OPEN

Two Instrument Comparison of Reagents From a US FDA-Approved Assay for the Assessment of Ki-67 in High-Risk Early Breast Cancer

Miglena Komforti, DO,* Erinn Downs-Kelly, DO,* Francisco Sapunar, MD,† Sameera R. Wijayawardana, PhD,‡ Aaron M. Gruver, MD, PhD,§ and Sunil S. Badve, MD||¶

Abstract: The objective of this study was to measure concordance of results obtained from the US Food and Drug Administration-approved Ki-67 immunohistochemistry MIB-1 pharmDx assay performed on the Dako Omnis automated staining instrument (Omnis) versus results produced from the assay reagents applied using an optimized protocol on the more widely available Autostainer Link 48 (ASL48) platform. Tissue sections obtained from 40 formalin-fixed paraffin-embedded breast carcinoma samples, with available Oncotype DX Breast Recurrence Score (RS) results, were stained. Three certified pathologists scored slides at 3 timepoints, totaling 360 observations for each instrument (N = 720 total) using the approved scoring approach. Using the $\geq 20\%$ cutoff, agreement was calculated with corresponding 2-sided 95% percentile bootstrap confidence intervals (CIs). Pairwise comparisons (N = 360) from the interinstrument evaluation, performed with

Received for publication January 31, 2022; accepted June 13, 2022.

- This work was supported by Eli Lilly and Company.
- M.K. reports receiving lecture fees from Agilent Technologies and has no additional potential conflicts of interest to disclose. E.D.-K. served on an oncology advisory board for Eli Lilly and Company. F.S. is an employee and company stockholder at Eli Lilly and Company. S.R.W. is an employee and company stockholder at Eli Lilly and Company. A.M.G. is an employee and company stockholder at Eli Lilly and Company; he also reports his spouse is an employee of Eli Lilly and Company. S.S.B. has received grant or research support from Agilent Technologies and is a paid speaker for Agilent Technologies.
- Reprints: Aaron M. Gruver, MD, PhD, Clinical Diagnostics Laboratory, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, 46285, IN (e-mail: gruver_aaron_m@lilly.com).
- Copyright © 2022 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

all observers, resulted in 325 (90.3%) concordant outcomes (244 negative and 81 positive) and 35 (9.7%) discordant outcomes. The overall agreement was 90.3% (95% confidence interval, 85.6% to 94.4%). No significant systematic differences were observed between instruments. Specimens scored from the Omnis were on average <1% higher than ASL48, with high correlation and little bias between the continuous Ki-67 scores (concordance correlation coefficient = 0.916). Most specimens with a Ki-67 score $\geq 20\%$ had a RS > 25. This study demonstrated that good concordance can be achieved with the reagents run on the ASL48 instrument when using an optimized protocol and standardized scoring.

Key Words: abemaciclib, early breast cancer, Ki-67, MIB-1, Omnis

(Appl Immunohistochem Mol Morphol 2022;30:577-583)

pproximately 90% of patients with breast cancer are A diagnosed at an early disease stage, and the hormone receptor-positive (HR+), human epidermal growth factor receptor 2-negative (HER2-) subtype accounts for nearly 73% of all breast cancers.¹ These patients are treated with curative intent and are candidates for surgery with or without radiotherapy depending upon the extent of regional disease.^{2,3} After surgery, adjuvant treatment is based on the estimated risk of disease recurrence and predicted sensitivity to available systemic therapies.^{2,3} Validated clinical and pathologic features that indicate a higher risk of distant disease recurrence include large primary tumor size, axillary lymph node involvement, and high histologic grade.⁴⁻⁶ As many as 41% of women with HR+ breast cancer initially diagnosed with early-stage disease experience distant recurrence after standard of care, including adjuvant endocrine therapy (ET), depending upon lymph node status and tumor grade.⁷

The Ki-67 antigen is a nuclear protein expressed during all active phases of the mammalian cell cycle (G1, S, G2, and M-phases) and downregulated in resting cells (G0-phase).⁸ Ki-67 expression has been widely studied as a marker of cell proliferation and as an independent prognostic factor in early breast cancer (EBC).⁹ In HR+ breast

Staining Step	Reagent	Omnis Procedure Incubation, min:s (Number of Cycles)	ASL48 Procedure Incubation, min: (Number of Cycles)	
Wash	Wash buffer	2:40 (2)	0:00 (1)	
Primary antibody	Ki-67 MIB-1 pharmDx	20:00 (1)		
Endogenous enzyme block	Envision FLEX peroxidase- blocking reagent	_``	5:00 (1)	
Wash	Wash buffer	2:00 (10)	0:00 (1)	
Endogenous enzyme block	Envision FLEX peroxidase- blocking reagent	3:00 (1)	_	
Primary antibody	Ki-67 MIB-1 pharmDx	_	20:00 (1)	
Wash	Wash buffer	2:00 (10)	0:00 (1)	
Labeled polymer	Envision FLEX/HRP	20:00 (1)	20:00 (1)	
Wash	Wash buffer	2:00 (10)	0:00 (1)	
Wash	DI water	0:31 (1)		
Wash	Wash buffer	2:00 (10)	5:00 (1)	
Wash	Wash buffer	2:00 (10)		
Substrate chromogen	Envision FLEX substrate working solution	5:00 (1)	5:00 (2)	
Wash	Wash buffer	2:00 (10)	0:00 (1)	
Wash	DI water	0:31 (1)		
Wash	Wash buffer	2:00 (10)	_	
Counterstain	FLEX hematoxylin	3:00 (1)	5:00 (1)	
Wash	DI water	2:00 (10)	0:00 (1)	
Wash	Wash buffer	2:00 (10)	5:00 (1)	
Wash	DI water	_	0:00 (1)	

TABLE 1. Immunohistochemistr	Procedure on the Omnis Versus ASL48 Automated IHC Instru	ment

cancer, patients with high levels of Ki-67 after surgery have been shown to have higher recurrence rates while receiving adjuvant ET.^{10}

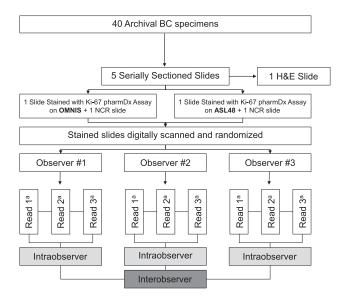


FIGURE 1. Study design for comparison of the Ki-67 immunohistochemistry (IHC) MIB-1 pharmDx reagents run on the Dako Omnis and Dako Autostainer Link48 instruments. ^aA minimum 14-day washout period was applied between reads. ASL48 indicates Dako Autostainer Link 48; BC, breast carcinoma; H&E, hematoxylin and eosin; NCR, negative control reagent; OMNIS, Dako Omnis automated staining instrument.

Abemaciclib was the first cyclin-dependent kinase 4 and 6 (CDK 4/6) inhibitor to receive US Food and Drug Administration (FDA) approval, in combination with ET (tamoxifen or an aromatase inhibitor), for the adjuvant treatment of adult patients with HR+, HER2-, nodepositive EBC at high risk of recurrence and a Ki-67 score of $\geq 20\%$, as determined by an FDA-approved test.¹¹ This approval was based on the data from the monarchE trial, which is a randomized, open-label, global phase 3 study comparing treatment with abemaciclib combined with standard adjuvant ET versus ET alone in patients with high risk, node-positive, early stage, HR+, HER2- breast cancer.¹² In this population, adjuvant abemaciclib combined with ET resulted in clinically meaningful improvements in invasive disease-free and distant recurrence-free survival outcomes.13,14

Until recently, the lack of standardized procedures or accepted cutoff definitions have limited the application of Ki-67 assessment during breast cancer workup in some geographies.^{9,15} Therefore, an investigational assay performed with a standardized protocol and interpretation guidelines was developed to help determine the eligibility of patients with high-risk EBC as part of the monarchE trial.¹⁶ With the abemaciclib approval, described earlier, came the simultaneous US FDA approval of the investigational Ki-67 immunohistochemistry (IHC) assay used in monarchE to serve as a companion diagnostic to aid in identifying patients with EBC at high risk of disease recurrence for whom adjuvant treatment with abemaciclib in combination with ET is being considered.¹⁷ Although Ki-67 IHC is a commonly performed assay using a variety of commercially available in vitro diagnostic

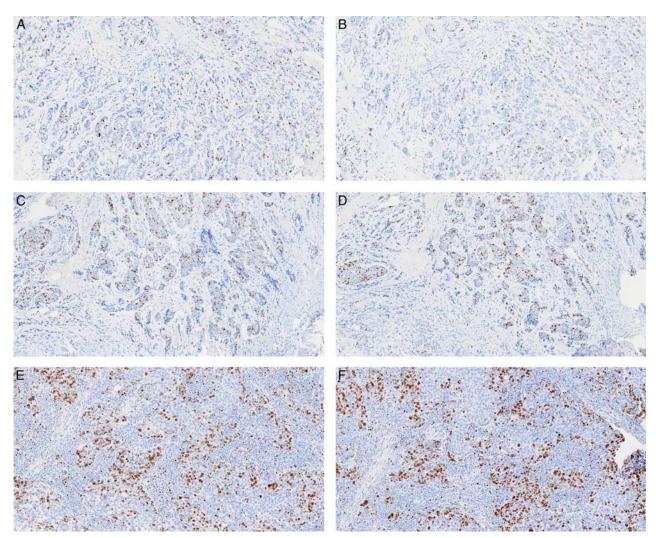


FIGURE 2. Representative images captured from tumors with low (A and B), medium (C and D), and high (E and F) Ki-67 immunoreactivity. Tissues shown in A, C, and E were stained with the ASL48, whereas those in B, D, and F were stained with the Omnis platform. Scanned slide images were obtained at ×10 zoom level.

reagents, many diagnostic laboratories do not have access to specific automated IHC instruments, which are a part of Class III devices approved by the US FDA (ie, the Dako Omnis). However, similar instruments such as the ASL48 are more available in US laboratories because of the proliferation of PD-L1 IHC testing.^{18,19} Knowing this limitation, the following study was conducted to evaluate the analytic performance of Ki-67 IHC MIB-1 pharmDx reagents applied using the Dako Omnis automated staining instrument (Omnis) compared with an optimized protocol on the more prevalent Dako Autostainer Link 48 (ASL48) instrument. Here, we test the hypothesis that these reagents can produce similar results on either staining platform.

MATERIALS AND METHODS

Deidentified human tissue was obtained according to the protocols and procedures of the Institutional Review Board at Indiana University School of Medicine

(#2001022404). The specimens selected were HR+, HER2-, as determined by the IHC staining for estrogen receptor, progesterone receptor, and HER2 expression before study inclusion. All specimens were also previously tested via the Oncotype DX assay (Exact Sciences, Redwood City, CA), and historical RS information was used for comparisons.²⁰ Sections from these blocks were cut at 4 µm thickness, mounted on positively charged glass slides, oven dried at $58 \pm 2^{\circ}$ C for ~1 hour, and stored in the dark at 2 to 8°C until staining (within 1 mo of microtomy). The Ki-67 IHC MIB-1 pharmDx (Dako Omnis) (GE020) kit, necessary additional materials referenced in the package insert,²¹ and automated IHC instruments were obtained from Agilent Technologies Inc. (Santa Clara, CA). An optimal protocol for the ASL48 instrument was selected before study initiation, and staining was performed as described in Table 1. The FDA-approved assay was performed according to the instructions for use.

The final test set comprised 30 negative (Ki-67 score <20%) and 10 positive (Ki-67 score \geq 20%) samples, as determined by the consensus IHC status of the Omnis scores across all observers and reads. Of the 40 samples, 19 (47.5%) were near the diagnostic cutoff established for the monarchE phase 3 clinical study (defined as a Ki-67 score in the 10% to 30% range¹⁶). Specimens were examined by 3 certified pathologists, at 3 different timepoints (40 slides \times 9 observations/slide = 360 readings for each instrument or 720 data points total) (Fig. 1). Evaluations were performed utilizing wholeslide images captured using a PANNORAMIC 250 scanner (3DHISTECH, Budapest, Hungary) at ×40 magnification and displayed utilizing the PathoTrainer whole-slide image viewer (CellCarta, Belgium). Immunoreactivity was assessed using the Ki-67 pharmDx Score, as described.²² The run-to-run variation for the Ki-67 IHC MIB-1 pharmDx assay (eg, interday, interinstrument, interlot, and interrack) has been previously described.¹⁶ For the accuracy of results and the reduction of recall bias, observers were blinded to the identifiers for specimens used in all studies. A minimum 14-day washout period was applied between reads (Fig. 1).

For the concordance analysis, 2-sided 95% confidence intervals (CIs) were calculated using intraobserver and intraread pairwise comparisons. Comparisons were made on the IHC status (positive/ negative) between each test condition (ASL48) and the consensus reference condition (Omnis) for each specimen within each read by each observer. The total number of comparisons per specimen is therefore equal to the total number of test scores, excluding the reference score (3 observers \times 3 reads per observer = 9 comparisons per sample). Calculations of negative percent agreement, positive percent agreement, and overall percent agreement were performed. As multiple ASL48 sections per specimen were compared with the reference. a nonparametric percentile bootstrap was used to calculate 2-sided 95% CIs. Continuous Ki-67 percentage scores from both instruments were evaluated using scatter and difference plots. The concordance correlation coefficient (CCC) with corresponding 95% CI was also calculated to evaluate correlation and bias in the data. The CCC values close to 1 represent paired data in which high correlation and no bias between the 2 conditions is observed.²³ Mean Ki-67 percentage scores (instrument/observer/read) were used to evaluate the relationship with RS.

RESULTS AND DISCUSSION

Qualitative assessment of the stained slides demonstrated the immunoreactivity produced with reagents run on the ASL48 was similar but slightly more intense when compared with the Omnis instrument on average (Fig. 2). Concordance analysis resulted in a point estimate for positive percent agreement of 79.4% (95% CI, 69.3% to 88.5%), negative percent agreement

Interinstrument, $\geq 20\%$ Cutoff (All Observers)							
	Specimen distri	bution summary	r				
Definition	Positive Range (%)	No. Blocks	Blocks (%)				
Negative	< 20	30/40	75.0				
Positive	≥ 20	10/40	25.0				
Near cutoff	10-30	19/40	47.5				
	Comparisons to stand	dard (standard:C	Omnis)				
Negat	Negative Standard Positive Standard						
True	False	True	False				
Negatives	Positives	Positives	Negatives	Total			
244	14	81	21	360			
	Agreemen	t Summary					
		95% Confidence Interval (Bootstrap)					
Performance	Point	Lower Bound:	Upper Bound:				
Criteria	Estimate	2.5%	97.5%				
NPA	94.6	91.1	97.5				
PPA	79.4	69.3	88.5				
OA	90.3	85.6	94.4				

NPA indicates negative percent agreement; OA, overall percent agreement; PPA, positive percent agreement.

of 94.6% (95% CI, 91.1% to 97.5%), and overall percent agreement of 90.3% (95% CI, 85.6% to 94.4%), indicating good overall concordance in samples when stained using either the Omnis or ASL48 instruments based on consensus reference scores (Table 2).

Continuous score analysis indicates high correlation and little to no bias between the 2 staining platforms. The CCC point estimate is 0.916 (Fig. 3A), and the average score obtained from the Omnis instrument was <1% higher than the corresponding ASL48 score (Fig. 3B). There was also no apparent bias in score variability between or within observers associated with the overall mean Ki-67 percentage score when results were evaluated by specimen (Fig. 3C). When comparing the mean Ki-67 percentage score against the RS (Fig. 4), the Ki-67 percentage score tended to increase with greater RS when RS ≥ 25 . In this limited data set, this tendency is suggestive of possible higher concordance of Ki-67 (using the assay cutoff of $\geq 20\%$) and the corresponding RS when RS ≥ 25 .

The recent FDA-approved Ki-67 IHC assay is a qualitative test using a MIB-1 monoclonal mouse anti-Ki-67 antibody with a polymer-based detection system on the Dako Omnis platform.²⁴ Although the results from the UK National External Quality study highlight the importance of matching primary antibody clones and other factors to ensure quality of Ki-67 testing.²⁵ access to specific automated IHC instruments is a real-world limitation for many laboratories. This report demonstrates that good concordance is achievable when samples are stained with these reagents using the ASL48 instrument. Furthermore, these data suggest that there is no inherent bias in pathologist scores from the same sample stained using the Ki-67 IHC MIB-1

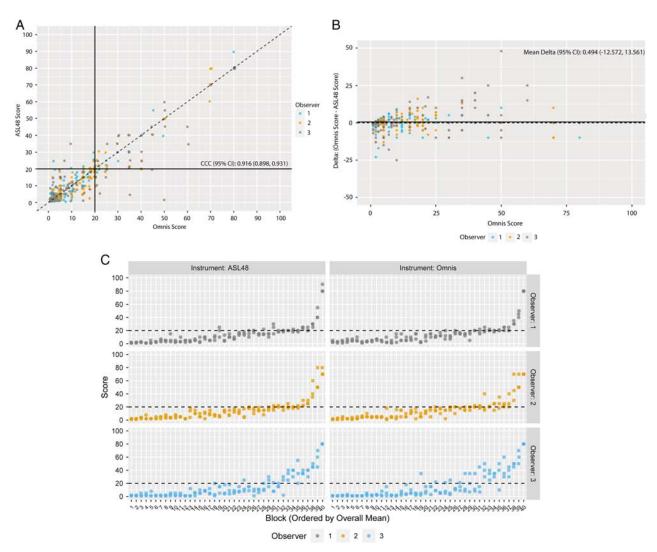


FIGURE 3. Continuous score analysis to evaluate correlation and bias between 2 staining platforms. A, The scatter plot visualizes the paired Omnis and ASL48 Ki-67 scores within each read and observer. Points that fall along the dotted diagonal line represent perfectly correlated scores between the instruments. The Ki-67 \geq 20% cutoff is visualized by the horizontal and vertical solid black lines. Points falling within the bottom left and top right quadrant of the cutoff lines represent true negative and true positive comparisons, respectively, when using the outcome of the Omnis score as a reference. Points that fall within the top left and bottom right quadrants represent false positive and false negative comparisons, respectively, when using the Omnis score as reference plot compares the delta in paired scores (*y*-axis = Omnis – ASL48) against the corresponding Omnis score (*x*-axis) for a given pair. The dashed horizontal line represents a delta of 0, meaning no difference in Ki-67 score between Omnis and ASL48. The solid horizontal line represents the mean delta across all observations. C, Intraobserver scores per specimen, split by instrument and observer.

pharmDx reagents on the ASL48 instrument when applying the $\geq 20\%$ cutoff with the standardized scoring instructions.

The results of this technical study should be considered within the context of some limitations, including a relatively small sample size adequate for the initial evaluation of nonpredictive factor assays,²⁶ and a distribution enriched toward cases with lower levels of Ki-67 immunoreactivity around the diagnostic cutoff. These limitations were counterbalanced by the use of multiple pathologists who provided independent reads for each tissue, assessed by both automated IHC instruments, during multiple reading sessions using controls to reduce recall bias. Pathologists should assess individual IHC assay performance in their local environment when evaluating the purpose and potential role of a laboratorydeveloped test.²⁷ Future studies to evaluate how the FDA-approved test performs compared with other commonly used in vitro diagnostic assays, and its correlation with the RS, will further inform local testing procedures for the assessment of Ki-67 IHC in highrisk EBC.

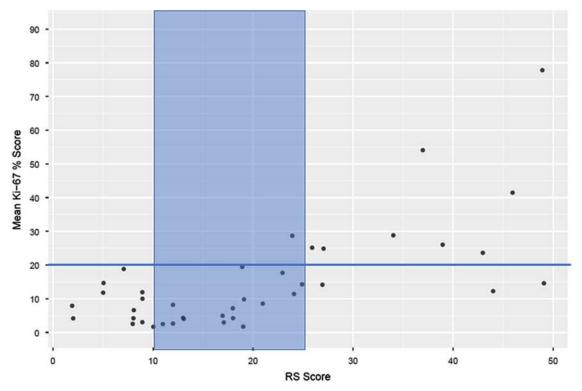


FIGURE 4. Mean Ki-67 percentage scores plotted with RS obtained from the same tissue blocks. The box insert highlights scores in the 11 to 25 RS range. RS, recurrence score.

ACKNOWLEDGMENTS

The authors thank Neeraj Kumar Singh for the assistance in staining on both the Omnis and the Autostainer Link 48 and Peter Krein, PhD, for the critical review of the manuscript. Eli Lilly and Company contracted with Syneos Health for the writing and editorial assistance from Andrea Metti, PhD, MPH and Dana Schamberger, MA. The authors also thank Agilent Technologies for all reagents and instrumentation.

REFERENCES

- Howlader N, Altekruse SF, Li CI, et al. US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. *J Natl Cancer Inst.* 2014;106:dju055.
- NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines), Breast Cancer (version 1.2022). November 24, 2021.
- Cardoso F, Kyriakides S, Ohno S, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2019;30:1194–1220.
- 4. Gluz O, Nitz UA, Christgen M, et al. West German Study Group Phase III PlanB Trial: First prospective outcome data for the 21-Gene Recurrence Score Assay and Concordance of Prognostic Markers by Central and Local Pathology Assessment. *J Clin Oncol.* 2016;34:2341–2349.
- Mamounas EP, Tang G, Paik S, et al. 21-Gene Recurrence Score for prognosis and prediction of taxane benefit after adjuvant chemotherapy plus endocrine therapy: results from NSABP B-28/NRG Oncology. *Breast Cancer Res Treat*. 2018;168: 69–77.
- 6. Marme F, Lederer B, Blohmer J-U, et al. Utility of the CPS + EG staging system in hormone receptor-positive, human epidermal growth factor receptor 2-negative breast cancer treated with neoadjuvant chemotherapy. *Eur J Cancer*. 2016;53:65–74.

- Pan H, Gray R, Braybrooke J, et al. 20-year risks of breast-cancer recurrence after stopping endocrine therapy at 5 Years. N Engl J Med. 2017;377:1836–1846.
- Gerdes J, Lemke H, Baisch H, et al. Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. *J Immunol.* 1984;133:1710–1715.
- Dowsett M, Nielsen TO, A'Hern R, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in breast cancer working group. J Natl Cancer Inst. 2011;103:1656–1664.
- Viale G, Giobbie-Hurder A, Regan MM, et al. Prognostic and predictive value of centrally reviewed Ki-67 labeling index in postmenopausal women with endocrine-responsive breast cancer: results from Breast International Group Trial 1-98 comparing adjuvant tamoxifen with letrozole. J Clin Oncol. 2008;26:5569–5575.
- FDA approves abemaciclib with endocrine therapy for early breast cancer. Available at: https://wwwfdagov/drugs/resources-informationapproved-drugs/fda-approves-abemaciclib-endocrine-therapy-earlybreast-cancer. Accessed December 15, 2021.
- Johnston SRD, Harbeck N, Hegg R, et al. Abemaciclib combined with endocrine therapy for the adjuvant treatment of HR+, HER2-, node-positive, high-risk, early breast cancer (monarchE). J Clin Oncol. 2020;38:3987–3998.
- Harbeck N, Rastogi P, Martin M, et al. Adjuvant abemaciclib combined with endocrine therapy for high-risk early breast cancer: updated efficacy and Ki-67 analysis from the monarchE study. *Ann Oncol.* 2021;32:1571–1581.
- 14. Harbeck N, Rastogi P, Shahir A, et al. Letter to the Editor for 'Adjuvant abemaciclib combined with endocrine therapy for highrisk early breast cancer: updated efficacy and Ki-67 analysis from the monarchE study'. Ann Oncol. 2022;33:227–228.
- Nielsen TO, Leung SCY, Rimm DL, et al. Assessment of Ki67 in breast cancer: updated recommendations from the International Ki67 in Breast Cancer Working Group. J Natl Cancer Inst. 2021; 113:808–819.

- 16. Polewski MD, Nielsen GB, Gu Y, et al. A standardized investigational Ki-67 immunohistochemistry assay used to assess high-risk early breast cancer patients in the monarchE phase 3 clinical study identifies a population with greater risk of disease recurrence when treated with endocrine therapy alone. *Appl Immunohistochem Mol Morphol.* 2022;30:237–245.
- 17. Royce M, Osgood C, Mulkey F, et al. FDA approval summary: abemaciclib with endocrine therapy for high-risk early breast cancer. *J Clin Oncol.* 2022;40:1155–1162.
- Roach C, Zhang N, Corigliano E, et al. Development of a companion diagnostic PD-L1 immunohistochemistry assay for pembrolizumab therapy in non-small-cell lung cancer. *Appl Immunohistochem Mol Morphol.* 2016;24:392–397.
- Phillips T, Simmons P, Inzunza HD, et al. Development of an automated PD-L1 immunohistochemistry (IHC) assay for non-small cell lung cancer. *Appl Immunohistochem Mol Morphol.* 2015;23: 541–549.
- Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004;351:2817–2826.
- Ki-67 IHC MIB-1 pharmDx (Dako Omnis) package insert. Available at: https://www.agilent.com/cs/library/packageinsert/ public/PD00903E_02.pdf. Accessed April 13, 2022.

- Ki-67 IHC MIB-1 pharmDx (Dako Omnis) Interpretation manual breast carcinoma. Agilent Technologies, Inc. Published November 18, 2021. Available at: https://www.agilent.com/en/product/pharmdx/ki-67-ihc-mib-1-pharmdx-dako-omnis/ki-67-ihc-mib-1-pharmdx-dakoomnis-1963684. Accessed December 10, 2021.
- Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45:255–268.
- Ki-67 IHC PharmDx Premarket Approval (PMA). US Food and Drug Administration. Available at: https://www.accessdata.fda.gov/scripts/ cdrh/cfdocs/cfpma/pma.cfm?id=P210026. Accessed December 10, 2021.
- Parry S, Dowsett M, Dodson A. UK NEQAS ICC & ISH Ki-67 data reveal differences in performance of primary antibody clones. *Appl Immunohistochem Mol Morphol.* 2021;29:86–94.
- 26. Fitzgibbons PL, Bradley LA, Fatheree LA, et al. Principles of analytic validation of immunohistochemical assays: Guideline from the College of American Pathologists Pathology and Laboratory Quality Center. Arch Pathol Lab Med. 2014;138: 1432–1443.
- Torlakovic EE, Cheung CC, D'Arrigo C, et al. Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine. Part 3: technical validation of immunohistochemistry (IHC) assays in clinical IHC laboratories. *Appl Immunohistochem Mol Morphol.* 2017;25:151–159.