process that could assure prescribing physicians that consistent and accurate test findings were produced across the landscape of clinical NGS testing.

Standardized reference samples were essential to this comparative effort; therefore, the Pilot focused on two specific aims. First, the Pilot would define the requirements for the development of traceable reference samples that could function as human reference materials for molecular pathology laboratories. Second, the Pilot would model the utility of the reference samples via an analytic challenge involving a cohort of clinical NGS laboratories using LDTs in the setting of gene-targeted therapy for colorectal adenocarcinoma.[35-38] To fulfill the first specific aim of the Pilot, two types of reference samples were designed to assess the performance of the "complete analytical NGS test cycle" and to separately assess the "bioinformatics analysis" of NGS assay data. Reference samples for the complete test cycle challenge were human cell lines (wet samples) with specific *KRAS* and *NRAS* genetic variants integrated into human genomic DNA via genetic engineering. The second type of reference sample (dry samples) was produced by introducing genetic variants into sequence data files using computerized algorithms designed to separately assess the accuracy of bioinformatics analysis pipelines. Both types of reference samples were distributed to laboratories participating in the Pilot, and each laboratory tested the reference samples using the validated LDT they employed in routine clinical laboratory practice.

The Pilot's overall design thus represents a novel approach for comparison of LDTs. First, the analytic challenge did not consist of unrelated variants tested without context to a specific disease process; rather, the challenge was intentionally designed around testing in a specific clinical setting to model an actual patient encounter. Consequently, the included variants were not selected to be representative of the types of mutations generally encountered in routine clinical NGS; instead, variants included mutations with demonstrated roles in gene-targeted therapy for a defined clinical diagnostic setting of colorectal adenocarcinoma.[35-38] Second, the included VAFs were intended to sample LDT accuracy not only at higher VAFs but also at VAFs near the limit of detection (LOD) of the LDTs. Third, the results of the analytic challenge were not evaluated via aggregated data but rather on an LDT-by-LDT basis. This latter approach is more aligned with the clinical reality that patient care is generally determined by the result of an NGS test performed by a single laboratory rather than by aggregate results from a group of NGS laboratories performing similar testing.

MATERIALS AND METHODS

Organization of the Diagnostic Quality Assurance Pilot Study

The infrastructure created to support and oversee the design and execution of the Pilot FIGURE 1 included a multistakeholder

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

Steering Committee and a Scientific and Technical Working Group (STWG) (Supplementary Material S2). The CAP coordinated the STWG and provided professional, logistical, and operational expertise in support of the Pilot.

CAP solicited participation in the Pilot from Clinical Laboratory Improvement Amendments (CLIA)–certified clinical laboratories registered in CAP's NGS proficiency testing survey programs. The solicitation packet included information describing the intent of the Pilot to evaluate LDT performance based on the clinical utility of a test approved by the US Food and Drug Administration (FDA) (Supplementary Material S3). Twenty-six laboratories volunteered from which the STWG selected 20 laboratories (the Pilot laboratories) to represent different sequencing platforms, test design (amplicon based vs hybrid capture based), laboratory setting (commercial vs academic medical center), and annual test volume.

Genetic Variants

The Pilot focused on *KRAS* and *NRAS* sequence variants TABLE 1 important for selection of patients with metastatic colorectal cancer likely to obtain therapeutic benefit with the epidermal growth factor receptor (EGFR) inhibitor panitumumab (Vectibix; Amgen).[35-38] Advantages of this model were that both SNVs and small indels (specifically, di- and trinucleotide sequence substitutions, also known as multinucleotide variants or MNVs) were included, variants within three different exons of each gene have clinical impact, and an FDA-approved companion diagnostic (the Praxis Extended RAS Panel) provided a comparator for NGS testing for selection of patients with colorectal cancer for targeted *EGFR* inhibitor therapy.[35]

Aim 1: Traceable Reference Samples

Wet Samples

Horizon Discovery Biosciences plc genetically engineered the wet lab samples using CRISP/Cas9 technology into the haploid eHAP cell line.[39] All variants were inserted at the endogenous genomic locations to minimize nonphysiologic sequence artifacts that could bias NGS bioinformatics pipelines. All engineered cell lines were isogenic except for each introduced variant to avoid complications secondary to bioinformatics filters designed to detect specimen contamination.[40-42] The cell lines were propagated, mixed with the parent cell line to achieve 5% VAF, and harvested as cell pellets, which were then formalin fixed and paraffin embedded (FFPE) to simulate FFPE tissue. In practice, the final VAF of the variants in the wet samples TABLE 1 ranged from 5% to 8.6% (as confirmed by digital droplet polymerase chain reaction [PCR] performed by the vendor). Laboratories were provided 20-µm FFPE sections to yield 1 µg DNA (Supplementary Material S4).

Dry Samples

P&V Licensing performed the in silico mutagenesis of NGS sequence data files. A 5% VAF was selected to mimic the VAFs of the wet laboratory challenge; a 15% VAF was also provided as an alternate to be more easily detectable. Each participating laboratory sequenced DNA from eHAP using their LDT and submitted the FASTQ or BAM sequence file via an online interface maintained by the CAP. The data files were accessed by the vendor from this interface (the laboratories' identities were unknown to the vendor), and *RAS* variants were inserted into the sequence file to generate a custom mutagenized file specific for each participant laboratory's LDT. Each mutagenized data file was then downloaded from the online interface and analyzed by the respective laboratory (Supplementary Material S4).

Aim 2: Analytic Challenge

Proof-of-Concept Laboratories

Three CLIA-certified clinical laboratories with expertise in clinical NGS for oncology patients were designated by the STWG as proof-of-concept (POC) laboratories to verify the performance of the reference samples. In addition, the reference samples were analyzed by Illumina (the CDx manufacturer) using the Praxis Extended RAS Panel. The POC laboratories also verified the data collection submission process using the CAP result form (Supplementary Material S5).

Participating Pilot Laboratories

The Pilot laboratories included 15 academic medical centers and 6 commercial laboratories, with 17 using an Illumina platform, 3 using Thermo Fisher Ion Torrent platforms, and 1 using another platform. Pilot laboratories reported performing between 9 and 21,000 NGS tests per year; 7 laboratories utilized an amplicon-based test, and 14

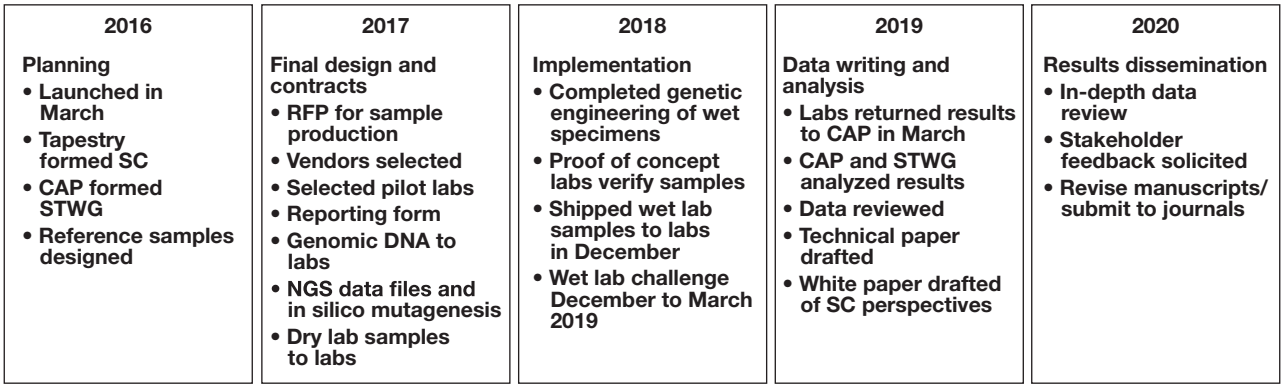

FIGURE 1 Sustainable Predictive Oncology Therapeutics and Diagnostics Pilot timeline and highlights. Developmental stages are described with time intervals of duration assigned to successive calendar years. CAP, College of American Pathologists; NGS, next-generation sequencing; RFP, request for proposal; SC, steering committee; STWG, Scientific and Technical Working Group.

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

TABLE 1 Variants Introduced Into the Wet and Dry Reference Samples

Sample	VAF %	SNV /Small Indel	*KRAS* (NM_004985.4)	Sample	VAF %	SNV /Small Indel	*NRAS* (NM_002524.4)
1A	5.5	SNV	*KRAS* p.Lys117Asn c.351A>C chr12:25378647T>G	4A	7.0	SNV	*NRAS* p.Ala146Thr c.436G>A
Wet				Wet			chr1:115252204C>T
	6.2	SNV	*KRAS* p.Ala59Thr c.175G>A		5.7	SNV	*NRAS* p.Gln61His c.183A>C
			chr12:25380283C>T				chr1:115256528T>G
	6.9	Small indel	*KRAS* p.Gly13Glu c.38_39GC>AA		6.0	SNV	*NRAS* p.Gly12Asp c.35G>A
			chr12:25398281_25398280CG>TT				chr1:115258747C>T
2A	6.3	SNV	*KRAS* p.Lys117Asn c.351A>T	5A	6.2	SNV	*NRAS* p.Lys117Asn c.351G>T
Wet			chr12:25378647T>A	Wet			chr1:115252289C>A
	7.4	SNV	*KRAS* p.Gln61Arg c.182A>G		7.1	SNV	*NRAS* p.Gln61Leu c.182A>T
			chr12:25380276T>C				chr1:115256529T>A
	5.4	SNV	*KRAS* p.Gly12Asp c.35G>A		8.6	Small indel	*NRAS* p.Gly12Trp c.34_36GGT>TGG
			chr12:25398284C>T				chr1:115258748_115258746CC A>ACC
3A	7.5	SNV	*KRAS* p.Ala146Pro c.436G>C	6A	5.6	SNV	*NRAS* p.Ala59Thr c.175G>A
Wet			chr12:25378562C>G	Wet			chr1:115256536C>T
	5.3	SNV	*KRAS* p.Gln61His c.183A>C		6.7	Small indel	*NRAS* p.Gly13Glu c.38_39GT>AA
			chr12:25380275T>G				chr1:115258744_115258743CA>TT
	5.1	Small indel	*KRAS* p.Gly12Trp c.34_36GGT>TGG		8.4	SNV	*NRAS* p.Gly12Val c.35G>T
			chr12:25398285_25398283CCA>ACC				chr1:115258747C>A
1B	5.0	Small indel	*KRAS* p.Gly13Glu c.38_39GC>AA	4B	15.0	SNV	*NRAS* p.Gly12Asp c.35G>A
Dry			chr12:25398281_25398280CG>TT	Dry			chr1:115258747C>T
	5.0	SNV	*KRAS* p.Ala59Thr c.175G>A		15.0	SNV	*NRAS* p.Ala59Thr c.175G>A
			chr12:25380283C>T				chr1:115256536C>T
	5.0	SNV	*KRAS* p.Lys117Asn c.351A>C		15.0	SNV	*NRAS* p.Ala146Thr c.436G>A
			chr12:25378647T>G				chr1:115252204C>T
2B	5.0	SNV	*KRAS* p.Gly12Asp c.35G>A	5B	5.0	Small indel	*NRAS* p.Gly12Trp c.34_36GGT>TGG
Dry			chr12:25398284C>T	Dry			chr1:115258748_115258746CC A>ACC
	5.0	SNV	*KRAS* p.Gln61Arg c.182A>G		5.0	SNV	*NRAS* p.Gln61Leu c.182A>T
			chr12:25380276T>C				chr1:115256529T>A
	5.0	SNV	*KRAS* p.Lys117Asn c.351A>T		15.0	SNV	*NRAS* p.Thr148Ser c.443C>G
			chr12:25378647T>A				chr1:115252197G>C
3B	15.0	SNV	*KRAS* p.Gly12Asp c.35G>A	6B	15.0	Small indel	*NRAS* p.Gly13Glu c.38_39GT>AA
Dry			chr12:25398284C>T	Dry			chr1:115258744_115258743CA>TT
	15.0	SNV	*KRAS* p.Gln61His c.183A>C		5.0	SNV	*NRAS* p.Gly12Val c.35G>T
			chr12:25380275T>G				chr1:115258747C>A
	15.0	SNV	*KRAS* p.Thr127Ile c.380C>T		15.0	SNV	*NRAS* p.Gly60Glu c.179G>A
			chr12:25378618G>A				chr1:115256532C>T
1C	15.0	Small indel	*KRAS* p.Gly13Glu c.38_39GC>AA	4C	5.0	Small indel	*NRAS* p.Gly13Glu c.38_39GT>AA
Dry			chr12:25398281_25398280CG>TT	Dry			chr1:115258744_115258743CA>TT
	15.0	SNV	*KRAS* p.Ala59Thr c.175G>A		5.0	SNV	*NRAS* p.Gly12Val c.35G>T
			chr12:25380283C>T				chr1:115258747C>A
	15.0	SNV	*KRAS* p.Lys117Asn c.351A>C		5.0	SNV	*NRAS* p.Ala59Thr c.175G>A
			chr12:25378647T>G				chr1:115256536C>T
2C	15.0	Small indel	*KRAS* p.Gly12Trp c.34_36GGT>TGG	5C	15.0	Small indel	*NRAS* p.Gly12Trp c.34_36GGT>TGG
Dry			chr12:25398285_25398283CCA>ACC	Dry			chr1:115258748_115258746CC A>ACC
	15.0	SNV	*KRAS* p.Gln61Arg c.182A>G		15.0	SNV	*NRAS* p.Gln61Leu c.182A>T
			chr12:25380276T>C				chr1:115256529T>A

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

TABLE 1 (cont)

Sample	VAF %	SNV /Small Indel	*KRAS* (NM_004985.4)	Sample	VAF %	SNV /Small Indel	*NRAS* (NM_002524.4)
	15.0	SNV	*KRAS* p.Ala146Pro c.436G>C		5.0	SNV	*NRAS* p.Lys117Asn c.351G>T
			chr12:25378562C>G				chr1:115252289C>A
3C	5.0	Small indel	*KRAS* p.Gly12Trp c.34_36GGT>TGG	6C	5.0	SNV	*NRAS* p.Gly12Asp c.35G>A
Dry			chr12:25398285_25398283CCA>ACC	Dry			chr1:115258747C>T
	5.0	SNV	*KRAS* p.Gln61His c.183A>C		5.0	SNV	*NRAS* p.Gln61His c.183A>C
			chr12:25380275T>G				chr1:115256528T>G
	5.0	SNV	*KRAS* p.Ala146Pro c.436G>C		5.0	SNV	*NRAS* p.Ala146Thr c.436G>A
			chr12:25378562C>G				chr1:115252204C>T

SNV, single-nucleotide variant; VAF, variant allele fraction.

laboratories used a hybrid capture–based test. All Pilot laboratories utilized an LDT.

The instructions to the participating laboratories (Supplementary Material S4) specifically introduced the analytic challenge in the context of a patient with metastatic colorectal cancer being considered for treatment with panitumumab, to remind participating laboratories that reports should align with the CDx labeling. Descriptive data about each LDT employed were also collected using a SPOT/Dx Validation Pilot Result Form. The data included the test design, specimen requirements, variant alleles, and validated LOD VAFs, sequencing platform, library preparation method, targeted depth of coverage, and bioinformatics pipeline features.

Analysis of Reported Data on Sequence Variants

For each laboratory, reported variants were compared with the genetically engineered variants for both the wet and dry samples. If a variant was correctly identified, the result was classified as a true positive. If a variant was not identified, the result was classified as a false negative (FN); when a sequence variant was reported that was not the correct sequence variant for that position (ie, when a laboratory incorrectly identified an engineered variant), the error was also classified as a FN. Technically, this latter class of FN represents a "miscall" consisting of a paired FN and false positive (FP), but to avoid double counting "miscalls," they were classified as a FN only. When a laboratory indicated that the variant was not included in their LDT (variant not tested), the absence of a result for a variant not tested was not classified as an error. If a variant was reported at a site not involving an engineered sequence change, it was classified as a FP. The total numbers for each of these classifications were tracked by laboratory, variant type (SNV or MNV), sample type (dry or wet), and engineered VAF. To clarify classification of FNs, participant laboratories were contacted and asked whether they detected the variant, but it was below their LDT's limit of detection.

RESULTS

Aim 1: Traceable Reference Samples

Production of the 18 genetically engineered cell lines was completed in 14 months; some cell lines' growth characteristics extended the production of the wet samples. The relative cost to produce the wet samples associated with engineering the mutant cell lines, growth of the cell lines, mixing cell lines to achieve the targeted VAFs, and preparing FFPE blocks and sections was significantly greater than the dry sample cost for in silico mutagenesis and exchange of data files (current CRISPR technology may be more cost-effective than in 2017-2018). In silico mutagenesis of submitted sequence files was completed within 2 to 3 days. The final VAF of the wet sample cell line mixtures ranged from 5.1% to 8.6% TABLE 1, as confirmed with digital PCR by the vendor prior to shipment.

Some Pilot laboratories had difficulty managing sequence file uploads and downloads via the CAP portal, negotiating institutional firewalls, introducing external sequence files into bioinformatics pipelines, and defining file formats to ensure compatibility with bioinformatics pipelines. These challenges were resolved via consultation with CAP and/or P&V Licensing.

Aim 2: Analytic Challenge

Although 20 laboratories were selected for participation as Pilot laboratories, 2 laboratories withdrew after the reference samples were distributed, citing insufficient resources to evaluate so many samples, bringing the number of Pilot laboratories to 18. In addition, two of the Pilot laboratories completed only the wet sample testing; they communicated that they did not have sufficient bioinformatics expertise to manage the required in silico file manipulations of the dry samples.

Variant detection data from the 21 participating laboratories are presented in TABLE 2 illustrating the results reported by each laboratory, gene, variant, sample type, and VAF. Two trends are discernable; first, there was variability in the accuracy of variant detection among the LDTs, and second, there was an increased frequency of errors for MNVs. FIGURE 2 is a plot of the error frequency for individual laboratories.

The data show that 7 (37%) of 19 laboratories correctly reported all variants for both wet and dry samples; Illumina also correctly reported all variants for both wet and dry samples using the Praxis Extended Ras Panel (data not shown). Of the laboratories that tested both the wet and dry samples, 3 (16%) of 19 had fewer than five errors, and 9 (47%) of 19 had five or more errors. Statistical comparison of the differences between the accuracy of variant detection of the LDTs is difficult given that the model variants were not selected randomly and that the LDTs had differences in LOD and reporting

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

TABLE 2 Genotyping Results From Pilot Laboratories[a]

	Laboratory Number																				
***KRAS* Variants**	**3**	**4**	**10**	**12**	**19**	**20**	**21**	**2**	**5**	**9**	**16**	**18**	**7**	**8**	**13**	**15**	**6**	**14**	**11**	**1**	**17**
Wet lab samples																					
KRAS p.Lys117Asn c.351A>C																					
KRAS p.Ala59Thr c.175G>A																					
KRAS p.Lys117Asn c.351A>T																					
KRAS p.Gln61Arg c.182A>G																					
KRAS p.Gly12Asp c.35G>A																					
KRAS p.Ala146Pro c.436G>C																					
KRAS p.Gln61His c.183A>C																					
KRAS p.Gly13Glu c.38_39GC>AA																					
KRAS p.Gly12Trp c.34_36GGT>TGG																					
Dry lab samples, 5% VAF																					
KRAS p.Lys117Asn c.351A>C																					
KRAS p.Ala59Thr c.175G>A																					
KRAS p.Lys117Asn c.351A>T																					
KRAS p.Gln61Arg c.182A>G																					
KRAS p.Gly12Asp c.35G>A																					
KRAS p.Ala146Pro c.436G>C																					
KRAS p.Gln61His c.183A>C																					
KRAS p.Gly13Glu c.38_39GC>AA																					
KRAS p.Gly12Trp c.34_36GGT>TGG																					
Dry lab samples, 15% VAF																					
KRAS p.Lys117Asn c.351A>C																					
KRAS p.Ala59Thr c.175G>A																					
KRAS p.Gln61Arg c.182A>G																					
KRAS p.Gly12Asp c.35G>A																					
KRAS p.Ala146Pro c.436G>C																					
KRAS p.Gln61His c.183A>C																					
KRAS p.Thr127Ile c.380C>T																					
KRAS p.Gly13Glu c.38_39GC>AA																					
KRAS p.Gly12Trp c.34_36GGT>TGG																					
	Laboratory Number																				
***NRAS* Variants**	**3**	**4**	**10**	**12**	**19**	**20**	**21**	**2**	**5**	**9**	**16**	**18**	**7**	**8**	**13**	**15**	**6**	**14**	**11**	**1**	**17**
Wet lab samples																					
NRAS p.Gly12Val c.35G>T																					
NRAS p.Gly12Asp c.35G>A																					
NRAS p.Ala59Thr c.175G>A																					
NRAS p.Ala146Thr c.436G>A																					
NRAS p.Gln61His c.183A>C																					
NRAS p.Gln61Leu c.182A>T																					
NRAS p.Lys117Asn c.351G>T																					
NRAS p.Gly13Glu c.38_39GT>AA																					
NRAS p.Gly12Trp c.34_36GGT>TGG																					
Dry lab samples, 5% VAF																					
NRAS p.Ala146Thr c.436G>A																					
NRAS p.Gln61His c.183A>C																					
NRAS p.Gly12Asp c.35G>A																					
NRAS p.Lys117Asn c.351G>T																					
NRAS p.Gln61Leu c.182A>T																					
NRAS p.Ala59Thr c.175G>A																					
NRAS p.Gly12Val c.35G>T																					
NRAS p.Gly12Val c.35G>T																					
NRAS p.Gly13Glu c.38_39GT>AA																					
NRAS p.Gly12Trp c.34_36GGT>TGG																					
Dry lab samples, 15% VAF																					
NRAS p.Ala146Thr c.436G>A																					
NRAS p.Gly12Asp c.35G>A																					
NRAS p.Gln61Leu c.182A>T																					
NRAS p.Ala59Thr c.175G>A																					
NRAS p.Thr148Ser c.443C>G																					
NRAS p.Gly60Glu c.179G>A																					
NRAS p.Gly13Glu c.38_39GT>AA																					
NRAS p.Gly12Trp c.34_36GGT>TGG																					

VAF, variant allele fraction.

[a]Green = correct variant reported. Orange = incorrect variant reported (paired false negative and false positive), Red = no variant reported (false negative alone). Yellow = no variant reported (false negative; laboratory noted that the variant was identified but at a variant allele fraction below the limit of detection of the laboratory developed test).

(see below). Nonetheless, the Praxis Extended Ras Panel CDx provides a comparator for clinical NGS identification of the variants.[35,37] The CDx has a published positive percent agreement of 98.7% and negative percent agreement of 97.6%, which, for a set of 54 variants, corresponds to three or fewer FPs and miscalls, as well as fewer than five FNs and miscalls (one-tailed z test, $P < .05$).

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

Incidental Findings

While the utility of the reference samples for comparison of LDTs was the goal of aim 2 of the Pilot, the data collected from participating laboratories included some incidental findings germane to the use of the reference samples in clinical NGS testing.

Error Trends

Both the wet and dry samples demonstrated a high rate of FN results for some SNVs and/or MNVs for some laboratories TABLE 2, frequently associated with VAFs at or near the LOD of the LDT (see Minimum VAFs, below). Most errors were associated with MNV detection TABLE 2.

Minimum VAFs

Participating laboratories self-reported VAFs from 2% to 5% for SNVs and 3% to 10% for indels as the LOD values for their LDT NGS assays. Most laboratories, but not all, reported variants when the VAF was below the LOD of their LDT. The information collected from participating laboratories did not include their clinical threshold for reporting.

Spectrum of Variants Reported

Several LDTs did not include all variants included in the CDx TABLE 2. In addition, the dry samples included three variants (TABLE 1 and Supplementary Material S6) that have been detected in *KRAS* and *NRAS* but are not included in the CDx,[35] and thus the impact on response to targeted therapy by these variants is uncertain.[43,44] Only 3 (17%) of 18 or 2 (11%) of 19 laboratories that detected a non-CDx variant indicated that the variant would not be reported.

Neoplastic Cellularity

Pilot laboratories reported acceptance of samples with a minimum tumor cell content ranging from 10% to 30% for their LDTs. The cell lines in the wet samples were nominally 100% neoplastic, and thus an online digital slide image challenge was a surrogate measure of laboratories' preanalytic assessment of specimens.[45] Depending on the challenge, up to 22% of laboratories incorrectly estimated tumor cell content (Supplementary Materials S7 and S8).

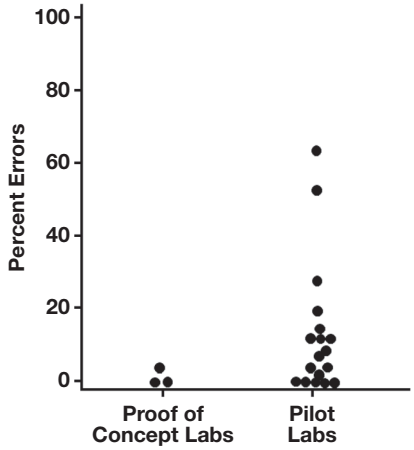

FIGURE 2 Error rates among laboratories. Data points represent the number of incorrectly reported variants divided by the number of variants included in the laboratory's laboratory-developed test (LDT). If a variant was not included in the LDT, lack of a correct reported variant was counted as an error. Lack of correct reported variants from a wet sample that failed DNA isolation or a dry sample for which a laboratory had insufficient bioinformatics capacity to complete the analyses were also not counted as errors.

DISCUSSION

In this Pilot model analytic challenge, traceable reference samples evaluated variability in the accuracy of genetic variant detection among laboratories that performed clinical NGS testing using LDTs. Differences among the LDTs (and a model CDx) included the LDT LOD, variants reported, comprehensiveness of testing, and required specimen tumor cellularity. The results suggested that some LDTs may not identify the same patient populations as candidates for targeted therapy as other LDTs or a CDx. Because the NGS analysis of the reference samples by participating laboratories occurred in December 2018 to March 2019 FIGURE 1, the variability in accuracy of variant detection identified by the Pilot represents a more contemporary interlaboratory comparison of clinical NGS than published PT results.[12-16]

The Pilot is a novel model for interlaboratory comparison of clinical NGS in several respects. The Pilot specified variants and VAFs representative of a disease-specific test to model an actual clinical encounter. An external, independent comparator with published performance metrics modeled the accuracy achievable with appropriate resources, expertise, and test validation. In addition, the results were assessed by individual LDTs, consistent with patient care decisions based on findings from a single laboratory. Because the Pilot differs from conventional laboratory PT in these fundamental ways, it is difficult to directly compare the results of the Pilot with those from formal PT programs. However, a number of advantages and disadvantages of the Pilot compared with conventional PT are nonetheless apparent. To begin with, an analytic challenge designed for testing within a specific clinical setting is useful to model an actual patient encounter but presents issues of scalability. It may be impractical to evaluate all clinical NGS on a disease-by-disease basis because such a paradigm of laboratory testing for somatic variants would require independent assessment of the same LDT for each of hundreds of different tumor types and treatment scenarios. The conventional PT approach of unrelated variants without context to a specific disease process may prove helpful to streamline some aspects of comparisons between laboratories. Next, while the Pilot shows that engineered cell lines isogenic except for the introduced variant are useful for interlaboratory comparisons, these standards are costly and time-consuming to produce (although both costs and the time required continue to decrease). The use of so-called spike-in standards to produce wet laboratory samples with a broader range of variants at a wider range of VAFs in formal PT has proven to be an advantageous alternative,[13] although the biologic equivalence of "spike-in" approaches remains uncertain. Similarly, the use of in silico mutagenized sequence files offers the advantage of rapid and cost-effective production of dry standards with a unlimited number, type, and combination of variants, at any VAF, but their utility is limited to evaluation of the bioinformatics component

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

of NGS.[19,21,29-31] In addition, the Pilot highlights that the utility of in silico standards is also currently disadvantaged by the lack in some clinical NGS laboratories of sufficient expertise to manage the required sequence file uploads and downloads. Finally, pooled results from a very large number of clinical NGS laboratories as provided by published reports of formal PT programs provide a useful overview of the landscape of clinical NGS in some detail.[12-16] However, evaluation of results on an LDT-by-LDT basis, as in this Pilot, has the advantage of providing the opportunity to discover systematic errors in some laboratories that may be undiscernible in aggregate data. The individual laboratory data analysis modeled by the Pilot may enable a more targeted approach of quality assurance activities to where improvement is most needed.

Aim 1: Traceable Reference Samples

Direct statistical comparison of wet vs dry samples was not possible because the cell lines could not be mixed to a target VAF with the precision achievable via computerized mutagenesis of sequence files, but the data suggest these two types of samples are complementary for evaluation of clinical NGS test performance (Supplementary Material S9).

Wet samples permit the evaluation of the total NGS testing process from nucleic acid extraction through results reporting. Engineered cell lines are more sustainable than residual clinical specimens as reference samples, may be used fresh or as an FFPE cell block, and may serve as intra- and interlaboratory standards. Such reagents have been extensively used by government agencies, professional organizations, and commercial entities for implementation in NGS test validation and quality control.[46-48] However, as this Pilot shows, engineered cell lines can be costly and time-consuming to develop and propagate; it took 14 months to develop a panel of engineered cell lines for a selected set of just 18 SNVs and small indels for only two genes. The requirement for sequence variants to be inserted into the same parental cell line at the endogenous genomic locations avoided possible problems caused by bioinformatics pipelines recognizing a mixture of genomic backgrounds, with subsequent rejection of the specimen as contaminated.[40-42] This approach also facilitated production of mixtures harboring several variants, with defined VAFs, without sequence artifacts associated with introduction of the variants, but contributed to the expense and time required. Other approaches for generating wet samples with variants at predefined allelic ratios have been described that are less expensive and time-consuming, including addition of chemically synthesized or plasmid-based nucleic acids to reference cell lines or DNA standards.[13,25-28] However, these types of reference samples may not represent the preanalytical conditions of specimen handling, nucleic acid isolation, bioinformatics, and range of variants and variant combinations encountered in routine clinical NGS workflows.[13,21,29]

Dry samples of the type used in this Pilot, produced by customized in silico mutagenesis of sequence files, may be economically and rapidly produced to facilitate more flexible design of sequence variant types, of various mixtures, and with different VAFs, as demonstrated by this Pilot and previous projects.[19,29-31] However, their utility is limited to assessment of the bioinformatics component of an NGS test, as has been emphasized.[19,21,29-31] The use of in silico reference samples requires a well-developed information technology infrastructure to support sequence file transfers between participating laboratories and the entity performing the computerized mutagenesis. The Pilot demonstrates that a subset of laboratories does not currently have the bioinformatics expertise to easily manage the file transfers or import mutagenized data files into their laboratory's bioinformatics pipeline for sequence variant detection and analysis, a finding consistent with prior observations.[19,20,29] Formally, for the subset of Pilot laboratories that were not facile with the use of file transfers or import of data files, there is uncertainty as to whether the LDT results for the dry samples are a reliable measure of the LDT's accuracy, although the data in TABLE 2 indicate that there was no systematic difference in performance between wet and dry samples for any participant laboratory. In addition, formal PT based on in silico mutagenized NGS sequence files has been offered commercially for several years,[49] indicating that while the logistics of the in silico approach can be problematic, they can routinely be overcome by clinical NGS laboratories. Nonetheless, engagement with NGS platform manufacturers and bioinformatics vendors to simplify the process of inserting external data files for performance quality measurements might remedy these issues.

Aim 2: Analytic Challenge

The LDTs from a subset of laboratories (7/19; 37%) correctly identified all variants in the reference samples, indicating that clinical NGS laboratories with appropriate resources and expertise can achieve a very high level of performance.

Variable accuracy in detection and interpretation of MNVs compared with SNVs was evident, consistent with prior studies.[16,18,50,51] Several laboratories also failed to correctly identify some SNVs, which is a novel observation, and the percentage of FNs and incorrect calls is significantly higher than previously reported.[12-16] Two differences between the Pilot and prior studies might explain this divergence. First, the Pilot focused on low VAFs near the reported LOD of the LDTs of the Pilot laboratories; other studies focused on higher VAFs, generally from 10% to 50%. Second, the Pilot employed a diversity of *RAS* variants, while prior studies evaluated a limited set of variants.

An important caveat to the data is that the potential clinical significance of variant identification errors was not evaluated. In the context of the Pilot, the presence of any of the 56 *RAS* gene variants included in the CDx would exclude the patient from panitumumab therapy (ie, the NGS test is performed to rule out therapy). Consequently, a FP could rule out panitumumab therapy, while a FN might have a detrimental effect on the treatment efficacy or clinical outcome.[52,53] The consequence of an incorrect result is likely to be different in other settings, for example, when identification of a specific variant is required for inclusion of a patient for targeted therapy.

Similarly, the frequency of di- and trinucleotide sequence changes vs SNVs in both the wet and dry samples in the Pilot was increased relative to the frequency of MNVs vs SNVs in *KRAS* and

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

NRAS exons 2, 3, and 4 in clinical cases of colorectal carcinoma.[43,44] While the Pilot's results reveal that a subset of LDTs cannot reliably detect some MNVs and SNVs, the results do not provide a direct measure of the proportion of patients with colorectal cancer who might be affected in clinical practice by these errors.

Incidental Findings

While not the focus of the Pilot, given their likely impact on patient care, the incidental findings indicate aspects of clinical NGS that may warrant further systematic evaluation.

Minimum VAFs

Information related to assay validation characteristics collected from participating laboratories did not capture details used to establish the LOD or the clinical threshold (CT) for reporting for each LDT. Although many laboratories reported variants that were detected at VAFs below the validated LOD for their respective LDT, at least 3 (14%) of 21 laboratories did not report variants based on the measured VAF established for their LDT when the VAF was below the LOD of their LDT. The approach of this latter group of laboratories would not be concordant with the scheme of the CDx (in which a variant is reported if its VAF is below the LOD of the test as long as the VAF is above the CT of the test). The Pilot did not collect data about interpretive comments that laboratories may include with clinical reports to explain variants with a VAF below the LOD of their LDT.

The LOD (ie, the minimum VAF) for which an assay is validated and the approach used to report variants detected below the LOD have become increasingly significant clinical issues as the extent of intratumoral heterogeneity and the impact of intratumoral heterogeneity in response to therapy have become increasingly clear.[54,55] Although the minimum tumor cellularity of the Praxis Extended Ras Panel CDx is 50%,[35] the range of tumor cellularity in the sample requirements of the LDTs performed by participating laboratories in the Pilot ranged from 10% to 30%. Because VAFs of 5% (or even lower) will be routinely encountered from clinical specimens (Supplementary Material S10), the lack of clarity regarding how laboratories report variants with VAFs below the LOD of their test is a significant finding and suggests areas for further evaluation to ensure quality NGS testing.

Variants Not Included in the CDx

The dry specimens in the Pilot included three sequence variants not targeted by the CDx, which would therefore not disqualify a patient for panitumumab therapy. The finding that only 11% to 17% of LDTs included a reporting scheme for these three variants aligned with FDA labeling for drug use is a novel observation and may be another source of variability among LDTs.

Neoplastic Cellularity

The finding that some laboratories were unable to correctly estimate tumor percentage in the range relevant to this Pilot is consistent with prior published results[45] and is problematic because tumor cellularity affects the sensitivity of NGS for identifying somatic variants. However, whole-slide estimates of tumor cell content as performed in the Pilot may not reflect the relative tumor cell content achieved by microdissection or coring of FFPE tissue samples, tumor cell enrichment steps that many clinical laboratories employ prior to nucleic acid preparation.

Limitations of the Pilot Study

The SPOT/Dx Pilot study had several limitations. First, the study focused on low VAFs and included a few rare sequence variants. This range of VAFs and variants may not represent those encountered in routine clinical practice. Second, the study consisted of a relatively small number of nonrandomly selected laboratories. The POC laboratories were chosen based on recognized experience and expertise in clinical NGS; the Pilot laboratories were selected to represent a cross section of different LDTs in different practice settings from laboratories that volunteered in response to a solicitation from CAP's NGS proficiency testing survey participants. These laboratories and their respective LDTs may not represent the broader landscape of laboratories that perform clinical NGS in CLIA-certified environments. Third, due to the small sample size, subgroup analysis was not feasible to determine whether specific patterns of errors were dependent on specific characteristics of LDTs. Fourth, the CDx was not commercially available in the United States at the time of the study; therefore, LDT performance could not be directly compared with the CDx by Pilot laboratories. Fifth, the SPOT/Dx Pilot design with wet samples and dry samples representing nominally 100% tumor cellularity did not permit direct evaluation of laboratories' determination of neoplastic cellularity of tissue samples. Sixth, the Pilot did not collect content from participating laboratories that would have been included in actual clinical reports. Additional information on problematic findings (eg, of a variant with a VAF below an LDT's LOD or of a sequence change outside the 56 specific mutations of the CDx) might be provided in the interpretive comments of clinical laboratory reports, which might mitigate some of the reporting issues identified by the Pilot.

CONCLUSIONS

The SPOT/Dx Pilot study demonstrated that reference samples can be developed to support an analytic challenge to evaluate interlaboratory performance of NGS LDTs. It is likely that these samples will also have utility for directly comparing the performance of NGS LDTs with companion diagnostic assays. While both wet samples and dry samples have advantages and disadvantages, the Pilot's results illustrate strategies for reference sample development that may be useful to measure comparable performance of LDTs and CDxs in a pragmatic, efficient, and sustainable manner.

The Pilot showed differences among NGS LDTs with respect to the correct identification of SNVs and MNVs, both with respect to one another and to the published accuracy of the model CDx. The pace of scientific advancement, the timelines required to obtain FDA approval for a CDx, and the intrinsic pairing of most CDxs with a narrow range of clinical interventions limit the availability of CDxs to encompass most routine clinical NGS testing, which will necessarily therefore be performed by LDTs. These

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

aspects of clinical NGS emphasize the need for reference standards to assist laboratories in the optimization and validation of LDTs, as well as for evaluation of performance among clinical NGS laboratories.

The differences in preanalytic variables, variant identification, and reporting indicate that the LDTs of participating laboratories for SNVs and MNVs may not be interchangeable with an FDA-approved CDx for identification of patients as candidates for gene-targeted therapy. Given that current clinical NGS testing encompasses a broader range of sequence variations (including CNVs, SVs, microsatellite instability, and tumor mutational burden) as well as a more extensive range of assay designs (including analysis of cell-free DNA, RNA sequence analysis, and even the whole genome[56]), the results of this Pilot suggest that additional reference sample designs and approaches to laboratory comparison will likely have utility for evaluation of laboratory performance more broadly.

Acknowledgments: The willingness of Illumina to share their experience and expertise with the Praxis Extended RAS Panel CDx is gratefully acknowledged. Illumina's participation in the POC component of the Pilot via testing of both the wet and dry reference samples in the Praxis CDx and in troubleshooting the logistics of the dry reference samples is also very much appreciated. The substantial support of Amgen throughout all phases of the Quality Assurance Pilot made this work possible.

REFERENCES

1. Berger MF, Mardis ER. The emerging clinical relevance of genomics in cancer medicine. *Nat Rev Clin Oncol*. 2018;15:353-365.
2. Roy S, Coldren C, Karunamurthy A, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J Mol Diagn*. 2018;20:4-27.
3. Gullapalli RR. Evaluation of commercial next-generation sequencing bioinformatics software solutions. *J Mol Diagn*. 2020;22:147-158.
4. US Food and Drug Administration. Considerations for design, development, and analytical validation of next generation sequencing (NGS)–based in vitro diagnostics (IVDs) intended to aid in the diagnosis of suspected germline diseases guidance for stakeholders and food and drug administration staff document issued April 13, 2018. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-design-development-and-analytical-validation-next-generation-sequencing-ngs-based. Accessed June 21, 2021.
5. New York Department of Health. Next generation sequencing (NGS) guidelines for somatic genetic variant detection. https://www.wadsworth.org/sites/default/files/WebDoc/3NextGenSeqONCOGuidelines%2012318.pdf. Accessed June 21, 2021.
6. Jennings LJ, Arcila ME, Corless C, 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;19:341-365.
7. Santani A, Simen BB, Briggs M, et al. Designing and implementing NGS tests for inherited disorders: a practical framework with step-by-step guidance for clinical laboratories. *J Mol Diagn*. 2019;21:369-374.
8. Hume S, Nelson TN, Speevak M, et al; Canadian College of Medical Geneticists (CCMG). CCMG practice guideline: laboratory guidelines for next-generation sequencing. *J Med Genet*. 2019;56:792-800.
9. Rehm HL, Bale SJ, Bayrak-Toydemir P, et al; Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Committee. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med*. 2013;15:733-747.
10. Aziz N, Zhao Q, Bry L, et al. College of American Pathologists' laboratory standards for next-generation sequencing clinical tests. *Arch Pathol Lab Med*. 2015;139:481-493.
11. Matthijs G, Souche E, Alders M, et al; EuroGentest; European Society of Human Genetics. Guidelines for diagnostic next-generation sequencing. *Eur J Hum Genet*. 2016;24:2-5.
12. Merker JD, Devereaux K, Iafrate AJ, et al. Proficiency testing of standardized samples shows very high interlaboratory agreement for clinical next-generation sequencing-based oncology assays. *Arch Pathol Lab Med*. 2019;143:463-471.
13. Keegan A, Bridge JA, Lindeman NI, et al. Proficiency testing of standardized samples shows high interlaboratory agreement for clinical next generation sequencing-based hematologic malignancy assays with survey material-specific differences in variant frequencies. *Arch Pathol Lab Med*. 2020;144:959-966.
14. Moncur JT, Bartley AN, Bridge JA, et al. Performance comparison of different analytic methods in proficiency testing for mutations in the BRAF, EGFR, and KRAS genes: a study of the College of American Pathologists Molecular Oncology Committee. *Arch Pathol Lab Med*. 2019;143:1203-1211.
15. Surrey LF, Oakley FD, Merker JD, et al. Next-generation sequencing (NGS) methods show superior or equivalent performance to non-NGS methods on BRAF, EGFR, and KRAS proficiency testing samples. *Arch Pathol Lab Med*. 2019;143:980-984.
16. Kim AS, Bartley AN, Bridge JA, et al. Comparison of laboratory-developed tests and FDA-approved assays for BRAF, EGFR, and KRAS testing. *JAMA Oncol*. 2018;4:838-841.
17. Nagarajan R, Bartley AN, Bridge JA, et al. A window into clinical next-generation sequencing-based oncology testing practices. *Arch Pathol Lab Med*. 2017;141:1679-1685.
18. Davies KD, Farooqi MS, Gruidl M, et al. Multi-institutional FASTQ file exchange as a means of proficiency testing for next-generation sequencing bioinformatics and variant interpretation. *J Mol Diagn*. 2016;18:572-579.
19. Duncavage EJ, Abel HJ, Merker JD, et al. A model study of in silico proficiency testing for clinical next-generation sequencing. *Arch Pathol Lab Med*. 2016;140:1085-1091.
20. Gutowska-Ding MW, Deans ZC, Roos C, et al. One byte at a time: evidencing the quality of clinical service next-generation sequencing for germline and somatic variants. *Eur J Hum Genet*. 2020;28:202-212.
21. Hardwick SA, Deveson IW, Mercer TR. Reference standards for next-generation sequencing. *Nat Rev Genet*. 2017;18:473-484.
22. Suzuki T, Tsukumo Y, Furihata C, et al. Preparation of the standard cell lines for reference mutations in cancer gene-panels by genome editing in HEK 293 T/17 cells. *Genes Environ*. 2020;42:42-48.
23. Coriell Institute for Medical Research. National Institute of Standards and Technology (NIST) reference materials: Genome in a Bottle Consortium. https://www.coriell.org/1/NIGMS/Collections/NIST-Reference-Materials. Accessed June 21, 2021.
24. Blackburn J, Wong T, Madala BS, et al. Use of synthetic DNA spike-in controls (sequins) for human genome sequencing. *Nat Protoc*. 2019;14:2119-2151.
25. Kudalkar EM, Almontashiri NA, Huang C, et al. Multiplexed reference materials as controls for diagnostic next-generation sequencing: a pilot investigating applications for hypertrophic cardiomyopathy. *J Mol Diagn*. 2016;18:882-889.
26. Sims DJ, Harrington RD, Polley EC, et al. Plasmid-based materials as multiplex quality controls and calibrators for clinical next-generation sequencing assays. *J Mol Diagn*. 2016;18:336-349.

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

27. Baum PD, Young JJ, Zhang Q, et al. Design, construction, and validation of a modular library of sequence diversity standards for polymerase chain reaction. *Anal Biochem.* 2011;411:106-115.
28. Zook JM, Samarov D, McDaniel J, et al. Synthetic spike-in standards improve run-specific systematic error analysis for DNA and RNA sequencing. *PLoS One.* 2012;7:e41356.
29. Duncavage EJ, Abel HJ, Pfeifer JD. In silico proficiency testing for clinical next-generation sequencing. *J Mol Diagn.* 2017;19:35-42.
30. Patil SA, Mujacic I, Ritterhouse LL, et al. insiM: in silico mutator software for bioinformatics pipeline validation of clinical next-generation sequencing assays. *J Mol Diagn.* 2019;21:19-26.
31. Escalona M, Rocha S, Posada D. A comparison of tools for the simulation of genomic next-generation sequencing data. *Nat Rev Genet.* 2016;17:459-469.
32. Tapestry Networks. Healthcare. https://www.tapestrynetworks.com/our-work/healthcare. Accessed June 21, 2021.
33. Tapestry Networks. SPOT/Dx working group. https://www.tapestrynetworks.com/our-work/healthcare/spotdx-working-group. Accessed June 21, 2021.
34. Tapestry Networks. Diagnostic quality assurance pilot. https://www.tapestrynetworks.com/our-work/healthcare/diagnostic-quality-assurance-pilot. Accessed June 21, 2021.
35. Illumina. Praxis extended RAS panel package insert. https://www.accessdata.fda.gov/cdrh_docs/pdf16/p160038c.pdf. Accessed June 21, 2021.
36. Zehnbauer B, Lofton-Day C, Pfeifer J, et al. Diagnostic quality assurance pilot: a model to demonstrate comparative laboratory test performance with an oncology companion diagnostic assay. *J Mol Diagn.* 2017;19:1-3.
37. Sorich MJ, Wiese MD, Rowland A, et al. Extended RAS mutations and anti-EGFR monoclonal antibody survival benefit in metastatic colorectal cancer: a meta-analysis of randomized, controlled trials. *Ann Oncol.* 2015;26:13-21.
38. Benson AB, Venook AP, Al-Hawary MM, et al. NCCN clinical practice guidelines in oncology; colon cancer. Version 2.2021. https://www.nccn.org/. Accessed June 21, 2021.
39. Essletzbichler P, Konopka T, Santoro F, et al. Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. *Genome Res.* 2014;24:2059-2065.
40. Cibulskis K, McKenna A, Fennell T, et al. ContEst: estimating cross-contamination of human samples in next-generation sequencing data. *Bioinformatics.* 2011;27:2601-2602.
41. Sehn JK, Spencer DH, Pfeifer JD, et al. Human specimen admixture in clinical next-generation sequencing data. *Am J Clin Pathol.* 2015;144:667-674.
42. Fiévet A, Bernard V, Tenreiro H, et al. ART-DeCo: easy tool for detection and characterization of cross contamination of DNA samples in diagnostic next-generation sequencing analysis. *Eur J Hum Genet.* 2019;27:792-800.
43. Haigis KM. KRAS alleles: the devil is in the detail. *Trends Cancer.* 2017;3:686-697.
44. Wellcome Sanger Institute. Catalog of somatic mutations in cancer. https://cancer.sanger.ac.uk/cosmic. Accessed June 21, 2021.
45. Viray H, Li K, Long TA, et al. A prospective, multi-institutional diagnostic trial to determine pathologist accuracy in estimation of percentage of malignant cells. *Arch Pathol Lab Med.* 2013;137:1545-1549.
46. Centers for Disease Control and Prevention. Clinical Laboratory Improvement Amendments (CLIA): genetic testing reference materials coordination program (GeT-RM). http://wwwn.cdc.gov/clia/Resources/GetRM/default.aspx. Accessed June 21, 2021.
47. National Institute of Standards and Technology. National Measurement Laboratory: standard reference materials. https://www.nist.gov/srm. Accessed June 21, 2021.
48. Medical Device Innovation Consortium. Cancer genomic somatic reference samples. https://mdic.org/project/cancer-genomic-somatic-reference-samples/. Assessed June 21, 2021.
49. College of American Pathologists proficiency testing. https://www.cap.org/laboratory-improvement/proficiency-testing. Accessed June 21, 2021.
50. Ghoneim DH, Myers JR, Tuttle E, et al. Comparison of insertion/deletion calling algorithms on human next-generation sequencing data. *BMC Res Notes.* 2014;7:864.
51. Wakeling MN, Laver TW, Colclough K, et al. Misannotation of multiple-nucleotide variants risks misdiagnosis. *Wellcome Open Res.* 2019;4:145.
52. Douillard JY, Oliner KS, Siena S, et al. Panitumumab-FOLFOX4 treatment and RAS mutations in colorectal cancer. *N Engl J Med.* 2013;369:1023-1034.
53. Bokemeyer C, Köhne CH, Ciardiello F, et al. FOLFOX4 plus cetuximab treatment and RAS mutations in colorectal cancer. *Eur J Cancer.* 2015;51:1243-1252.
54. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell.* 2017;168:613-628.
55. Vitale I, Shema E, Loi S, et al. Intratumoral heterogeneity in cancer progression and response to immunotherapy. *Nat Med.* 2021;27:212-224.
56. Duncavage EJ, Schroeder MC, O'Laughlin M, et al. Genome sequencing as an alternative to cytogenetic analysis in myeloid cancers. *N Engl J Med.* 2021;384:924-935.

HTTPS://DOI.ORG/10.1093/AJCP/AQAB164

HOLOGIC®

ThinPrep® Processors

Meeting the Demands of Your Lab

THINPREP® GENESIS™ PROCESSOR

THINPREP® 5000 PROCESSOR

THINPREP® 5000 AUTOLOADER

ThinPrep® processors are versatile and scalable solutions for laboratories of all volumes.

The portfolio offers reliable, best-in-class systems that help cytology labs automate their processes – decreasing the need for hands-on labor and increasing laboratory efficiencies.

Explore the portfolio

ThinPrep®