# Ki67 assessment based on the International Ki67 Working Group (IKWG) recommendations and correlation with 21-gene assay results in a large integrated health care system: We might not be there yet.

Veronica C Shim[1], Robin J Baker[2], Wen Jing[2], Sally S Agersborg[3], Thomas K Lee[3], Wamda Goreal[3], Róisín Puentes[3], Ninah Achacoso[4], Catherine Lee[4], Marvella Villaseñor[4], Amy Lin[2], Malathy Kapali[5], and Laurel A Habel[4]

[1]The Permanente Medical Group, Oakland, CA; [2]The Permanente Medical Group, San Francisco, CA; [3]NeoGenomics Laboratories, Inc., Aliso Viejo, CA; [4]The Permanente Medical Group, Division of Research, Oakland, CA; [5]The Permanente Medical Group, Sacramento, CA

## BACKGROUND

While substantial evidence indicates that Ki67, a marker of proliferation, is strongly associated with breast cancer outcomes, its clinical utility has been limited given concerns about scoring inter-rater reliability and appropriate cut points.

Nonetheless, Ki67 has been used in multiple clinical trials and results from POETIC indicated that low baseline Ki67 (ie, <10%) predicts good prognosis in postmenopausal women with hormone-sensitive, early breast cancer.

Further, the International Ki67 Working Group (IKWG) has developed website-based training materials to improve the reproducibility of Ki67 scoring by immunohistochemistry (IHC) and recently considered the marker to be sufficiently validated to support treatment decisions in early ER+ breast cancer (≤ 5 no chemotherapy, ≥30 chemotherapy indicated).

## AIMS

Our aims were to examine Ki67 scoring reproducibility following IKWG training and the extent to which low or high scores could be used to accurately identify patients with low or high 21-gene assay Recurrence Scores (RS) who could selectively avoid this test.

## METHODS

Setting and study population: The study was conducted within the membership of Kaiser Permanente Northern California (KPNC), an integrated health care system with over 4.4 million enrollees. We included a random sample of women aged 50+ years at diagnosis of node-negative, ER+, PR+, HER2- breast cancer with the 21-gene assay done on their surgical specimen from 2018-2020 (n=307).

Ki67 staining, training and scoring. We retrieved archived core biopsy specimens, which were sent to NeoGenomics for Ki67 staining (Dako clone MIB1) and scoring by image analysis (IA) using the hot spot counting method. In addition, two KPNC pathologists specializing in IHC scoring underwent IKWG training and independently scored all slides using the global counting method, blinded to each other and to readings by AI; weighted Ki67 scores were calculated.

Analysis. We examined inter-rater reproducibility across pathologists using intraclass correlation (ICC) and Kappa statistics. We also examined the percent of patients with low Ki67 scores (≤ 5, <10) by each pathologist and by IA who also had low RS (<26) and the percent who had high Ki67 scores (≥30) who also had high RS (≥26).

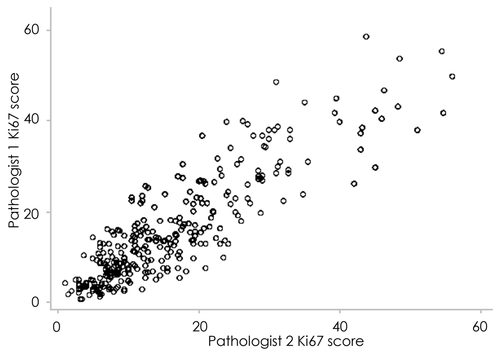
## RESULTS

**Selected characteristics of breast cancer cases, overall and stratified by Ki67 and 21-gene recurrence scores**

| Characteristic | | Overall N=307 (Column %) | Ki67 ≤5 N=33 (Row %) | Ki67 6-29 N=205 (Row %) | Ki67 ≥30 N=69 (Row %) | RS < 26 N=267 (Row %) | RS ≥ 26 N=40 (Row %) |
|---|---|---|---|---|---|---|---|
| **Age at diagnosis** | 50-59 | 113 (36.8) | 16 (14.2) | 73 (64.6) | 24 (21.2) | 94 (83.2) | 19 (16.8) |
| | 60-69 | 142 (46.3) | 13 (9.2) | 95 (66.9) | 34 (23.9) | 126 (88.7) | 16 (11.3) |
| | 70-79 | 50 (16.3) | 4 (8.0) | 36 (72.0) | 10 (20.0) | 45 (90.0) | 5 (10.0) |
| | 80+ | 2 (0.7) | 0 | 1 (50.0) | 1 (50.0) | 2 (100.0) | 0 |
| **Race** | Non-Hispanic white | 189 (61.6) | 21 (11.1) | 127 (67.2) | 41 (21.7) | 165 (87.3) | 24 (12.7) |
| | Black | 10 (3.3) | 2 (20.0) | 7 (70.0) | 1 (10.0) | 9 (90.0) | 1 (10.0) |
| | Asian | 61 (19.9) | 3 (4.9) | 41 (67.2) | 17 (27.9) | 52 (85.2) | 9 (14.8) |
| | Hispanic | 34 (11.1) | 5 (14.7) | 23 (67.6) | 6 (85.3) | 29 (85.3) | 5 (14.7) |
| | Other/unknown | 13(3.9) | 2 (15.4) | 7 (53.8) | 4 (30.8) | 11 (91.7) | 1 (8.3) |
| **Tumor size** | <1 cm (T1a) | 6 (2.0) | 2 (33.3) | 3 (50.0) | 1 (16.7) | 6 (100.0) | 0 |
| | 1-1.9 cm (T1b, T1c) | 207 (67.4) | 27 (13.0) | 144 (69.6) | 36 (17.4) | 186 (90.1) | 21 (9.9) |
| | 2-5.0 cm (T2) | 86 (28.0) | 4 (4.7) | 52 (60.5) | 30 (34.9) | 69 (80.2) | 17 (19.8) |
| | 5.1 + cm (T3) | 8 (2.6) | 0 | 6 (75.0) | 2 (25.0) | 6 (75.0) | 2 (25.0) |
| **Tumor grade** | Low | 109 (35.5) | 25 (22.9) | 81 (74.3) | 3 (2.8) | 106 (97.2) | 3 (2.8) |
| | Intermediate | 166 (54.1) | 8 (4.8) | 111 (66.9) | 47 (28.3) | 143 (86.1) | 23 (13.9) |
| | High | 32 (10.4) | 0 | 13 (6.3) | 19 (27.5) | 18 (56.2) | 14 (43.8) |
| **ER** | 1-10 | 2 (0.7) | 0 | 2 (100.0) | 0 | 1 (50.0) | 1 (50.0) |
| | 11-20 | 1 (0.3) | 0 | 0 | 1 (100.0) | 0 | 1 (100.0) |
| | 21-30 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 31+ | 304 (99.0) | 33 (10.9) | 203 (66.8) | 68 (22.4) | 266 (87.5) | 38 (12.5) |
| **PR** | 1-10 | 39 (12.7) | 5 (12.8) | 22 (56.4) | 12 (30.8) | 21 (53.8) | 18 (46.2) |
| | 11-20 | 15 (4.9) | 3 (20.0) | 8 (53.3) | 4 (26.7) | 13 (86.7) | 2(13.3) |
| | 21-30 | 17 (5.5) | 0 | 9 (52.9) | 8 (47.1) | 12 (70.6) | 5 (29.4) |
| | 31+ | 235 (76.6) | 25 (10.6) | 165 (70.2) | 45 (19.2) | 221 (94.0) | 14 (6.0) |
| **BRCA/other positive** | Yes | 5 (1.6) | 0 | 3 (60.0) | 2 (40.0) | 3 (60.0) | 2 (40.0) |
| | No | 302 (98.4) | 33 (10.9) | 202 (66.9) | 67 (22.2) | 264 (87.4) | 38 (12.6) |
| **Low tumor grade plus:*** | Ki67 ≤5 | 25 (8.1) | | | | 24 (96.0) | 1 (4.0) |
| | Ki67 <10 | 51 (16.6) | | | | 50 (98.0) | 1 (2.0) |

*Low grade and Ki67 <5; low grade and Ki67<10

**Comparison of pathologist 1 weighted global Ki67 scores to pathologist 2 global weighted global Ki67 scores**




ICC for Ki67 scores by the two pathologists was 0.82 (95% CI 0.78-0.85).

When categorizing scores as ≤5, 6-29, 30+, the Kappa was 0.67 (95% CI 0.56-0.78).

**Ki67 scores by reader and 21-gene recurrence scores (RS)**

| Reader | Ki67 scores | Total N=307 (Column %) | RS <26 N=267 (Row %) | RS ≥26 N=40 (Row %) |
|---|---|---|---|---|
| **Visual Scoring – Pathologist 1** | Weighted ≤ 5 | 49 (16.0) | 45 (91.8) | 4 (8.2) |
| | Weighted 6-29 | 217 (70.7) | 198 (91.2) | 19 (8.8) |
| | Weighted ≥ 30 | 41 (13.4) | 24 (58.5) | 17 (41.5) |
| | Weighted < 10 | 120 (39.1) | 115 (95.8) | 5 (4.2) |
| | Weighted ≥ 10 | 187 (60.9) | 152 (81.3) | 35 (18.7) |
| **Visual Scoring – Pathologist 2** | Weighted ≤ 5 | 27 (8.8) | 25 ( 92.6) | 2 ( 7.4) |
| | Weighted 6-29 | 245 (79.8) | 225 (91.8) | 20 (8.2) |
| | Weighted ≥ 30 | 35 (11.4) | 17 ( 48.6) | 18 (51.4) |
| | Weighted < 10 | 111 (36.2) | 105 (94.6) | 6 (5.4) |
| | Weighted ≥ 10 | 196 (63.8) | 162 (82.6) | 34 (17.4) |
| **Image Analysis** | ≤ 5 | 33 (10.7) | 30 (90.9) | 3 ( 9.1) |
| | 6-29 | 205 (66.8) | 187 (91.2) | 18 (8.8) |
| | ≥ 30 | 69 (22.5) | 50 (72.5) | 19 (27.5) |
| | < 10 | 81 (26.4) | 77 (95.1) | 4 (4.9) |
| | ≥ 10 | 226 (73.6) | 190 (84.1) | 36 (15.9) |

## KEY FINDINGS

- Approximately 83% of patients were ages 50-69 years (median 63), 61% were non-Hispanic white and 93% were stage 1A.
- The ICC for Ki67 scores (log-transformed) by the two pathologists was 0.82 (95% CI 0.78-0.85); using cut points of ≤5, 6-29, 30+, the Kappa was 0.67 (95% CI 0.56-0.78).
- Among patients with Ki67 scores of <10% by IA (n=81), pathologist 1 (n=120) or pathologist 2 (n=111), the percent with a RS of <26 was 95.1% for IA, 95.8% for pathologist 1, and 94.6% for pathologist 2.
- Among patients with Ki67 scores ≤5, the percentages were 90.9%, 92.6% and 91.8% for IA, pathologist 1 and pathologist 2, respectively.
- Among patients with Ki67 scores ≥30 by IA (n=69), pathologist 1 (n=41) or pathologist 2 (n=35), the percent who had a RS >26 was 27.5% for IA, 41.5% for pathologist 1 and 51.4% for pathologist 2.
- Results are improved if we excluded all 50 patients with weak PR by IHC (1-10%); for example, among patients with Ki67 scores of <10%, the percent with a RS of <26 was 97.1% for IA, 98.1% for pathologist 1, and 97.9% for pathologist 2.

## CONCLUSIONS

Among women aged 50+ years with node-negative, ER+PR+HER2- breast cancer in our setting, approximately 5-10% of patients with Ki67 scores of ≤5% or <10% on core biopsies would have high RS (≥26) on surgical specimens and over 48% of cases with Ki67 scores ≥30 would have low RS (<26), which may be insufficiently accurate for avoiding the 21-gene or other multi-gene assays.

## FUTURE DIRECTIONS

Future studies are needed to examine whether restricting Ki67 testing to ER+, HER2- patients with PR >10% would improve its clinical validity.

## FUNDING

The study was funded through a Delivery Science Grant from The Permanente Medical Group (TPMG).

## CONTACT

Dr. Veronica C Shim: Veronica.Shim@kp.org
Dr. Laurel A Habel: Laurel.Habel@kp.org