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

Joel T. Moncur, MD, PhD; Angela N. Bartley, MD; Julia A. Bridge, MD; Suzanne Kamel-Reid, PhD; Alexander J. Lazar, MD, PhD; Neal I. Lindeman, MD; Thomas A. Long, MPH; Jason D. Merker, MD, PhD; Alex J. Rai, PhD; David L. Rimm, MD, PhD; Paul G. Rothberg, PhD; Patricia Vasalos, BS; Annette S. Kim, MD, PhD

• ***Context.***—**The performance of laboratory testing has recently come under increased scrutiny as part of important and ongoing debates on regulation and reimbursement. To address this critical issue, this study compares the performance of assay methods, using either commercial kits or assays designed and implemented by single laboratories ("home brews"), including next-generation sequencing methods, on proficiency testing provided by the College of American Pathologists Molecular Oncology Committee.**

Objective.—**To compare the performance of different assay methods on College of American Pathologists proficiency testing for variant analysis of 3 common oncology analytes: *BRAF*, *EGFR*, and *KRAS*.**

Design.—**There were 6897 total responses across 35 different proficiency testing samples interrogating 13 different variants as well as wild-type sequences for *BRAF*, *EGFR*, and *KRAS*. Performance was analyzed by test method, kit manufacturer, variants tested, and preanalytic and postanalytic practices.**

Results.—**Of 26 reported commercial kits, 23 achieved greater than 95% accuracy. Laboratory-developed tests with no kit specified demonstrated 96.8% or greater accuracy across all 3 analytes (1123 [96.8%] acceptable of 1160 total responses for *BRAF*; 848 [97.5%] acceptable of 870 total responses for *EGFR*; 942 [97.0%] acceptable of 971 total responses for *KRAS*). Next-generation sequencing platforms (summed across all analytes and 2 platforms) demonstrated 99.4% accuracy for these analytes (165 [99.4%] acceptable of 166 total next-generation sequencing responses). Slight differences in performance were noted among select commercial assays, dependent upon the particular design and specificity of the assay. Wide differences were noted in the lower limits of neoplastic cellularity laboratories accepted for testing.**

Conclusions.—**These data demonstrate the high degree of accuracy and comparable performance across all laboratories, regardless of methodology. However, care must be taken in understanding the diagnostic specificity and reported analytic sensitivity of individual methods.**

(*Arch Pathol Lab Med*. 2019;143:1203–1211; doi: 10.5858/arpa.2018-0396-CP)

The performance of laboratory testing has recently come under increased scrutiny, with ongoing debates on the role of the US Food and Drug Administration (FDA) in the

Accepted for publication December 17, 2018.
Published online April 10, 2019.
Supplemental digital content is available for this article at www.archivesofpathology.org in the October 2019 table of contents.

From the Office of the Director, The Joint Pathology Center, Silver Spring, Maryland (Dr Moncur); the Department of Pathology, St Joseph Mercy Hospital, Ypsilanti, Michigan (Dr Bartley); the Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha (Dr Bridge); the Department of Pathology, University Health Network, University of Toronto, Toronto, Ontario, Canada (Dr Kamel-Reid); the Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston (Dr Lazar); the Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts (Drs Lindeman and Kim); Biostatistics (Mr Long) and Proficiency Testing (Ms Vasalos), College of American Pathologists, Northfield, Illinois; the Department of Pathology and Laboratory Medicine, University of North Carolina, Chapel Hill (Dr Merker); the Department of Pathology and Cell Biology, Columbia University Medical Center, New York, New York (Dr Rai); the Department of Pathology, Yale University School of Medicine, New Haven, Connecticut (Dr Rimm); and the Department of Pathology and Laboratory Medicine, Strong Memorial Hospital, University of Rochester Medical Center and Pathology and Laboratory Medicine, Molecular Diagnostic Laboratory, Rochester, New York (Dr Rothberg).

Dr Moncur is employed by the US Army. The identification of specific products or scientific instrumentation is considered an integral part of the scientific endeavor and does not constitute endorsement or implied endorsement on the part of the authors, the Department of Defense, or any component agency. The views expressed in this article are those of the authors and do not reflect the official policy of the Department of Army/Navy/Air Force, Department of Defense, or US government. The other authors have no relevant financial interest in the products or companies described in this article.

All authors are current or past members of the College of American Pathologists Molecular Oncology Committee.

Corresponding author: Annette S. Kim, MD, PhD, Department of Pathology, Brigham & Women's Hospital, 75 Francis St, Thorn 613A, Boston, MA 02115 (email: askim@bwh.harvard.edu).

regulation of laboratory-developed tests (LDTs) as well as the impact of FDA approval on Medicare reimbursement.[1–17] The FDA has recommended greater oversight of all in vitro testing. Several FDA-approved companion diagnostics sold under the designation of in vitro diagnostics are available within the realm of molecular oncology testing. However, there are many more clinical LDTs that are designed, validated, and performed in a single laboratory. Proprietary commercial assays may also fall under the LDT umbrella. Under currently proposed regulations, LDTs would undergo greater FDA review.

In 2010, the FDA announced plans to reassess its role in the regulation of LDTs as in vitro diagnostics under FDA medical device regulation.[18] Two draft guidance documents in 2014 proposed to phase out the enforcement discretion policy that allowed Clinical Laboratory Improvement Amendments of 1988 oversight of all clinical laboratory testing. The FDA would phase in regulation, with the most complex and high-risk testing undergoing the same level of scrutiny as new medical devices with a full 510(k) premarket approval process.[19] The FDA expressed the intention to evaluate "both the analytic validity (eg, analytic specificity and sensitivity as well as accuracy and precision) and clinical validity (eg, positive and negative predictive value) of diagnostic tests through its premarket clearance or approval process."

This debate has been rejuvenated by the recent FDA approval of 2 next-generation sequencing (NGS) assays that previously fell under the category of LDTs.[20–23] The Centers for Medicare and Medicaid Services propose to reimburse NGS testing under specific circumstances, which include the criterion of FDA approval.[24,25]

Although debate is focused on both regulatory and performance aspects, a central question relates to the relative performance of the various methods of testing for a given analyte. There are a broad number of laboratories to which clinicians may send their testing requests. Similarly, a range of technical approaches, from single-analyte to panel (including NGS) testing, exists for laboratories to use, as well as a breadth of options for testing kits, including FDA-CDs and LDTs (non–FDA-approved commercial kits and home-brew LDTs). These selection decisions should be made in a data-driven fashion whenever possible.

The College of American Pathologists (CAP) is a well-recognized provider of external proficiency testing (PT) materials that provide a mechanism for laboratories to fulfill the Clinical Laboratory Improvement Amendments requirement for laboratories to assess the analytic validity of clinical assays during initial development and ongoing clinical use. Preanalytic, analytic, and postanalytic characteristics of the assay may also be assessed. The CAP Molecular Oncology Committee previously demonstrated that there was excellent overall performance (>97% acceptable results) of laboratories in the variant analysis of 3 oncology analytes: *BRAF*, *EGFR*, and *KRAS*.[26] In this study, specific comparisons of analytic performance, preanalytic and postanalytic practices, and reported sensitivity (lower limit of neoplastic cellularity) across a wide range of testing modalities are investigated.

METHODS

Sample Definition and Time Frame

All CAP PT for *BRAF* from series 2011B through series 2015A was included in this study (14 samples, 8 mailings, 2524 total responses, 2404 responses with methodology details). For *EGFR*, data from 2013A through 2015A were examined (11 samples, 5 mailings, 2216 total responses, 2176 responses with methodology details). For *KRAS*, only data from 2013B through 2015A were examined (10 samples, 4 mailings, 2157 total responses, 2112 responses with methodology details), because the specific option for respondents to indicate the use of the FDA-approved platform was not available until 2013B. For all 3 analytes, 2 PTs (A and B) were provided each year. During this time frame, PTs 2011B through 2013B contained only a single specimen. Beginning with 2014A, each mailing contained 3 specimens. In total, we included data from 14 samples for *BRAF*, 11 samples for *EGFR*, and 10 samples for *KRAS* (ranging from ≥50% to 100% neoplastic cellularity, the latter in cell lines carrying the variant). Samples for *BRAF* were composed of formalin-fixed, paraffin-embedded neoplastic tissue prior to the 2013A survey and formalin-fixed, paraffin-embedded cell lines thereafter. Samples for *EGFR* and *KRAS* were composed of formalin-fixed, paraffin-embedded cell lines for the entirety of the period of this study.

Data Definition

See Table 1 for specific variants assessed (see also supplemental digital content for a brief discussion of the clinical implications of the variants at www.archivesofpathology.org in the October 2019 table of contents). Every result was counted as an individual data point, so each laboratory could be represented by as many as 14 results for *BRAF*, 11 for *EGFR*, and 10 for *KRAS*.

Results on CAP PT are considered either good, acceptable, or unacceptable. These terms are defined differently depending upon the type of sample used for the PT. For tissue samples used prior to 2013, a good result (detected or not detected) was one that matched the consensus response, defined as concordance of 80% or greater of respondents. Beginning with the 2013A survey for *BRAF* (*EGFR* and *KRAS* surveys used cell lines for the duration of the study), when cell lines with defined mutations were introduced for the PTs, a good result was defined as the identification of the correct result.

The responses "does not detect" and "does not discriminate" were also considered acceptable. Some assays that interrogate, for instance, the *BRAF* c.1799T position cannot discriminate between single-nucleotide variants at this position and those that occur in cis with either c.1798G or c.1800G variants, resulting in different codon V600 amino acid alterations. Laboratories using an LDT that missed the p.V600K (c.1798_1799delGTinsAA) variant covered by this study were queried as to whether or not their assay could discriminate between these possibilities. In addition, because some platforms do not detect all possible variants, the PT response form included a "test not performed" option for assay methods that do not interrogate a particular variant. However, only laboratories that reported results for a specific variant were included in the subsequent studies on the specific variants; laboratories reporting "does not discriminate variants" or "test not performed" were excluded from further consideration. Therefore, if a laboratory reported "not detected" for a *BRAF* p.V600K variant when using an FDA-CD that does not discriminate p.V600K from p.V600E, this was counted in the analysis as unacceptable. For the purposes of this study, all good and acceptable results were considered acceptable and were included in the overall rates of acceptability that are used throughout the manuscript.

For the assessment of laboratory practice, select qualitative and quantitative preanalytic and postanalytic considerations were also surveyed. Of note, not all participants answered all questions; only respondents were considered in the analysis.

Method Definition

Based upon the participant responses, the testing methodology was categorized as either FDA-CD or LDT (see supplemental materials for the specifications of the FDA-CDs). Although attempts were made to determine if laboratories using a kit produced by a manufacturer with an FDA-CD were actually using

Table 1. Summary of the Specific Mutations Used in the Proficiency Testing Surveys for *BRAF*, *EGFR*, and *KRAS*

	Survey Series							
Gene	**2011B**	**2012A**	**2012B**	**2013A**	**2013B**	**2014A**	**2014B**	**2015A**
BRAF	p.V600E	WT	p.V600E	p.V600E	p.V600E	p.V600E	p.V600E	WT
	NA	NA	NA	NA	NA	p.V600K[a]	WT	p.V600E
	NA	NA	NA	NA	NA	WT	p.V600E	p.V600K[a]
EGFR	...	...	...	p.G719S	p.L858R	Exon 19 del[b]	p.L858R/p.T790M	WT
	...	...	...	NA	NA	WT	p.G719S	p.G719S
	...	...	...	NA	NA	p.L861Q	WT	exon 19 del[b]
KRAS	...	...	...	...	p.G13D	p.G12C	p.G12V	p.G12A
	...	...	...	...	NA	WT	p.G12S	p.G12R
	...	...	...	...	NA	p.G12S	p.G12R	p.G12C

Abbreviations: NA, only one sample was provided in that mailing; WT, the materials provided contained only the wild-type sequence at this locus; ..., these mailings were not included in the study.

[a] The responses "does not detect" and "does not discriminate" were counted as acceptable.

[b] Exon 19 del is c.2235_2249del15.

the FDA-CD or a research use–only version, too few laboratories responded to enable assignment of laboratories to the proper category. Therefore, for all 3 surveys, all laboratories using kits purchased from a vendor with an FDA-CD were considered in the FDA cohort for the purposes of this study, acknowledging that some laboratories (in particular for *EGFR* and *KRAS*) may have been using the research use–only version (and therefore should have been categorized as using an LDT). Laboratories using a kit manufactured by a vendor with FDA approval for that kit after the FDA approval date were analyzed as FDA-CDs. All other assays were analyzed as LDTs, but broken down by vendor whenever possible. In addition, the use of NGS for testing was inferred from participant responses that indicated use of reagents from NGS suppliers, such as Illumina, Truseq, Ion Torrent, or Ion Ampliseq.

Statistical Analysis

A χ^2 test of association compared acceptability across FDA-CDs and LDTs for all tests. For testing acceptability across FDA and LDT by sample type, Fisher exact test was used because of counts higher than 5 in at least one table cell. A significance level of .05 was used.

RESULTS

A total of 6897 laboratory responses for *BRAF*, *EGFR*, and *KRAS* CAP PT were recorded, of which 6692 were associated with a listed manufacturer (205 survey responses did not designate their methodology). During the time period of this study, there were 2 FDA-CDs each for *BRAF* and *EGFR* and 1 FDA-CD for *KRAS* reported. Next-generation sequencing reagents were reportedly used by a subset of laboratories for all 3 analytes.

BRAF Survey Results

As previously reported, the number of responding laboratories in the CAP *BRAF* PTs ranged from 123 to 212, for a total of 2524 responses (2224 LDTs, 300 FDA-CDs).[26] The overall *BRAF* CAP PT acceptable rate was 96.2%, with LDT results demonstrating 96.6% acceptability and FDA-CD results significantly lower at 93.0% ($P = .002$). In this previous publication,[26] all laboratories that failed to designate their methodology were grouped with the LDTs. In the current study, only laboratories that report their methodology were considered. Using these criteria, there were a total of 2404 responses (2104 LDTs; 300 FDA-CDs) with essentially identical results. Laboratory-developed tests demonstrated 96.6% (2032 of 2104) acceptability; FDA-CDs were lower at 93.0% (279 of 300) (Table 2), including the FDA-CDs Roche cobas *BRAF* assay (n = 282) and the bioMérieux THxID BRAF test (n = 18). The main cause of this discrepancy between LDT and FDA-CD performance was the p.V600K analysis, with LDT acceptable rates at 88.0% (271 of 308) whereas FDA-CDs were 66.1% acceptable (39 of 59; $P < .001$).[26] Laboratories using the response option "does not discriminate" were not included in this analysis (see Methods). Laboratories using LDTs that missed the p.V600K were queried (n = 27) individually as to whether their assay discriminates between p.V600E and p.V600K. Of the 7 respondents, 3 discriminated all variants at c.1799 and 4 did not. There was no significant difference in performance between LDTs and FDA-CDs for specimens with the common p.V600E mutation ($P = .25$) and wild-type sequence (no mutation, $P = .99$).[26] Of the 2 *BRAF* FDA-CDs included in this study, the Roche cobas assay reached a 92.6% acceptability rate (261 of 282), and the bioMérieux ThxID assay achieved a 100% acceptability rate (18 of 18) in a much smaller set of tested samples. The majority of participants used LDTs, and more than half of LDT laboratories (1160 of 2104) did not use specified commercial kit reagents ("No kit specified" in Table 2); as a group these laboratories showed 96.8% acceptability (1123 of 1160). For those LDT laboratories that did use reagents from a commercial kit, the range of acceptable results was 89% to 100%. Responses that corresponded to the use of either Illumina or Ion Torrent (NGS) reagents (n = 54) demonstrated 100% acceptability.

EGFR Survey Results

The number of responding laboratories ranged from 170 to 197 for a total of 2216 PT responses (1667 LDT; 549 FDA-CD), as previously reported.[26] The overall acceptable rate was 98.0%.[26] Laboratory-developed tests performed significantly less well than the FDA-CDs overall (97.6% acceptability for LDT versus 99.1% for FDA-CDs; $P = .03$). When only those that reported their methodology are considered, there were a total of 2176 responses (1627 LDTs; 549 FDA-CDs) with a 97.5% (1586 of 1627) acceptability rate for LDTs and a 99.1% (544 of 549) acceptability rate for FDA-CDs (Table 2). The total number and performance of FDA-CD laboratories remains unchanged but are now broken down by manufacturer: the Roche cobas *EGFR* v1 assay (n = 82) and the Qiagen Therascreen RGQ assay (n = 467). This discrepancy between LDT and FDA-CD laboratory performance was driven by the detection of the *EGFR*

Table 2. Percentage Acceptable Proficiency Testing Results of FDA-CD Versus LDT by Test Manufacturer for *BRAF*, *EGFR*, and *KRAS* (No. of Responses in Parentheses)

	BRAF	*EGFR*	*KRAS*	All Analytes
FDA-CD				
All	93.0 (300)	99.1 (549)	98.8 (331)	97.5 (1180)
bioMérieux	100 (18)	...	...	
Roche	92.6 (282)	100 (82)	...	
Qiagen Therascreen	...	98.9 (467)	98.8 (331)	
LDT				
All	96.6 (2104)	97.5 (1627)	97.5 (1781)	97.2 (5512)
No kit specified	96.8 (1160)	97.5 (870)	97.0 (971)	
Amoy Diagnostics	100 (25)	98.2 (56)	97.1 (34)	
ThermoFisher Scientific (Applied Biosystems)	95.0 (161)	100 (16)	...	
Asuragen	94.9 (39)	...	100 (62)	
Autogenomics	100 (11)	...	...	
BioSewoom	100 (35)	...	...	
DxS Therascreen	...	100 (4)	95.1 (82)	
EntroGen, Inc	89.4 (94)	91.2 (91)	96.3 (107)	
Illumina	100 (30)	100 (37)	95.5 (22)	
ThermoFisher Scientific (Ion Torrent)	100 (24)	100 (14)	100 (39)	
Medical & Biological Laboratories (MEBGEN)	...	...	96.9 (32)	
Qiagen	97.3 (150)	100 (53)	...	
Qiagen (Pyro)	98.5 (66)	97.3 (186)	99.5 (190)	
Qiagen (RGQ)	95.5 (156)	100 (51)	...	
Roche (cobas)	...	97.7 (130)	98.7 (74)	
Seegene	100 (2)	...	...	
Sequenom	97.4 (38)	100 (26)	100 (39)	
SNaPshot	...	...	100 (53)	
ThermoFisher Scientific (TaqMan)	97.7 (43)	...	...	
ThermoFisher Scientific (not otherwise specified)	...	98.1 (52)	...	
TIB Molbiol	100 (4)	...	...	
Trimgen	98.5 (66)	100 (41)	96.7 (60)	
ViennaLab	...	...	100 (16)	

Abbreviations: DxS, Diagnostics Innovations, distributed by Qiagen; FDA-CD, Food and Drug Administration–approved companion diagnostic; LDT, laboratory using a laboratory-developed test; PYRO, PyroMark polymerase chain reaction for pyrosequencing; RGQ, Rotor-Gene Q; ..., no commercial kit for this gene available from this specific manufacturer.

p.L861Q variant (90.7% ([117 of 129)] for LDTs versus 100% ([46 of 46)] for FDA-CDs; $P = .04$).[26] There were no significant differences for other variants (P values all $>.05$).[26] The 2 *EGFR* FDA-CDs performed well overall, with the Roche cobas assay reaching a 100% acceptability rate (82 of 82) and the Qiagen RGQ assay reaching a 98.9% acceptability rate (462 of 467). Again, the majority of participants used LDTs. Laboratory-developed test laboratories using a commercial kit demonstrated 91% to 100% acceptable results. More than half of LDT laboratories (870 of 1627) did not specify if they used commercial kit reagents ("No kit specified" in Table 2), and as a group these demonstrated 97.5% accuracy (848 of 870). Next-generation sequencing responses (n = 51) showed 100% acceptability.

KRAS Survey Results

The number of responding laboratories ranged from 237 to 282, for 2157 total responses for the CAP *KRAS* PT (1826 LDT; 331 FDA-CD Qiagen Therascreen RGQ PCR Kit). As reported previously, the overall acceptable rate was 97.6%, with a 97.4% acceptability for LDTs and 98.8% acceptability (327 of 331) for the Qiagen Therascreen *KRAS* FDA-CDs ($P = .16$, Fisher exact test; Table 2).[26] When only those that reported their methodology were considered, there were a total of 2112 responses (1781 LDTs, 331 FDA-CDs), with a 97.5% acceptability rate (1736 of 1781) for LDTs and a 98.8% acceptability rate (327 of 331) for FDA-CDs (Table 2). There was no significant difference between LDT and FDA-CD acceptability rates when the data were broken down by acceptable response rates for detection of a normal (no mutation) genotype or individual variants (P values all $>.1$).[26] As with the other 2 analytes, the majority of participants used LDTs. Laboratory-developed test laboratories using a commercial kit demonstrated 95% to 100% acceptable results. More than half of LDT laboratories (971 of 1781) did not indicate using commercial kit reagents ("No kit specified" in Table 2), and as a group these demonstrated 97.0% accuracy (942 of 971). Next-generation sequencing responses (n = 61) showed 98.4% acceptability (60 of 61).

Preanalytic and Postanalytic Practices

A summary of data for acceptable specimen preparations, acceptable tumor types, pathologist review of tissue, DNA quantification, and the use of tissue dissection has been previously reported.[26] Across all 3 tests, 97.9% of all respondents (6707 of 6849 total responses) had a pathologist review the tissue for testing at least some of the time, with no clear trend of either FDA-CD or LDT laboratories more

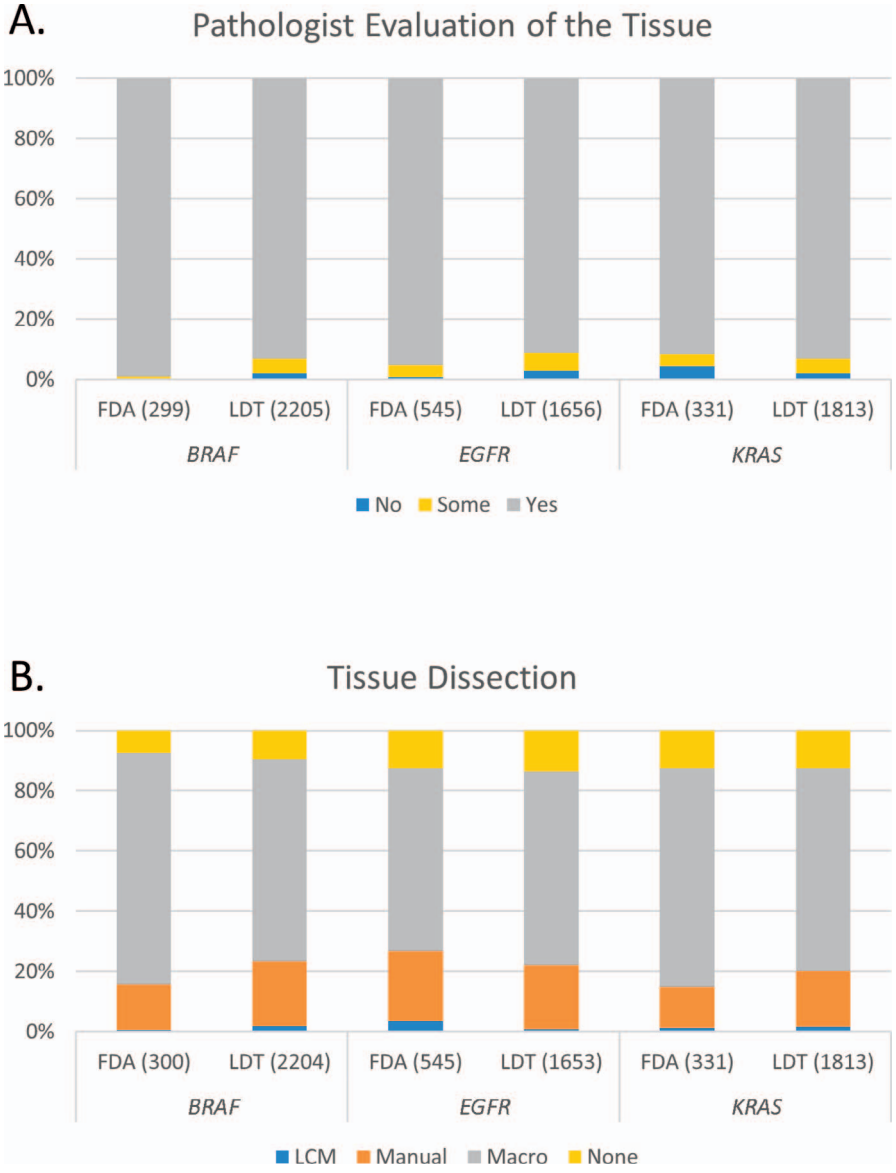

Figure 1. *A, Percentage of laboratories reporting that a pathologist evaluates the tissue morphologically prior to testing always (response = Yes), in some of the cases (response = Some), or never (response = No). Numbers in parentheses report the total number of responses (denominator). B, Percentage of laboratories reporting that the tissue is dissected prior to testing to enrich for neoplastic cells. Numbers in parentheses report the total number of responses (denominator). Abbreviations: FDA, laboratory using a Food and Drug Administration–approved companion diagnostic; LCM, laser capture microdissection; LDT, laboratory using a laboratory-developed test; Macro, macrodissection.*

consistently performing this important step across the 3 surveys (Figure 1, A). Similarly, 89.7% of all laboratories performed tissue dissection (6049 of 6846 total responses across all 3 assays) to enrich for tumor cells, with macrodissection and manual dissection as the most common methods and virtually no laboratories performing laser capture microdissection (Figure 1, B). Again, there was no clear trend for either FDA-CD or LDT laboratories performing this more consistently. Trends appeared more analyte specific than assay specific.

When asked about the minimum acceptable neoplastic cellularity (ie, lowest tumor cell fraction that can be detected by the assay assuming that all tumor cells carry the mutation), laboratory responses ranged from 0% to 90% when summed across all 3 analytes (Figure 2, A through C). In several cases, the FDA-CD laboratories reported minimum acceptable neoplastic cellularities that were below the reported minimum tumor content required for the assay by the manufacturer (indicated by the vertical lines in Figure 2). There was no specific trend for FDA-CD or LDT laboratories to have greater or lesser analytic sensitivity.

Across the 3 assays, 83.7% of laboratories (5721 of 6838 total responses) performed DNA quantification prior to testing, including a wide range of methods for both FDA-CD and LDT laboratories, with the exception that most laboratories using the *KRAS* FDA-CD (157 [47.4%] of 331 *KRAS* responses) did not quantify DNA as prescribed in the in vitro diagnostic instructions (Figure 3, A). The methods for DNA quantification varied extensively, with spectrophotometry, including specifically NanoDrop, as the most common method (Figure 3, B).

Postanalytically, 82.0% of all laboratories (1228 of 5612 total responses) provided an interpretive comment for their results, with no clear trend for laboratories performing an FDA-CD or LDT demonstrating superior practice (Figure 4). In summary, there were no consistent differences in preanalytic or postanalytic clinical practice between laboratories using FDA-CDs or LDTs.[27]

DISCUSSION

As previously reported, both FDA-CDs and LDTs exhibited similar and excellent performance, with both test

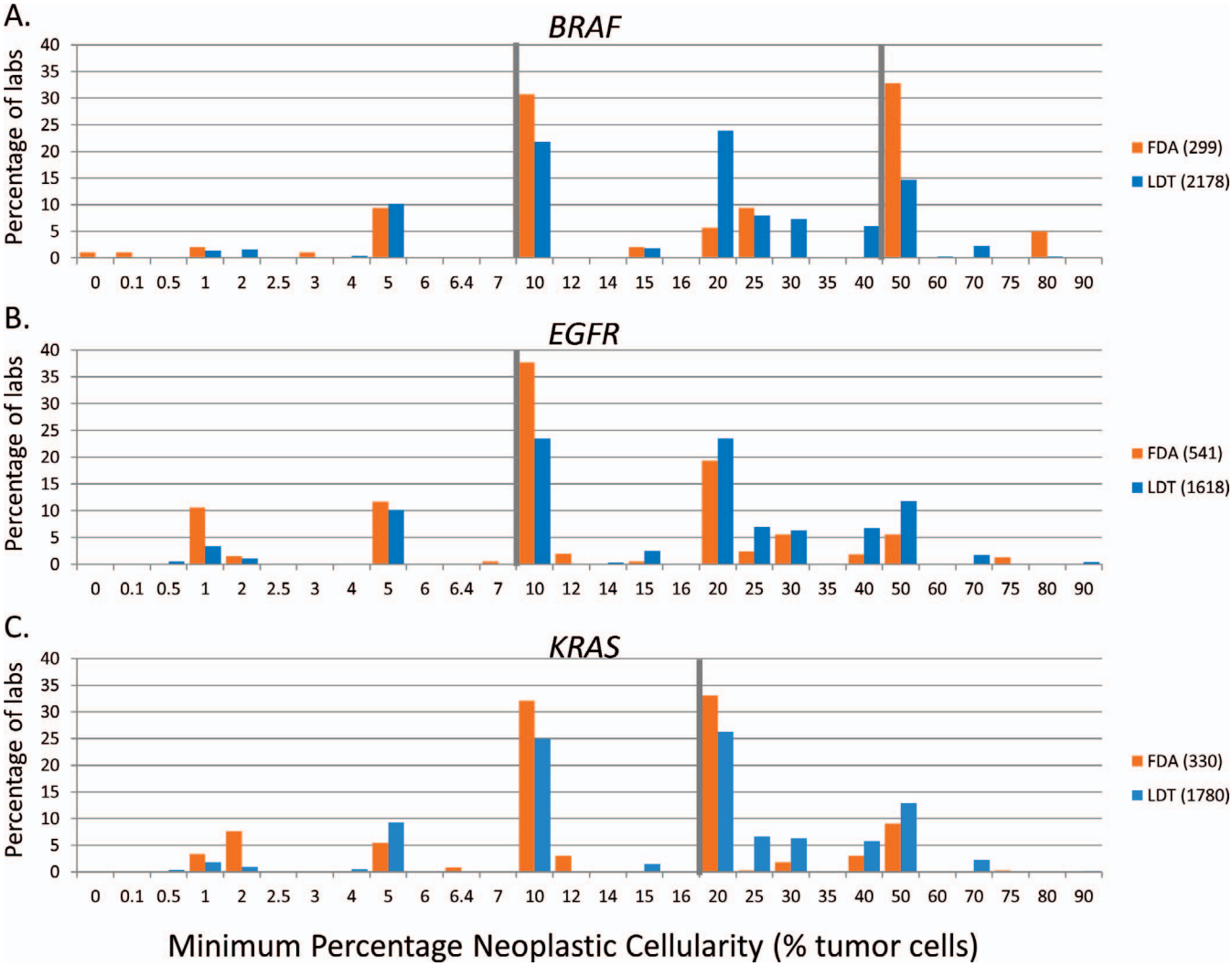

Figure 2. *Reported minimum percentage neoplastic cellularity accepted for testing. The gray vertical lines demarcate the stated minimum tumor content per the Food and Drug Administration companion diagnostic (FDA-CD) assay instruction manuals. There are 2 lines for the 2 FDA-CDs for* BRAF *(A) and 1 line for the 1* KRAS *FDA-CD (C). There is only 1 line for* EGFR *(B) because one assay does not specify a minimum tumor content for testing. Validated assay sensitivities for the second* EGFR *FDA-CD vary from locus to locus, with a range from 0.81% to 16.87% variant allele fraction (corresponding to 1.6% to 33.7% neoplastic cellularity). Abbreviations: FDA, laboratory using an FDA-CD; LDT, laboratory using a laboratory-developed test.*

types exceeding 97% accuracy overall on PT across all 3 analytes: *BRAF*, *EGFR*, and *KRAS* (Table 2).[26] These findings led us to question the comparability of individual test methods and assay manufacturers in this current study of the same data set, with enhanced granularity on individual FDA-CDs, NGS, and other commercial or home-brew LDTs.

As previously reported, for 11 of 13 variants across the 3 PT surveys, there were no significant differences between the acceptable rates of FDA-CD and LDTs. For 1 of 13 variants (*BRAF* p.V600K), LDTs performed statistically better than FDA-CDs, whereas for the second variant (*EGFR* p.L861Q), FDA-CDs performed statistically better than LDTs (Table 2; see also Kim et al[26]). It should be noted that across all 3 analytes, NGS results from multiple participants and across platforms demonstrated uniformly near-perfect accuracy (165 of 166; 99.4%) for the tested variants.

In the case of *BRAF* p.V600K, the LDT laboratories identified the acceptable/correct response in 88.0% of cases, compared with 66.1% of the FDA-CD laboratories ($P < .001$). The vast majority of FDA-CD laboratories used the Roche platform. This assay does not claim to distinguish between p.V600E and p.V600K and does not detect all instances of p.V600K because the assay is less sensitive for this variant. Because the PT provided a response option of "assay does not discriminate," the failure to select this option could reflect a lack of understanding of assay limitations. This problem was not unique to FDA-CD laboratories as a similar issue was reported by some LDT laboratories when queried. The use of assays that cannot distinguish between these 2 variants and have decreased ability to identify p.V600K is clinically problematic because *BRAF* p.V600K represents 15% to 20% of *BRAF* mutations in melanoma and is responsive to inhibitors. Missing this variant could result in failure to receive a therapeutically effective BRAF inhibitor.[28–32] Subsequent surveys since this

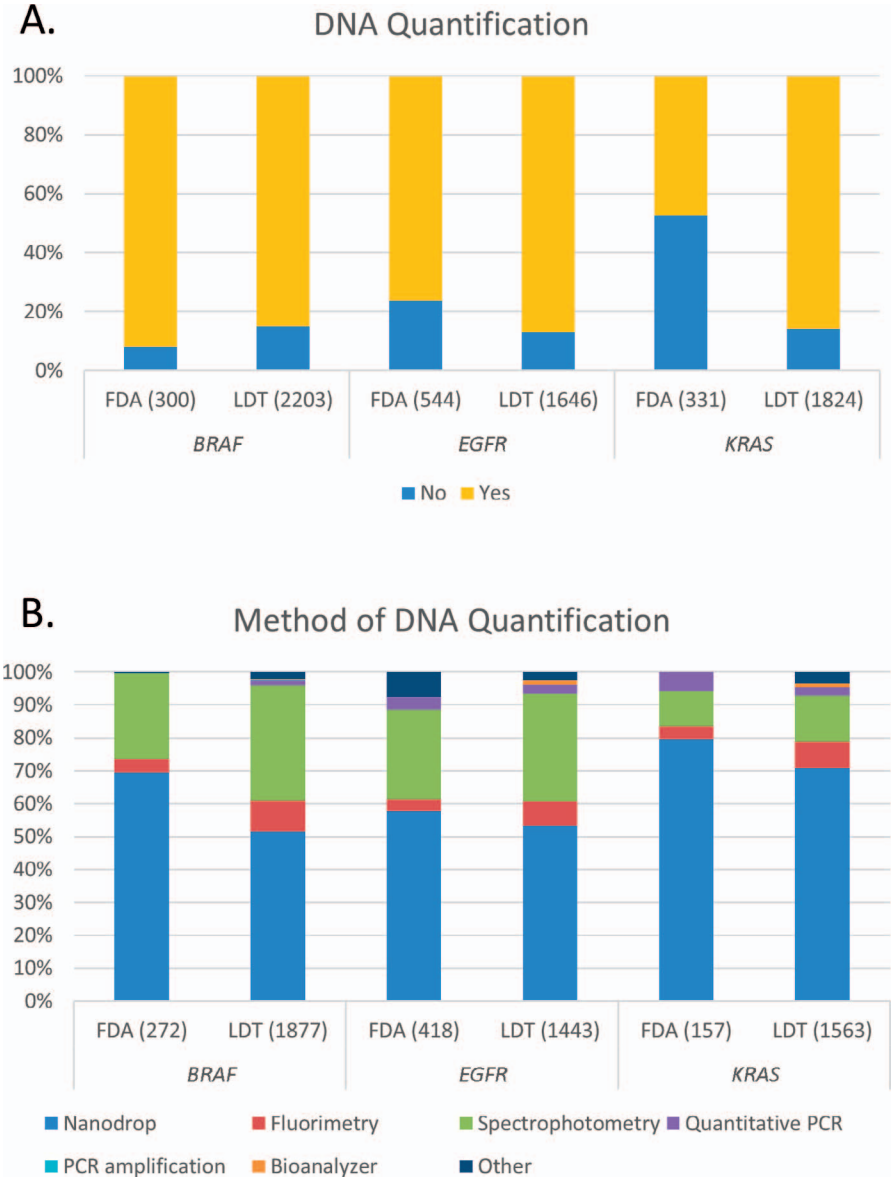

Figure 3. *A, Percentage of laboratories reporting that the DNA isolated from the tissue is quantified prior to testing. B, Percentage of laboratories reporting a specific method of DNA quantification. Numbers in parentheses report the total number of responses. Abbreviations: FDA, laboratory using a Food and Drug Administration–approved companion diagnostic; LDT, laboratory using a laboratory-developed test; PCR, polymerase chain reaction.*

study period have continued to demonstrate laboratories, including both FDA-CD and LDT laboratories, that either mistakenly identify a p.V600K as a p.V600E or fail to detect the p.V600K variant at lower tumor burdens (specifically, 59% of laboratories [29 of 49] that reported that their assay did not discriminate p.V600 variants failed to detect the p.V600K at 40% neoplastic cellularity).

For *EGFR* p.L861Q, the FDA-CD laboratories identified the acceptable/correct response in 100% of cases compared with 90.7% of LDT laboratories ($P = .04$). Investigation failed to identify an assay method or innate sequence issue that could account for this discrepancy. This variant is found in approximately 2% of *EGFR* mutated non–small cell lung cancers[33] and confers increased kinase activity to EGFR. Although these tumors lack the enhanced sensitivity to EGFR tyrosine kinase inhibitors associated with p.L858R and exon 19 deletions, the EGFR tyrosine kinase inhibitors are still superior to traditional platinum-based therapies for patients with *EGFR* p.L861Q; thus, missing this variant could result in failure to receive the optimal treatment.[34–37] It should be noted that in subsequent CAP PT specimens with this variant, LDTs have performed universally well without this discrepancy (data not shown).

When overall acceptability was examined for each kit manufacturer, both FDA-CD and LDT manufacturers showed excellent results overall, with typically more than 95% acceptable rates across all manufacturers with only 3 exceptions. The notable exceptions included 1 FDA-CD manufacturer of a *BRAF* assay (p.V600K issue, described above, n = 282) and 1 LDT manufacturer with lower rates of acceptability for its *BRAF* and *EGFR* assays from multiple laboratories (n = 94 and 91, respectively) using these kits (Table 2).

This study also compared FDA-CDs and LDTs for preanalytic and postanalytic factors, including specimen preparations used for testing, pathologist review, DNA quantification, tissue dissection, and interpretive comments. As previously published, more than 60% of all laboratories in this data set using an FDA-CD kit reported modifications from the FDA-approved protocol, effectively converting the test to an LDT.[26] Overall, neither FDA-CDs nor LDTs were more consistently associated with best laboratory practices,

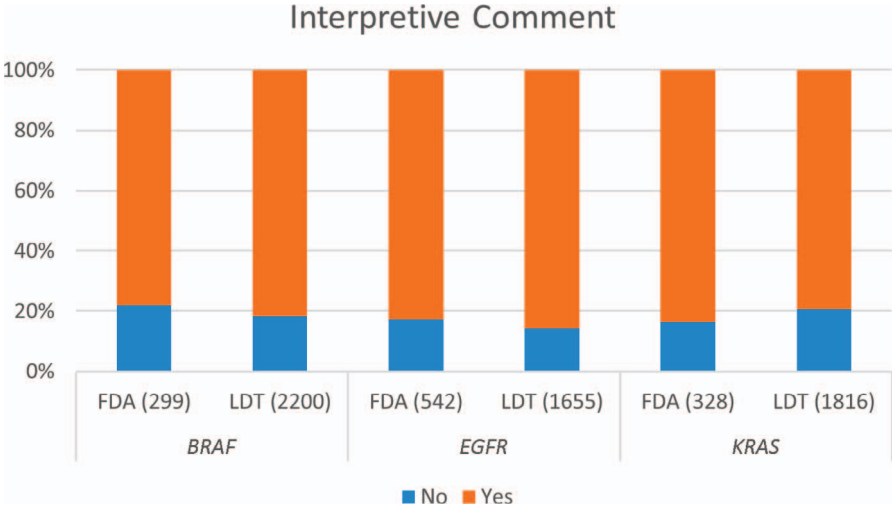

Figure 4. *Percentage of laboratories providing an interpretive comment to guide understanding of the clinical significance of the results. Numbers in parentheses report the total number of responses. Abbreviations: FDA, laboratory using a Food and Drug Administration–approved companion diagnostic; LDT, laboratory using a laboratory-developed test.*

as delineated by the CAP accreditation standards with regards to pathology review (CAP checklist MOL.32395), DNA quantification (CAP checklist MOL.32430), and the use of an interpretive comment (CAP checklist MOL.49570).[38,39] The only notable differences were the lower percentages of FDA-CD laboratories that performed preanalytic DNA quantification for *EGFR* and *KRAS*. However, the Qiagen and bioMérieux FDA-CDs for these analytes do not require a DNA quantification step because they use a control amplification, still keeping them compliant with CAP requirement MOL.35360.[38,39] Overall, the absence of consistent differences between FDA-CDs and LDTs, when considered in aggregate for these preanalytic and postanalytic factors, suggests that neither FDA-CD nor LDT tests are superior overall with respect to preanalytic and postanalytic laboratory practices that contribute to test quality.

One interesting modification to the FDA-approved protocols was the acceptance of specimens with lower tumor content than required for the approved assay. Although there is a broad range in the minimum acceptable fraction of neoplastic cells indicated by laboratories, several laboratories using FDA-CD tests report accepting a minimum tumor-cell burden that is less than the FDA-approved lower limit of detection for that specific assay (see Supplemental Table 1 and Figure 1). This adaptation to use a lower percentage neoplastic cellularity should be validated as an acceptable limit of detection in the individual laboratory as an LDT to ensure that there is no risk of reporting a false-negative result.[40] Although values of less than 10% neoplastic cellularity (5% allele limit of detection) may be allowable for only certain tested positions of *EGFR* and *KRAS* using Qiagen FDA-CDs, overall, no laboratory using an FDA-CD should be indicating the acceptance of tumor burdens below 10% and still have the test considered an FDA-approved assay. The data demonstrate that there was no trend for either FDA-CD or LDT laboratories to report superior analytic sensitivity.

There are several limitations of this study that bear additional consideration. First, each PT response was treated as an individual data point rather than grouping the responses by laboratory. As a result, each laboratory may be represented by as many as 14 data points. This could potentially result in one or more laboratories having a disproportionate impact on some variables, particularly those variables with smaller numbers of data points. Second, all participants using an assay produced by a vendor with an FDA-CD after the FDA approval date were categorized as FDA-CDs although some laboratories may have been using an alternate unapproved kit or have validated the FDA-CD as an LDT. Third, the assignment of NGS methods was based solely upon the report of use of commercial reagents from vendors known to provide solely NGS reagents. Therefore, the number of NGS laboratories is likely underestimated, as laboratories may purchase synthesized oligonucleotide primers and/or baits from other companies and would thus fall under the category of "no kit specified." A more definitive study on the performance of NGS platforms compared with non-NGS methods has been conducted, examining data from a time frame following the institution of survey questions specifically designed to identify laboratories using NGS on single-analyte as well as multianalyte surveys prior to the development of NGS-directed PT surveys.[41] Fourth, some laboratories failed to provide responses to all of the PT questions. This resulted in inconsistent numbers of data points for many of the variables, although the discrepancies were small and thus likely had minimal impact on the final results. Fifth, some results could be confounded by a failure to understand one or more of the survey questions. This was suggested by some participants submitting responses such as requiring "0% neoplastic cellularity" for their assay. Lastly, the PT did not include questions about all aspects of the performance of each assay. Therefore, this survey cannot determine if there are additional ways that laboratories are using the FDA-CDs off label.

In conclusion, this study demonstrates the overall high degree of accuracy and comparable performance across all laboratories, regardless of methodology. However, it should be noted that the specific assay design or platforms may affect the identification of certain clinically relevant variants, so both clinicians and laboratories should be aware of the limitations of the assay being used to test their patients. In addition, laboratories report a wide range of minimum neoplastic cellularity requirements for sample acceptability for testing. Thus, this study highlights the need for clinicians and laboratories to be aware of both the clinical and analytic limitations of their assay.

References

1. College of American Pathologists. AACC warns of "extraneous" LDT oversight. *CAP Today*. February 2015. http://captodayonline.com/aacc-warns-extraneous-ldt-oversight/. Accessed August 13, 2018.
2. Kaul KL, Sabatini LM, Tsongalis GJ, et al. The case for laboratory developed procedures: quality and positive impact on patient care. *Acad Pathol*. 2017;4: 2374289517708309.
3. Burton TM. Is lab testing the "Wild West" of medicine? *Wall Street Journal*. December 10, 2015. https://www.wsj.com/articles/is-lab-testing-the-wild-west-of-medicine-1449800707. Accessed August 13, 2018.
4. Evans JP, Watson MS. Genetic testing and FDA regulation: overregulation threatens the emergence of genomic medicine. *JAMA*. 2015;313(7):669–670.
5. Sharfstein J. FDA regulation of laboratory-developed diagnostic tests: protect the public, advance the science. *JAMA*. 2015;313(7):667–668.
6. O'Leary TJ. Regulating laboratory-developed tests. *J Mol Diagn*. 2014;16(6): 595–598.
7. Allen TC. Food and Drug Administration approval of laboratory tests. *Arch Pathol Lab Med*. 2013;137(1):13–18.
8. Bayefsky M, Berkman BE. FDA's proposed guidance for laboratory developed tests: how should regulators balance the risks and promise of innovation? *FDLIs Food Drug Policy Forum*. 2015;5(2). http://www.fdli.org/resources/resources-order-box-detail-view/regulating-laboratory-developed-tests-(ldts). Published February 25, 2015. Accessed August 13, 2018.
9. Ratner M. FDA pushes for control over laboratory-developed tests. *Nat Biotechnol*. 2014;32(9):855.
10. Ray T. Alternative proposals for lab test regulation: are there opportunities for consensus? https://www.genomeweb.com/sites/default/files/downloads/news/ldt_alternative_proposals.pdf. Accessed August 13, 2018.
11. O'Reilly KB. FDA's LDT proposal means "whole new ballgame" for labs. *CAP Today*. October 2014. http://captodayonline.com/fdas-ldt-proposal-means-whole-new-ballgame-labs/. Accessed August 13, 2018.
12. Sobel ME. ASIP response to FDA draft LDT guidance. https://www.asip.org/SciencePolicy/documents/ASIP.RESPONSE.TO.FDA.LDT.GUIDANCE.January2015.pdf. Published January 28, 2015. Accessed August 13, 2018.
13. Hwang TJ, Lehmann LS, Kesselheim AS. Precision medicine and the FDA's draft guidance on laboratory-developed tests. *Nat Biotechnol*. 2015;33(5):449–451.
14. Herbek GN. A need for clarity on regulation of LDTs. *CAP Today*. February 2015;29(2):11. https://captoday.epubxp.com/i/462136-feb-2015/10. Accessed August 13, 2018.
15. US Food and Drug Administration. The public health evidence for FDA oversight of laboratory developed tests: 20 case studies. http://wayback.archive-it.org/7993/20171115144712/https://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Reports/UCM472777.pdf. Accessed August 13, 2018.
16. Ferreira-Gonzalez A, Emmadi R, Day SP, et al. Revisiting oversight and regulation of molecular-based laboratory-developed tests: a position statement of the Association for Molecular Pathology. *J Mol Diagn*. 2014;16(1):3–6.
17. Evans BJ, Burke W, Jarvik GP. The FDA and genomic tests—getting regulation right. *N Engl J Med*. 2015;372(23):2258–2264.
18. Joly Y, Koutrikas G, Tassé AM, et al. Regulatory approval for new pharmacogenomic tests: a comparative overview. *Food Drug Law J*. 2011;66(1): 1–24.
19. US Food and Drug Administration (FDA). FDA notification and medical device reporting for laboratory developed tests (LDTs): draft guidance. www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM416684.pdf. Accessed August 13, 2018.
20. US Food and Drug Administration. CDRH'S approach to tumor profiling next generation sequencing tests. https://www.fda.gov/downloads/medicaldevices/productsandmedicalprocedures/invitrodiagnostics/ucm584603.pdf. Accessed August 13, 2018.
21. US Food and Drug Administration (FDA). List of cleared or approved companion diagnostic devices (in vitro and imaging tools). https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm301431.htm. Accessed August 13, 2018.
22. US Food and Drug Administration (FDA). FDA grants marketing approval to FoundationOne CDx in vitro diagnostic. https://www.fda.gov/drugs/informationondrugs/approveddrugs/ucm587387.htm. Accessed August 13, 2018.
23. US Food and Drug Administration (FDA). FDA unveils a streamlined path for the authorization of tumor profiling tests alongside its latest product action. https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm585347.htm. Accessed August 13, 2018.
24. Centers for Medicare & Medicaid Services. CMS finalizes coverage of next generation sequencing tests, ensuring enhanced access for cancer patients. https://www.cms.gov/Newsroom/MediaReleaseDatabase/Press-releases/2018-Press-releases-items/2018-03-16.html. Accessed August 13, 2018.
25. Centers for Medicare & Medicaid Services. Proposed decision memo for next generation sequencing (NGS) for Medicare beneficiaries with advanced cancer (CAG-00450N). https://www.cms.gov/medicare-coverage-database/details/nca-proposed-decision-memo.aspx?NCAId=290&bc=AAAAAAAAAAQAAA%3D%3D. Accessed August 13, 2018.
26. Kim A, Bartley A, Bridge J, et al. Comparison of laboratory-developed tests and FDA-approved assays for BRAF, EGFR, and KRAS Testing. *JAMA Oncol*. 2018; 4(6):838–841.
27. Treece AL, Gulley ML, Vasalos P, et al. Reporting results of molecular tests: a retrospective examination of BRAF mutation reporting. *Arch Pathol Lab Med*. 2017;141(5):658–665.
28. Tiacci E, Park JH, De Carolis L, et al. Targeting mutant BRAF in relapsed or refractory hairy-cell leukemia. *N Engl J Med*. 2015;373(18):1733–1747.
29. Hyman DM, Puzanov I, Subbiah V, et al. Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. *N Engl J Med*. 2015;373(8): 726–736.
30. Rosove MH, Peddi PF, Glaspy JA. BRAF V600E inhibition in anaplastic thyroid cancer. *N Engl J Med*. 2013;368(7):684–685.
31. Rush S, Foreman N, Liu A. Brainstem ganglioglioma successfully treated with vemurafenib. *J Clin Oncol*. 2016;31(10):159–160.
32. Peters S, Michielin O, Zimmermann S. Dramatic response induced by vemurafenib in a BRAF V600E-mutated lung adenocarcinoma. *J Clin Oncol*. 2013;31(20).
33. Mitsudomi T, Yatabe Y. Epidermal growth factor receptor in relation to tumor development: EGFR gene and cancer. *FEBS J*. 2010;277(2):301–308.
34. Chiu C-H, Yang C-T, Shih J-Y, et al. Epidermal growth factor receptor tyrosine kinase inhibitor treatment response in advanced lung adenocarcinomas with G719X/L861Q/S768I mutations. *J Thorac Oncol*. 2015;10(5):793–799.
35. Watanabe S, Minegishi Y, Yoshizawa H, et al. Effectiveness of gefitinib against non-small-cell lung cancer with the uncommon EGFR mutations G719X and L861Q. *J Thorac Oncol*. 2014;9(2):189–194.
36. Otsuka T, Mori M, Yano Y, et al. Effectiveness of tyrosine kinase inhibitors in Japanese patients with non-small cell lung cancer harboring minor epidermal growth factor receptor mutations: results from a multicenter retrospective study (HANSHIN Oncology Group 0212). *Anticancer Res*. 2015;35(7):3885–3891.
37. Wu J-Y, Yu C-J, Chang Y-C, Yang C-H, Shih J-Y, Yang P-C. Effectiveness of tyrosine kinase inhibitors on "uncommon" epidermal growth factor receptor mutations of unknown clinical significance in non-small cell lung cancer. *Clin Cancer Res*. 2011;17(11):3812–3821.
38. Jennings L, Van Deerlin VM, Gulley ML. Recommended principles and practices for validating clinical molecular pathology tests. *Arch Pathol Lab Med*. 2009;133(5):743–755.
39. Commission on Laboratory Accreditation, ed. *Molecular Pathology Checklist*. Northfield, IL: College of American Pathologists; 2016.
40. Viray Dr. 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(11):1545–1549.
41. 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 [published online March 13, 2019]. *Arch Pathol Lab Med*. doi: 10.5858/arpa.2018-0394-CP