Supplemental Digital Content

This supplement has been provided by the authors to give readers additional information about the development of the assay.

Supplement to: Polewski, 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

Contents

1 Supplemental Methods

1.1 Investigational Assay Design Input

Early input for the assay prototype was provided from a qualitative comparison of commercially acquired formalin-fixed, paraffin-embedded (FFPE) human breast carcinoma tissue assessed by the Ki-67 immunohistochemistry (IHC) assay used in the neoMONARCH clinical trial.[1] A limited sample set of 18 invasive breast carcinoma resection specimens, a tissue microarray comprised of 49 invasive breast carcinoma cores, and tonsil control tissue were assayed with the laboratory developed test (LDT) at University of Southern California, and were assayed with a prototype assay (Agilent). Alternating section levels were stained with both assays prior to performance comparisons. The study was performed for informational purposes only, and no formal acceptance criteria were applied.

1.2 Investigational Assay Sensitivity Study

The expression level of Ki-67 protein in procured FFPE breast carcinoma tissue sections using the Investigational Use Only (IUO) Ki-67 IHC assay was evaluated. One hundred and forty-eight specimens were stained using the assay. Specimens reflected a range of Ki-67 expression levels and included both resection (n=100) and core needle biopsy (CNB) (n=48) specimens. A subset of the specimens were sister blocks from the same case; therefore, prevalence analysis was conducted using the average score from these blocks. Prevalence analysis was performed on 113 unique specimens (resection, n=80; CNB, n=33).

1.3 Investigational Assay Specificity Studies

1.3.1 Western Blot

Cell lysates from 2 cancer cell lines, SKBR3 and IM-9, were used for a Western immunoblot to establish the specificity of the monoclonal mouse antibody (clone MIB-1) for Ki-67. The samples were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDSPAGE) along with a protein molecular weight marker ladder (Novex Hi-Mark, Fisher Scientific) and then transferred onto a polyvinylidene difluoride membrane. To show inhibition of Ki-67 MIB-1 antibody binding in the presence of a peptide derived from a region of the Ki-67 protein containing the binding epitope (Abcam 15581), the blot was cut into 3 identical strips, each containing the transferred SKBR3 and IM9 cell lysate proteins. One strip was incubated with MIB-1 clone alone as primary antibody, the other 2 with MIB-1 antibody premixed with different amounts of inhibitory peptide (5x and 15x excess by weight). Following primary antibody incubation, Ki-67 protein was detected with a goat anti-mouse fluorescent-tagged secondary antibody using SuperSignal West Femto substrate (Fisher Scientific). After imaging, the blots were stripped of Ki-67 primary antibody and re-stained with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) housekeeping protein antibody to show equal loading of cell lysate proteins. The detection method for the housekeeping protein was a goat anti-mouse fluorescent-tagged secondary antibody.

Immunoreactivity: 31 Normal Tissue Testing

Immunohistochemistry-based specificity testing was conducted on a sample set consisting of resection specimens derived from three different cases each of the 30 tissue types recommended by the United States Food and Drug Administration,[2] as well as 3 normal bladder specimens.

1.4 Investigational Assay Robustness Study

The assay was evaluated under various laboratory conditions (diluted Target Retrieval Solution pH range 5.9-6.5, tissue section thickness from 3-6 µm, and overnight and over-weekend staining procedures) to assess the robustness of the assay.

The effect of preanalytical variables was also studied. For evaluation of fixative type and fixation time, archived tonsil specimens (n=4 per condition tested in triplicate) that had been processed with various fixatives (i.e. 10% neutral-buffered formalin [NBF], acetic formalin alcohol, Bouin's solution, 10% unbuffered formalin, and 10% NBF followed by 4-5 days in 70% ethanol) and fixation times (24, 48, or 72 hours) were stained. Additionally, fixation and ischemia times were evaluated on fresh tonsil tissue cut into approximately equally sized pieces and placed under wet gauze at ambient temperature for various ischemia times before being placed into NBF fixative. Fixation times between 6 and 72 hours and ischemia times between 30 minutes and 72 hours were studied. Prospectively collected tonsil tissue processed under these defined conditions were obtained by the Cooperative Human Tissue Network.

1.5 Tumor Heterogeneity Study

Tumor heterogeneity was assessed in FFPE breast carcinoma specimens stained with IUO Ki-67 IHC assay. Intrablock tissue heterogeneity was assessed in 36 specimens between nonserial sections across a span of at least 200 µm. Intra-case heterogeneity was assessed between 25 unique breast carcinoma specimen cases in sister block pairs (consisting of 50 total unique block IDs). Sister blocks are defined as paraffin blocks prepared from the same specimen. The positive/negative diagnostic status of each slide was determined based on the cutoff (≥20%). Comparisons to consensus were made using the diagnostic status of each block and were used to

calculate negative percent agreement (NPA), positive percent agreement (PPA), and overall percent agreement (OA).

1.6 Investigational Assay Precision Study

Precision testing was conducted internally at Agilent Technologies to demonstrate the assay produces consistent results in normal day-to-day testing. Repeatability was measured within instrument, staining rack, and day. Assay precision was measured between Dako Omnis instruments, testing days, and different lots of primary antibody and accessory reagent lots. Since the Dako Omnis platform is fully automated, interoperator variability has been determined to be negligible. Breast carcinoma specimens with a range of expression regarding percentage of positive tumor cells were selected for internal analytical precision studies. Efforts were made to include approximately 20% to 25% of specimens considered to be in the near cutoff range (10%-30%). Repeatability and interinstrument, interday, and accessory reagent interlot precision studies were performed using replicates from 32 breast carcinoma specimens. Assay interlot testing utilized replicates from 40 specimens. Precision testing was conducted over 5 nonconsecutive days. The stained slides were assessed for their diagnostic concordance in positivity and negativity of Ki-67 expression using a cutoff of ≥20%.

Interobserver precision was evaluated by testing scoring reproducibility between 3 trained and certified pathologists, who performed 3 independent Ki-67 evaluations on a set of 60 specimens. Information regarding the hormone receptor (HR) status and human epidermal growth factor receptor 2 (HER2) status was not available for all

specimens. The consensus for interobserver analysis was determined as the majority call for the sample across all 9 observations.

2 Supplemental Results

2.1 Prototype Investigational Assay Comparison with the Laboratory Developed Test used in neoMONARCH

Prior to finalizing the investigational assay format, an informal comparison of the prototype with the LDT used in the phase 2 neoMONARCH trial was performed. A similar quality of immunoreactivity was observed between both assays (Supplemental Figure 1). In some cases, minor differences were observed in hematoxylin appearance and nuclear morphology. When the same scoring method was applied to both assays, a high degree of concordance was observed across the small sample set (Supplemental Table 1).

2.2 Investigational Assay Specificity

The specificity of the MIB-1 clone for the Ki-67 antigen has been described previously.[3] This study confirmed that MIB-1 detected protein bands corresponding to the expected sizes for Ki-67 antigen of 345 kDa and 395 kDa on a Western blot using lysates of the cell line, IM-9. As expected, cell line SKBR-3 did not express detectable Ki-67 protein as the level of Ki-67 RNA expressed is minimal (Supplemental Figure 2).[4] The binding of Ki-67 MIB-1 antibody to the Ki-67 protein could be completely abrogated by the addition of a peptide containing the region reported to contain the epitope of this antibody. When MIB-1 antibody was preincubated with 5 times excess by weight of the peptide, partial blocking of the binding was achieved, and when the peptide was added at 15 times excess by weight, no binding was detected.

Supplemental Table 2 summarizes the IHC-based specificity testing on 31 normal tissues demonstrating that Ki-67 IHC detects Ki-67 protein in the appropriate tissue elements and cellular compartments. Nuclear staining was observed in most specimens, and there were no unexpected results observed in cell types or tissue types tested. The observed staining was consistent with the reported literature for Ki-67 IHC expression in normal tissues.[5,6]

2.3 Investigational Assay Sensitivity

The prevalence of Ki-67 protein was evaluated in 113 unique FFPE specimens stained for Ki-67 with the clone MIB-1 monoclonal antibody, irrespective to HR or HER2 status. Specimens reflected a range of Ki-67 expression levels (Supplemental Figure 3 and Supplemental Figure 4) and included both resection (n=80) and CNB (n=33) specimens. Ki-67 expression was observed in tumor cells and a subset of benign elements including lymphocytes, stromal cells, and epithelium. Ki-67 IHC consistently detected Ki-67 protein across a relevant expression range (0% to 75%) in commercially procured specimens. Seventy-four percent of the specimens were negative based on the cutoff (<20% tumor cells positive) and 26% were positive (≥20% tumor cells positive). This prevalence is consistent with reported literature.[7,8] The distribution was similar in resection specimens and in CNB specimens (data not shown).

2.4 Investigational Assay Robustness

Robustness testing was conducted to evaluate the staining performance of the assay under various laboratory conditions. The assay demonstrated compatibility with Dako Omnis workflow options that allow the instrument to be preprogrammed and preloaded for a delayed start and used overnight or over a weekend, with slides being

removed from the loading tray when the staining is complete. In addition, the assay achieved highly consistent results when tested with a range of target-retrieval solution pH values between 5.9 and 6.5 and tissue section thicknesses between 3 µm and 6 µm. Analysis results are summarized in Supplemental Table 3. Evaluation of preanalytic variables demonstrated an ischemic time of 1 hour or less is tolerated and fixation with 10% NBF is required, which is consistent with the optimal testing recommendations for estrogen receptor, progesterone receptor, and HER2 biomarker testing in breast cancer.[9,10] Tonsil specimens fixed with acetic formalin alcohol, 10% unbuffered formal and Bouin's solution, or tonsil held in 70% ethanol for 4 to 5 days following fixation in 10% NBF, demonstrated altered immunostaining intensity in Ki-67 positive cells when compared to the reference condition, 10% NBF for 24 hours. Fixation times between 6 and 72 hours produced equivalent results, when 10% NBF was used and ischemia times were kept at or below 1 hour.

2.5 Tumor Heterogeneity Assessed with the Investigational Assay

Intrablock and intracase heterogeneity studies showed high agreement rates (Supplemental Table 3). Results demonstrate that Ki-67 diagnostic classification was consistent within tissue blocks and between sister blocks from the same case.

2.6 Precision Studies

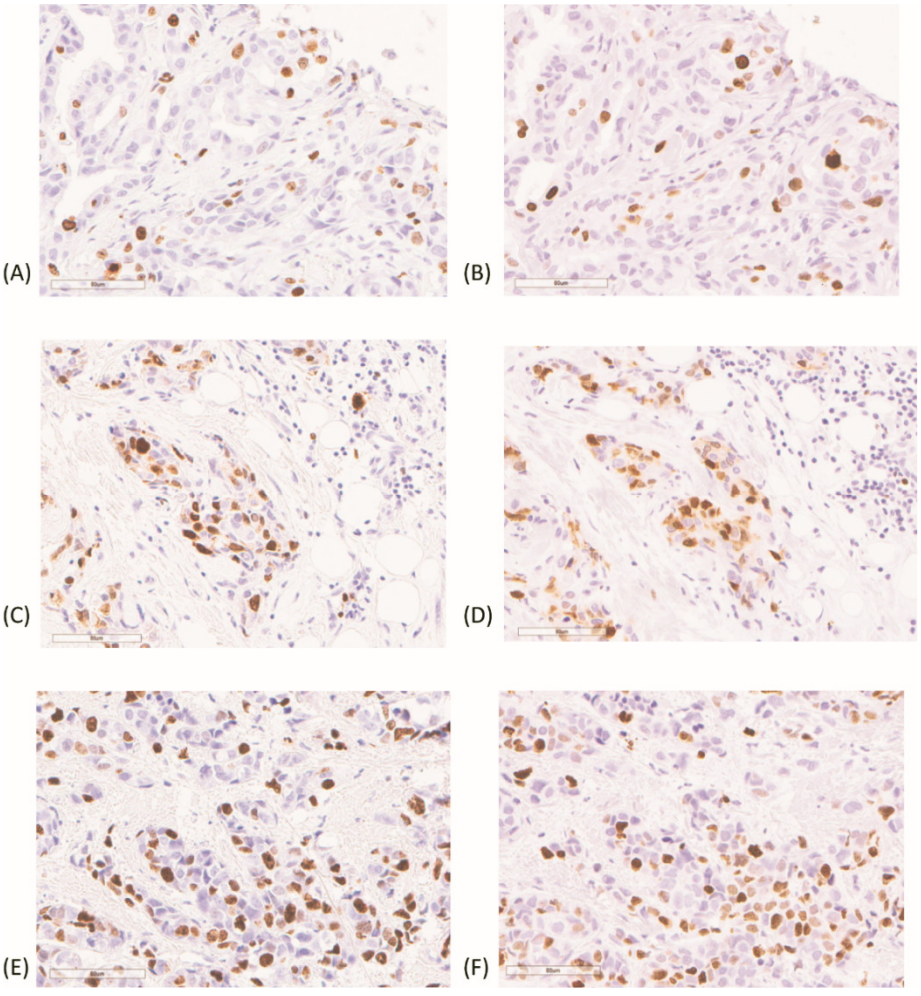
All precision studies achieved 95% lower-bound confidence interval (LBCI) of >90% for NPA, PPA, and OA. Analysis results for internal studies are summarized in Supplemental Table 4. The results for internal inter-/intraobserver studies are illustrated in Supplemental Figure 5. In 3 instances a specimen was determined to have <200 viable tumor cells required for evaluation, leaving 537 observations included in the

analysis. Interobserver reproducibility testing achieved NPA, PPA, and OA point estimates (95% LBCI) of 98.9% (97.2%), 95.2% (91.7%), and 97.2% (95.4%), respectively. Intraobserver reproducibility achieved NPA, PPA, and OA point estimates (95% LBCI) of 99.3% (98.3%), 96.8% (94.4%), and 98.1% (96.8%), respectively.

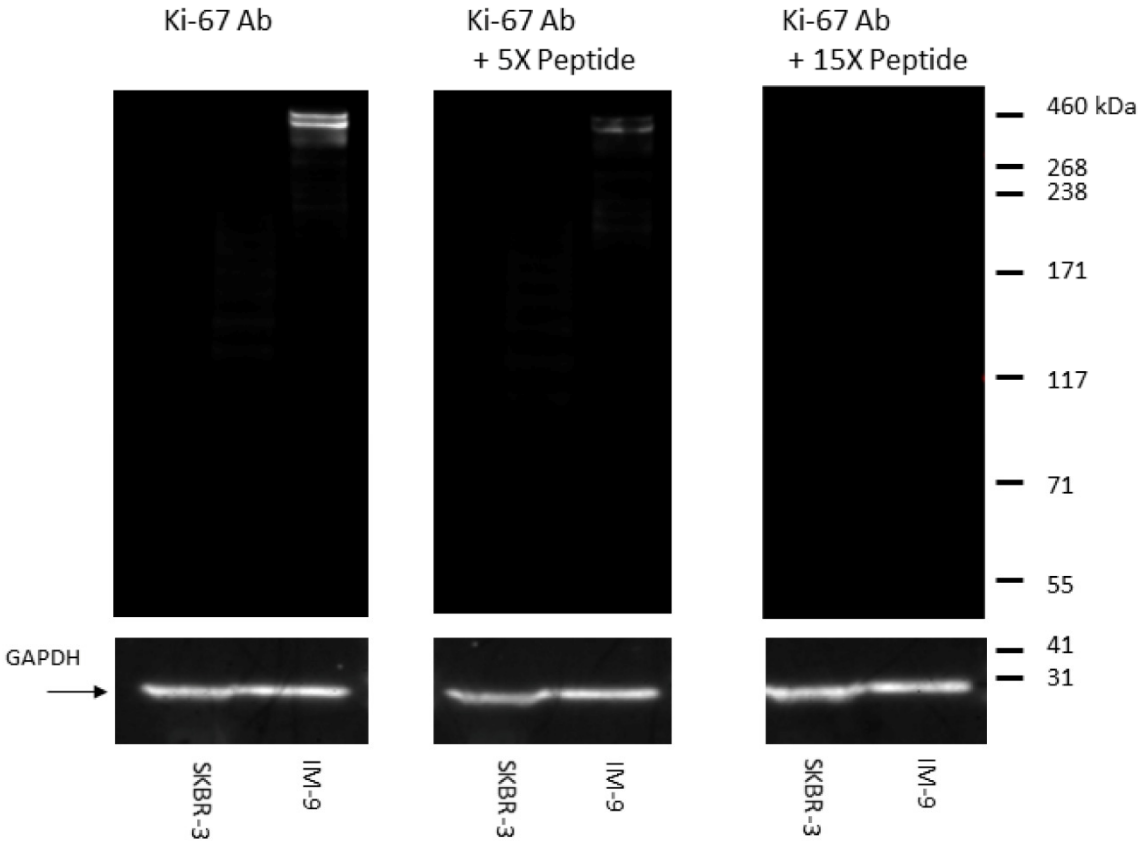
Additional exploratory scoring was performed on a subset of early studies prior to assay validation to assess the potential utility of a hot spot scoring approach. Hot spot scores tended to produce larger average standard deviations across low and high Ki-67 expressing tumors and resulted in lower concordance among observers (data not shown), which could in part be explained by less stringent observer training. These findings are congruent with previous observations of similar scoring comparisons.[11] As described in the Results section of the manuscript, score variability increased as Ki-67 expression range increased, but there was not marked interobserver variability by observer/read combination (Supplemental Figure 6).

3 Supplemental Tables and Figures

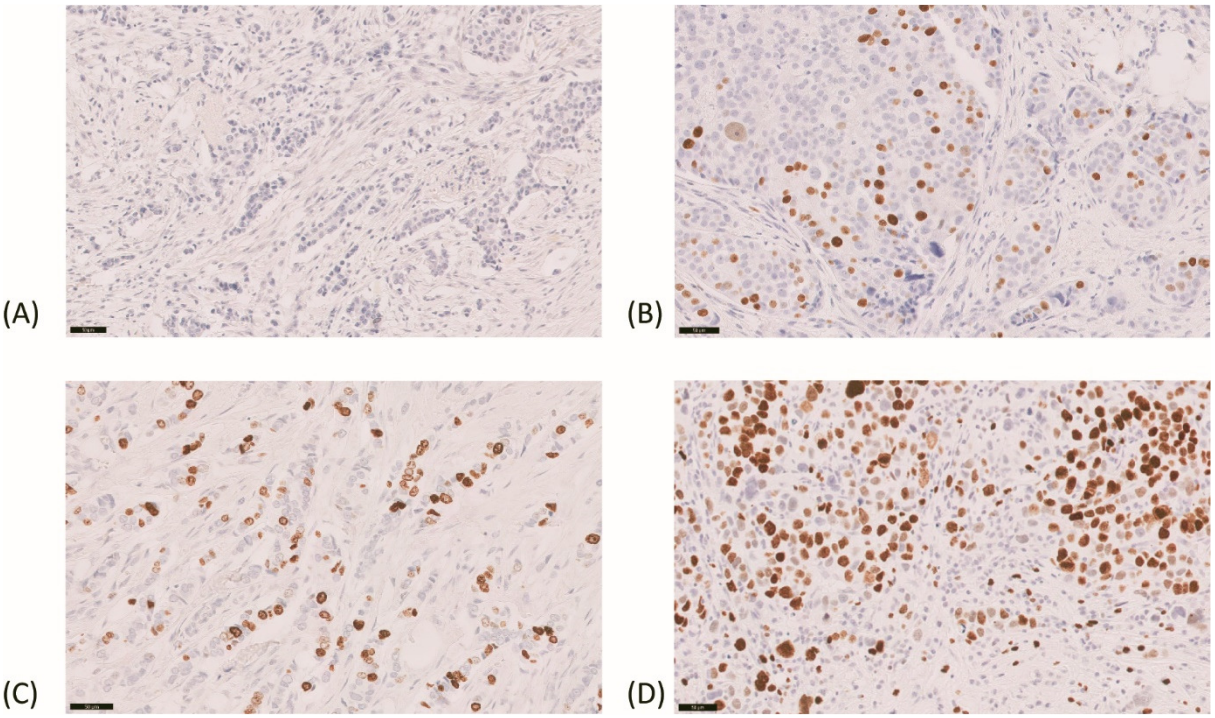
3.1 Supplemental Figure 1. Qualitative comparison of the prototype investigational Ki-67 assay with the neoMONARCH LDT. Representative images were captured from tumors with a Ki-67 positivity near the ≥20% cutoff. Breast carcinoma tissue sections in the left column were assayed with the prototype assay (A, 12%; C, 25%; E, 35%); nearby tissue levels were assayed with the LDT as shown on the right (B, 10%; D, 15%; F, 35%). Respective scores correspond to the entire specimen. Images A-F are representative fields captured using a 20x objective; scale bar is 80 µm. LDT, laboratory developed test.



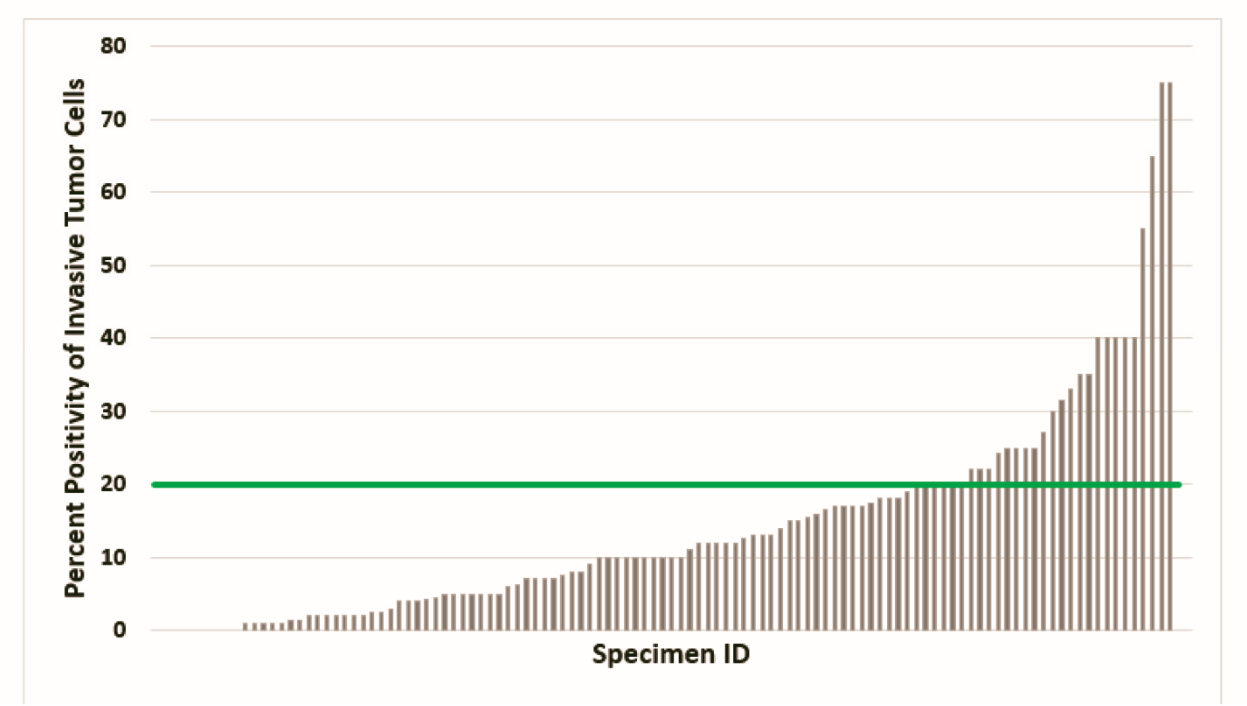
3.2 Supplemental Figure 2. Ki-67 protein is expressed in IM-9 cells, a cell line from which the immunizing protein for MIB-1 was derived,[3] and is not expressed in SKBR-3 cells, a cell line with a very low level of Ki-67 RNA expression.[4] A peptide comprising a region within the immunizing protein when pre-incubated with the antibody reduces the binding to the IM-9 protein. The GAPDH loading control demonstrates equal loading of the cell lysates. GAPDH, glyceraldehyde-3-phosphate dehydrogenase.



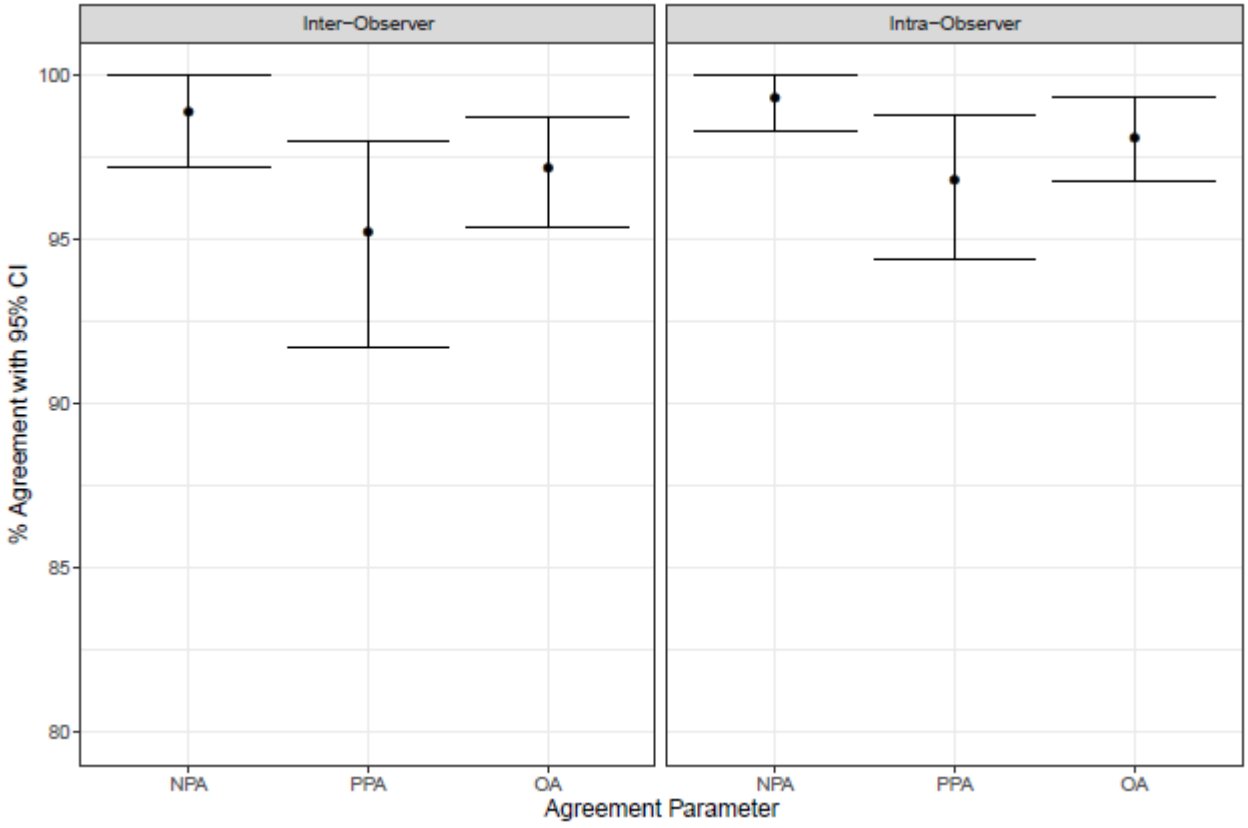
3.3 Supplemental Figure 3. Expression levels in breast carcinoma tissue bank specimens stained with the Investigational Use Only Ki-67 assay. Ki-67 was detected over a relevant range of breast cancer FFPE specimens. Images shown are from archived tumor samples with a Ki-67 assay score of 0% (A); a Ki-67 assay score of 19% (B); a Ki-67 assay score of 28% (C); and a Ki-67 assay score of 52% (D), respectively (20x objective; scale bar is 50 µm). FFPE, formalin-fixed, paraffin-embedded.



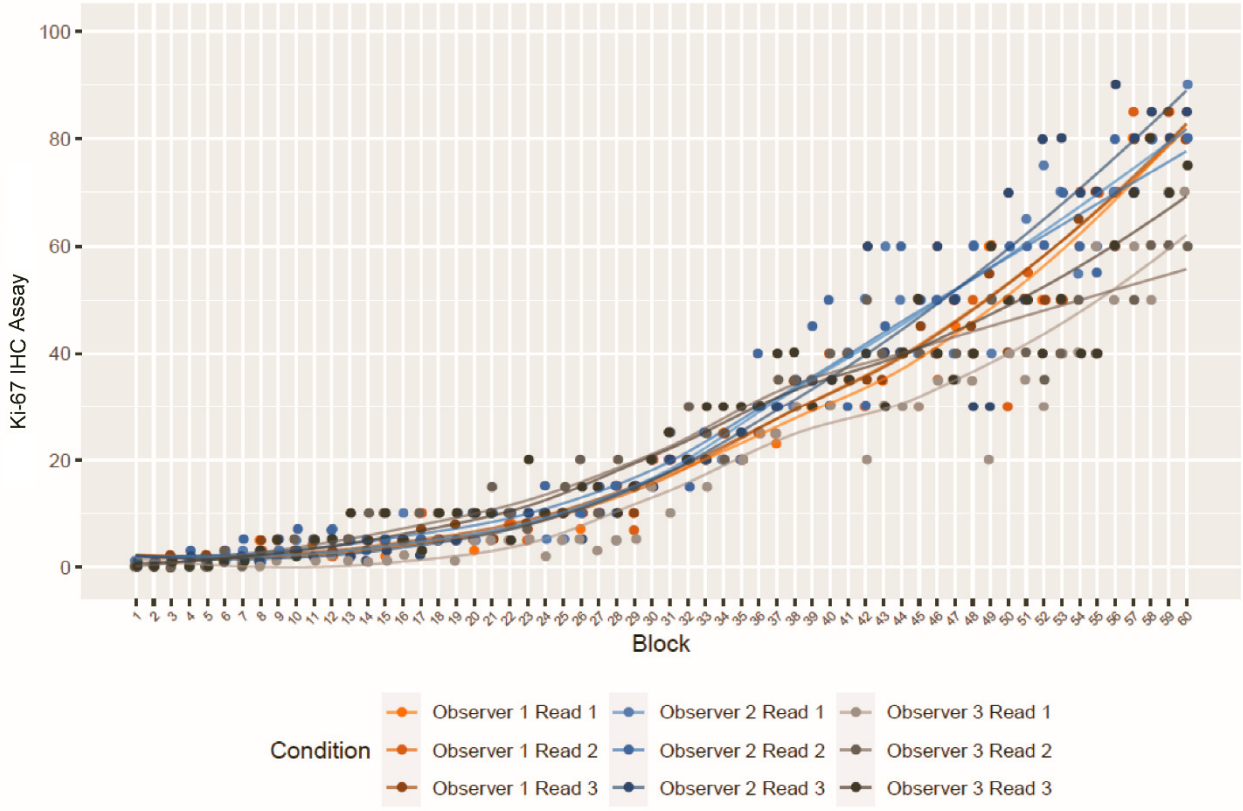
3.4 Supplemental Figure 4. Investigational Use Only Ki-67 IHC assay sensitivity. Dynamic range distribution of scores across 113 breast cancer samples, including resections and core needle biopsies. The green line indicates the diagnostic cutoff (≥20%).



3.5 Supplemental Figure 5. Summary of percent agreement for Investigational Use Only Ki-67 IHC assay observer precision studies performed internally at Agilent Technologies. CI, confidence interval; IHC, immunohistochemistry



3.6 Supplemental Figure 6: LOESS plot of external reproducibility interobserver Ki-67 continuous scores grouped by observer/read combination. LOESS lines demonstrate average trends over interobserver data using locally weighted regression. IHC, immunohistochemistry; LOESS, locally estimated scatterplot smoothing.



3.7 Supplemental Table 1: Concordance Between the Prototype Assay and the neoMONARCH Laboratory Developed Test When Assessed with the Same Scoring Method by a Single Observer.

		LDT		
		Positive	**Negative**	**Total**
Ki-67 IHC Prototype	**Positive**	24	4	28
	Negative	0	32	32
	Total	24	36	60

Abbreviations: IHC, immunohistochemistry; LDT, laboratory developed test.

Note: Seven cores were not evaluable.

3.8 Supplemental Table 2: Summary of Investigational Use Only Ki-67 IHC Assay Normal Tissue Reactivity

Tissue Type (# tested)	Positive Cell Staining: Tissue Elements	Nonspecific Staining
Adrenal (3)	2/3 Scattered adrenal cortex cells	0/3
Bone marrow (3)	2/3 Marrow cells	0/3
Breast (3)	2/3 subset ductal epithelial cells* 2/3 rare myoepithelial cells*	0/3
Cerebellum (3)	0/3	0/3
Cerebrum (3)	1/3 Rare oligodendrocytes	0/3
Cervix (3)	2/3 Rare epithelial cells 1/3 parabasal squamous cells	0/3
Colon (3)	3/3 Crypt epithelial cells	0/3
Esophagus (3)	3/3 Parabasal squamous cells	0/3
Kidney (3)	3/3 Rare tubule cells	0/3
Liver (3)	3/3 Rare hepatocytes	0/3
Lung (3)	3/3 Rare type 1 alveolar cells 1/3 Pulmonary macrophages	0/3
Mesothelial cells (3)	1/3 Rare mesothelial cells	0/3
Muscle, cardiac (3)	1/3 Few cardiac myocytes	0/3
Muscle, skeletal (3)	0/3	0/3
Nerve, peripheral (3)	0/3	0/3
Ovary (3)	2/3 Follicle cyst lining cells 1/3 Rare ovarian stromal cells	0/3
Pancreas (3)	3/3 Rare acinar cells	0/3
Parathyroid (3)	3/3 Rare endocrine cells/Chief cells	0/3
Pituitary (3)	1/3 Lymphocytes 2/3 Pituicytes	0/3
Prostate (3)	3/3 Rare epithelial cells and stromal cells	0/2
Salivary gland (3)	1/3 Rare acinar epithelial cells	0/3
Skin (3)	3/3 Suprabasal squamous cells	0/3
Small intestine (3)	2/3 Crypt epithelium	0/3
Spleen (3)	3/3 Few red pulp cells 1/3 white pulp cells	0/3
Stomach (3)	3/3 Gastric epithelial cells	0/3
Testis (3)	3/3 Spermatogonia	0/3
Thymus (3)	3/3 Thymic cortex	0/3
Thyroid (3)	0/3	2/3 Scattered interspersed inflammatory cells
Tonsil (3)	3/3 Squamous parabasal cells 3/3 Germinal centers 3/3 interfollicular cells	0/3

Tissue Type (# tested)	Positive Cell Staining: Tissue Elements	Nonspecific Staining
Urinary bladder (3)	3/3 Urothelial cells	2/3 Interspersed inflammatory cells
Uterus (3)	3/3 Endometrial epithelium and stroma	0/3

Abbreviations: IHC, immunohistochemistry.

*May serve as an internal tissue control element.

3.9 Supplemental Table 3: Agreements and 95% Confidence Intervals of Investigational Use Only Ki-67 IHC Assay Robustness and Heterogeneity Studies

Study	Number of Comparisons (Total)	% Agreement (95% CI)
Overnight/over-weekend (delayed start)	270	NPA 99.3% (97.8%-100.0%) PPA 95.6% (90.4%-100.0%) OA 97.4% (94.4%-99.6%)
1x Target Retrieval Solution pH (pH 5.9-6.5)	384	NPA 100% (98.4%-100.0%) PPA 97.9% (94.4%-100.0%) OA 99.2% (97.9%-100.0%)
Tissue section thickness (3 µm – 6 µm)	384	NPA 98.8% (96.7%-100.0%) PPA 97.2% (91.7%-100.0%) OA 98.2% (95.6%-100.0%)
Intrablock heterogeneity	144	NPA 94.0% (86.9%-100.0%) PPA 100% (94.0%-100.0%) OA 96.5% (92.4%-100.0%)
Intracase heterogeneity	50	NPA 94.1% (85.3%-100.0%) PPA 100% (80.6%-100.0%) OA 96.0% (90.0%-100.0%)

Abbreviations: CI, confidence interval; IHC, immunohistochemistry; NPA, negative percent agreement; OA, overall percent agreement; PPA, positive percent agreement.

3.10 Supplemental Table 4: Agreements and 95% Confidence Intervals of Investigational Use Only Ki-67 IHC Assay Precision Studies

Internal Repeatability	Number of Comparisons (Total)	% Agreement (95% CI)
Interday	480	NPA 98.3% (96.2%-100.0%) PPA 99.2% (97.5%-100.0%) OA 98.8% (97.5%-99.8%)
Interinstrument	480	NPA 98.8% (96.7%-100.0%) PPA 97.5% (94.6%-100.0%) OA 98.1% (96.5%-99.6%)
Interlot (Ki-67 IHC)	360	NPA 98.9% (97.2%-100.0%) PPA 97.2% (92.8%-100.0%) OA 98.1% (95.8%-99.7%)
Interlot (accessory reagents)	192	NPA 95.1% (88.2%-100.0%) PPA 100% (95.9%-100.0%) OA 97.4% (93.8%-100.0%)
Intrainstrument/rack/day	480	NPA 98.8% (96.7%-100.0%) PPA 97.5% (94.6%-100.0%) OA 98.1% (96.2%-99.6%)

Abbreviations: CI, confidence interval; IHC, immunohistochemistry; NPA, negative percent agreement; OA, overall percent agreement; PPA, positive percent agreement.

4 Supplemental References

1. Hurvitz SA, Martin M, Press MF, et al. Potent cell-cycle inhibition and upregulation of immune response with abemaciclib and anastrozole in neoMONARCH, phase II neoadjuvant study in HR(+)/HER2(-) breast cancer. *Clin Cancer Res.* 2020;26(3):566-580.
2. FDA. Guidance for Submission of Immunohistochemistry Applications to the FDA 1998.
3. Key G, Becker MH, Baron B, et al. New Ki-67-equivalent murine monoclonal antibodies (MIB 1-3) generated against bacterially expressed parts of the Ki-67 cDNA containing three 62 base pair repetitive elements encoding for the Ki-67 epitope. *Lab Invest.* 1993;68(6):629-636.
4. Klijn C, Durinck S, Stawiski EW, et al. A comprehensive transcriptional portrait of human cancer cell lines. *Nat Biotechnol*. 2015;33(3):306-312.
5. Cattoretti G, Becker MH, Key G, et al. Monoclonal antibodies against recombinant parts of the Ki-67 antigen (MIB 1 and MIB 3) detect proliferating cells in microwave-processed formalin-fixed paraffin sections. *J Pathol*. 1992;168(4):357-363.
6. Hsu CY, Yang CF, Liao LR, et al. Tonsil surface epithelium is ideal for monitoring Ki-67 immunohistochemical staining. *Histopathology*. 2013;63(6):810-816.
7. Kurbel S, Dmitrovic B, Marjanovic K, et al. Distribution of Ki-67 values within HER2 & ER/PgR expression variants of ductal breast cancers as a potential link between IHC features and breast cancer biology. *BMC Cancer*. 2017;17(1):231.

8. Regan MM, Francis PA, Pagani O, et al. Absolute benefit of adjuvant endocrine therapies for premenopausal women with hormone receptor-positive, human epidermal growth factor receptor 2-negative early breast cancer: TEXT and SOFT trials. *J Clin Oncol*. 2016;34(19):2221-2231.
9. Allison KH, Hammond MEH, Dowsett M, et al. Estrogen and progesterone receptor testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists Guideline Update. *Arch Pathol Lab Med*. 2020;144(5):545-563.
10. Wolff AC, Hammond MEH, Allison KH, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *Arch Pathol Lab Med*. 2018;142(11):1364-1382.
11. Leung SCY, Nielsen TO, Zabaglo L, et al. Analytical validation of a standardized scoring protocol for Ki67: phase 3 of an international multicenter collaboration. *NPJ Breast Cancer.* 2016;2:16014.