Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

GUVAKOVA FISHing out HER2-low breast cancer for HER2-targeted therapies

Marina A. Guvakova

guvakova@pennmedicine.upenn.edu

Article

Keywords: HER2-low breast cancer, FISH assay, HER2-targeted therapies, model-based clustering, Gaussian mixture modeling

Posted Date: July 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9859221/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

HER2-targeted therapies require precise tumor diagnosis due to variability in HER2 expression.

This study introduces a fully automated approach for analyzing fluorescent *in situ* hybridization (FISH) to determine HER2 status and resolve inconsistencies in immunohistochemistry (IHC) assessments. The ***g3mclass*** software was employed to identify HER2-low cases among HER2-negative breast cancers. Rather than imposing a predetermined number of FISH result groups, the software generated probability models using FISH data from 515 invasive cancers with diverse IHC scores and 52 non-cancerous tissues. The software established a normal HER2 copy number threshold of 2.11 and assigned each sample to the most probable group. Two distinct HER2-negative groups were identified. The HER2-normal group (217 cases) exhibited HER2 levels below the amplification threshold, consistent with normal gene activity. The HER2-low group (125 cases) demonstrated HER2 levels above the normal range ($P < 0.0001$) but below the amplification threshold. IHC scores did not distinguish between HER2-low and HER2-normal groups, although the HER2-low group contained fewer IHC 0 cases (82.4% vs. 95.9%) and more IHC 1+ cases (15.2% vs. 3.2%). This software-assisted FISH methodology facilitates the identification of HER2-low tumors that may benefit from novel therapies and excludes those unlikely to respond, thereby supporting personalized treatment strategies.

INTRODUCTION

For two decades, medical experts have assessed epidermal growth factor receptor 2 (HER2) status to categorize invasive breast cancer (IBC) as either HER2-positive or HER2-negative, using the U.S. Federal Food and Drug Administration (FDA)-approved immunohistochemistry (IHC) and *in situ* hybridization (ISH) methods. A HER2-positive status indicates that breast cancer cells have extra copies of the HER2 gene, which encodes the HER2 protein, and overexpress this protein, leading to a more aggressive disease [1]. Only patients with HER2-positive tumors were considered eligible for anti-HER2 therapy, such as trastuzumab (Herceptin) and pertuzumab (Perjeta).

Recently, clinical trial investigators have expanded the spectrum of HER2 positivity by introducing the terms HER2-low and HER2-ultralow to identify patients who may benefit from a new generation of HER2-directed antibody-drug conjugates (ADCs). The FDA approved fam-trastuzumab-deruxtecan-nxki (Enhertu, Daiichi Sankyo, Inc.) for the treatment of HER2-low and HER2-ultralow unresectable or metastatic hormone receptor-positive breast cancer [2–4]. Both terms refer to HER2 ISH-negative tumors, with HER2-low defined as IHC 1 + or 2+, and HER2-ultralow as IHC 0 with weak membrane staining [5]. The rest of HER2-negative breast cancer with an IHC score of 0 has been referred to as the HER2-0 (a.k.a. HER2-null) [6]. Patients with HER2-0 tumors should not be selected for anti-HER2 targeted therapies because HER2-0 tumor cells have HER2 levels similar to those in normal breast epithelium [7]. Although clinical rationale dictates the distinction among the three HER2-negative subcategories, standard IHC score-based subgrouping is challenging in real-world settings. Many studies have reported poor concordance among pathologists worldwide in classifying tumors as either HER2 IHC 1 + or 0,

often due to subtle differences in IHC staining [8–13]. An urgent need remains for a new assay or a testing algorithm to reliably evaluate the full spectrum of HER2 expression in routine clinical practice.

HER2 ISH is a diagnostic technique used to assess the degree of HER2 gene amplification. It overcomes many of the interpretive limitations of HER2 IHC. Furthermore, the continuous measurements in ISH tests enable analysis of the full spectrum of HER2 gene expression. Reports have suggested repurposing ISH assays as an add-on tool for selecting HER2-low tumors by using the statistical upper bound of benign breast tissue as the lower limit for HER2-low [14, 15]. This premise is reasonable because, intuitively, HER2-low/ultralow tumors should have HER2 expression higher than that of normal mammary glands and HER2-0 cancers, but lower than that of HER2-overexpressing tumors. Currently, the integration of such an approach into clinical practice is complicated by overlapping ISH assay values between IHC 0 and IHC 1 + cases [14], lack of automated computation for the nonparametric distribution of ISH results in cancer [16, 15], and common practice of performing HER2 ISH only as a reflex testing for HER2 IHC2 + cases [17].

We hypothesized that subgroups within the HER2-negative category could be inferred from HER2 ISH assay data. To test this hypothesis, we utilized a large HER2 ISH dataset comprising 515 clinical samples of invasive breast cancer across all IHC scores obtained in proficiency tests (PTs) for breast cancer pathology. We employed **g3mclass**, an advanced statistical software package specifically developed and validated for biomarker grouping [18, 19]. The **g3mclass** algorithm uses a model-based clustering approach to analyze the distribution of biomarker data across a spectrum of assay results in both test (cancer) and reference (noncancer) samples. It applies Gaussian mixture modeling to calculate data-driven cutoffs and identify distinct groups in the data without requiring the number of groups to be specified in advance. Each sample is then assigned to the group with the highest probability based on its observed biomarker profile, enabling an unsupervised, reproducible classification process.

RESULTS

Automated classification on the HER2 to chromosome 17 centromere (CEP17) ratio

Based on a dual-probe fluorescence ISH (FISH) ratio data, the software sorted 515 invasive breast cancer cores into five groups: 350 (67.96%) cases in the group 0 (cutoff ≤ 1.41), 94 (18.25%) in the group 1 (1.41 < cutoff ≤ 5.03), and 71 (13.79%) in the group 2 and higher (cutoff > 5.03) as illustrated in Fig. 1A and summarized in Table 1. Notably, the software classified 51 of 52 (98%) benign cores into group 0, and one benign case into group 1 **(**Fig. 1B**).** We assessed the guideline-concordant HER2 amplification status of breast cancer samples within each ratio-related group.

Table 1

Automated classification on ratios of average HER2:CEP17 signals per nucleus.

ratio cutoff	**Group 0, n = 350**	**Group 1, n = 94**	**Group 2–4, n = 71**
	≤ 1.41	**≤ 5.03 and > 1.41**	**> 5.03**
group characteristic	No (%)	No (%)	No (%)
HER2/CEP17 ratio < 2.0	350 (100)	29 (30.9)	0 (0.0)
HER2/CEP17 ratio ≥ 2.0	0 (0.0)	65 (69.1)	71 (100)
HER2 signal < 4,	320 (91.4)	31 (33)	0 (0.0)
HER2 signal ≥ 4 and < 6	8 (2.0)	20 (21.3)	0 (0.0)
HER2 signal ≥ 6	22 (6.3)	43 (45.7)	71 (100)
IHC 0	299 (85.4)	33 (35.1)	0 (0.0)
IHC 1+	25 (7.1)	2 (2.1)	0 (0.0)
IHC 2+	3 (0.9)	7 (7.4)	1 (1.4)
IHC3+	23 (6.6)	52 (55.3)	70 (98.6)
HER2 FISH status	**non-amplified and non-classical amplified**	**amplified and non-amplified**	**classical amplified**

In group 0, where the software assigned most of the benign cases, all 350 (100%) cases would be considered FISH-negative based on a ratio cutoff of < 2.0, per the manufacturer's guidelines for the dual-probe PathVision FISH kit. Considering a 2018 ASCO-CAP two-tiered algorithm for dual-probe ISH testing and non-classical FISH results, 320 (91.4%) tumors were definitively FISH negative (ratio < 2.0, signal < 4.0), whereas 8 (2%) would be regarded as equivocal (ratio < 2.0, 4.0 ≤ signal < 6.0), and 22 (6.3%) of cases (ratio < 2.0, HER2 signal > 6.0) as non-classical FISH positive test results. The latter were co-amplified cases, characterized by increased HER2 and CEP17 signals, resulting in a negative ratio. The FISH findings, which align with IHC data, are summarized in Table 1. Most group 0 cancers were IHC 0 (299, 85.4%) and IHC 1+ (25, 7.1%); however, 3 (0.9%) cases were IHC 2 + and 23 (6.6%) were IHC 3+. Although group 0 primarily included non-amplified HER2 IHC-negative cases, the presence of co-amplified cases with low ratios led to discordance with the HER2 IHC status.

Group 1, with statistically higher ratios than those in group 0, included a mixture of HER2-amplified and non-amplified cases (Table 1). There was a trend towards a higher rate of amplified tumors, with a ratio greater than 2.0 in 65 (69.1%) cases and a HER2 signal greater than 6.0 in 43 (45.7%) cases. The range of IHC scores from 0 in 35 (37.2%) to 3 + in 50 (53.2%) reflected the mixed HER2 amplification/overexpression status in group 1.

In groups 2, 3, and 4, all 71 tumor cases were HER2-amplified, with increased HER2:CEP17 ratios greater than 2.0 and HER2 signals greater than 6.0. Concurrently, nearly 99% of tumors in these groups were HER2-positive cases with IHC 3+, and one case was HER2 IHC2+ (Table 1).

Automated classification on the average HER2 copy number, a.k.a. HER2 signal

Using the average HER2 copy numbers (i.e., a mean HER2 signals/cell) data, the **g3mclass** software reclassified 515 cancer samples into nine groups (Fig. 2). The group characteristics are summarized in Table 2. These included 217 (42.14%) in group 0 (cutoff ≤ 2.11), 125 (24.27%) in group 1 (2.11 < cutoff ≤ 3.36), 85 (16.5%) in group 2 (3.36 < cutoff < 10.0), and 88 (17.09%) in the rest of the groups (cutoff ≥ 10).

Table 2
Automated classification on HER2 signal, the averaged HER2 copy number per nucleus.

HER2 signal cutoff	**Group 0, n = 217**	**Group 1, n = 125**	**Group 2, n = 85**	**Group 3–8, n = 88**
	≤ 2.11	**≤ 3.36 and > 2.11**	**< 10.0 and > 3.36**	**≥ 10.0**
group characteristic	No (%)	No (%)	No (%)	No (%)
HER2/CEP17 ratio < 2.0	216 (99.5)	123 (98.4)	31 (36.5)	8 (9.1)
HER2/CEP17 ratio ≥ 2.0	1 (0.5)	2 (1.6)	54 (63.5)	80 (90.9)
HER2 signal < 4,	217 (100)	125 (100)	9 (10)	0 (0.0)
HER2 signal ≥ 4 and < 6	0 (0)	0 (0.0)	28 (33)	0 (0.0)
HER2 signal ≥ 6	0 (0)	0 (0.0)	48 (57)	88 (100)
IHC 0	208 (95.9)	103 (82.4)	21 (24.7)	0 (0.0)
IHC 1+	7 (3.2)	19 (15.2)	1 (1.2)	0 (0.0)
IHC 2+	1 (0.5)	3 (2.4)	4 (4.7)	3 (3.4)
IHC3+	1 (0.5)	0 (0.0)	59 (69.4)	85 (96.6)
HER2 FISH status	**non-amplified**, **normal HER2 signal**	**non-amplified**, **increased HER2 signal**	**classical and non-classical amplified, equivocal**	**classical amplified**

In group 0, 216 (99.5%) of cases were FISH-negative according to a guideline-concordant two-tiered algorithm. HER2 IHC scores in the group 0 cases ranged from 0, 1+, 2+, and 3 + in 95.9%, 3.2%, 0.5%, and 0.5%, respectively. One case of a potential monosomy/non-classical FISH result (borderline ratio 2.0; HER2 signal 2.1) was IHC negative (score 0). One IHC 2 + was FISH-negative (ratio 1.03; HER2 signal 2.05). One discordant IHC-positive (score 3+) FISH-negative (ratio 1.14 and HER2 signal 2.1) case qualified as a false positive. Importantly, 96.15% (50/52) of benign samples belonged to group 0, as depicted in Fig. 2B. Thus, the group 0 invasive breast cancer with the normal average HER2 copy numbers qualified as HER2-negative by both FISH and IHC.

The group 1 breast cancers had significantly higher HER2 signal than group 0 cases ($P < 0.0001$). Nevertheless, 123 (98.4%) were classical HER2 non-amplified cases by both a ratio < 2.0 and the average HER2 copy number < 4.0. The remaining two specimens would be classified as equivocal or non-classical amplified HER2 FISH cases (ratio 2.0; signal 2.15 and ratio 2; signal 2.7), as they were HER2-negative by IHC. Overall, HER2 IHC scores of 0, 1+, 2+, and 3 + were reported in 82.4%, 15.2%, 2.4%, and 0.0% of breast cancer cores, respectively. All IHC2 + were negative for HER2 gene amplification (ratio < 2; signal < 4). Notably, breast cancer with a HER2 low signal (group 1) compared to breast cancer with a normal HER2 signal (group 0) had a lower rate of IHC 0 (82.4% vs. 95.9%), but a higher rate of IHC 1 (15.2% vs. 3.2%) cases. Thus, group 1 was another distinct subgroup of HER2 FISH-negative tumors, characterized by an increased average HER2 copy number and higher HER2 protein expression by IHC.

Compared to group 1, the group 2 cancers had significantly higher HER2 signal ($P < 0.0001$) with heterogeneous ratios (0.9 to 7.62). Classical and non-classical HER2-amplified cases with increased HER2 signal levels formed group 2 (Table 2). Among 85 cases, 35 (41.2%) were classical HER2 amplified cases (ratio ≥ 2.0; HER2 signal ≥ 6.0). Additionally, 43 (50.6%) cases would be considered non-classical FISH results: 17 low-amplified (ratio ≥ 2.0, 4.0 ≤ HER2 signal < 6.0), two monosomies (ratio ≥ 2.0, HER2 signal < 4.0), 13 co-amplified/polysomy (ratio < 2.0, HER2 signal ≥ 6.0), and 11 equivocal (ratio < 2.0, 4.0 < HER2 signal < 6.0). Seven specimens with a ratio < 2.0 and borderline signal values (3.5 < HER2 signal < 3.9) would be classified as classical non-amplified or equivocal cases. The HER2 IHC results were also heterogeneous with scores of 0, 1+, 2+, and 3 + in 24.7%, 1.2%, 4.7%, and 69.4% of cases, respectively. FISH negative (86%, 6/7), equivocal (100%, 11/11), and some low-amplified (6%, 4/17) results were among the IHC 0 cases. Monosomies were IHC1 + and IHC2+. Classical amplified (97%, 34/35), co-amplified (100%, 13/13), and the majority of low-amplified (71%, 12/17) cases contributed to a high percentage of IHC3 + cases. Although most of the group 2 cases were HER2-positive, non-classical FISH results led to discordance with HER2 IHC positivity.

In groups 3 and higher, all 88 tumors with the highest HER2 signal levels (≥ 10) were HER2-amplified, including 9 co-amplified/polysomy cases (ratio < 2.0) and HER2 signal level ≥ 10). These cases were HER2-positive with IHC score of 0, 1+, 2+, and 3 + in 0%, 0%, 3.4%, and 96.6%, respectively.

DISCUSSION

This analytical study addressed the critical problem raised by healthcare professionals. The advent of new HER2-targeted drugs and the evolution of guidelines concerning HER2-low reporting created unforeseen challenges in clinical practice. Although experts agree that HER2-0, HER2-ultralow, and HER2-low breast tumors are considered classic HER2-negative tumors that lack HER2 gene amplification, there is no consensus regarding which method(s) are most accurate for patient stratification [8]. We report the successful performance of a fully automated algorithm for interpreting the FDA-cleared ISH method in HER2 assessment. This novel interpretation approach achieved early unattainable subgrouping of canonical HER2-negative as well as HER2-positive invasive breast cancer. The automated assessment of amplification levels in classic HER2-positive breast cancer may predict the effectiveness of HER2-directed therapies and patient outcomes in advanced cancer [20].

In the present article, we focused on the analysis of the HER2-negative HER2-nonamplified category. The results from a large cohort of patients strongly suggest that historically considered HER2-negative (IHC score 0, 1+, 2+/ISH-) are represented by two distinct groups: HER2-normal and HER2-low. Our findings align with the expert's opinion that it is premature to establish multiple reporting categories for the low end of HER2 expression (HER2-null, HER2-low, and HER2-ultralow) [21].

Our findings provide insights into the diagnostic challenge of assessing HER2-low status through IHC. Early suggestions included evaluating HER2 IHC staining at 40x magnification, providing extensive training for pathologists to improve the reproducibility of scoring HER2-ultralow/low, and applying artificial intelligence models - although these face limits in distinguishing HER2 IHC 0 from 1 + [22–24, 12]. We found that non-amplified, canonical HER2-negative cases were predominantly characterized by an IHC score of 0. A large proportion of these cancers had a normal HER2 gene copy number and exhibited the lowest levels of HER2 protein expression (IHC 0), similar to findings in normal breast epithelia [7]. Another cancer group had elevated levels of the HER2 signal and a nearly 5-fold higher rate of IHC 1+, though some cases were IHC 0. Furthermore, IHC 0 tumors were found among non-classical FISH cases. Thus, using IHC 0 and 1 + to interpret low-range HER2 expression is insufficient to distinguish HER2-normal from HER2-low tumors, as these scores occur in both groups.

We showed that maximum-likelihood classification can stratify HER2-negative breast cancer cases through the FISH assay. By introducing a novel algorithm that distinguishes HER2-low from HER2-normal cases, our study represents a significant advance in both clinical diagnostics and breast cancer research. We demonstrated that the **g3mclass** software automatically estimates an upper limit for the normal HER2 signal from the FISH data. Users can automatically deselect HER2-negative cases with a normal HER2 signal, which are unlikely to benefit from anti-HER2 therapies, which may be toxic and costly [25, 26]. Integration of this automated FISH-based method into the existing diagnostic workflow could be achieved by incorporating the image analysis and probabilistic classification software as an add-on module to the standard FISH evaluation. Pathologists could analyze FISH-stained slides using the established imaging platform, after which the algorithm would process the data and generate standardized reports indicating HER2 status, including HER2-low results. This would complement gold-

standard IHC and FISH practices without replacing them, enabling laboratories to adopt the system with minimal disruption and to provide clearer, more objective classification to guide therapy selection.

Accurate assessment of HER2 status is critical for therapy selection in advanced metastatic disease with low HER2 expression, and FISH HER2 copy number may serve as a better primary test for HER2-low breast cancer. It is essential to emphasize that the FISH ratio-based grouping led to the mixing of HER2-negative cases with non-classical FISH results, such as HER2 IHC-positive polysomy cases. This aligns with reports that non-classical amplification may lead to discordance with HER2 IHC results [27]. High HER2 copy number alone is considered important in the ASCO/CAP HER2 testing guidelines [17]. Increased HER2 copy number may raise the risk of death from breast cancer, even with a negative HER2 status by ratio [28]. Genomic analysis of a large population with metastatic breast cancer evaluated at Dana-Farber Cancer Institute showed higher HER2 copy count in HER2-low tumors (IHC 1 + or IHC 2/ISH-negative) than in HER2-0 tumors (IHC 0). HER2 copy count may be a promising biomarker for predicting response to T-DXd [29]. Our findings concur with the notion that HER2 gene copy number data may be a preferred selection parameter for HER2-low-expressing breast cancer. This approach also has potential for broader use, as the automated FISH-based algorithm could be adapted to assess low-level HER2 expression in other tumor types or pathologic settings. Such versatility may facilitate more precise patient stratification and support the development of targeted therapies in diverse clinical and research contexts.

Our current study is limited to secondary analysis of PT microarray samples, which lack clinicopathological features. We used peritumoral tissue as a reference, though ideally, breast tissue from women without breast cancer is preferred. We analyzed data from a regulatory-approved dual-probe HER2 FISH test with automated scanning available. In the future, the analysis can extend to other gold-standard methods, such as CISH or SISH, which are less time- and resource-intensive than FISH. Running the software-grouping trials on single-probe ISH data is warranted. Prospective clinical trials are also needed to assess how automated FISH-based stratification impacts treatment decisions and patient outcomes among those with HER2-low breast cancer.

In summary, the **g3mclass** software algorithm enables objective, accurate identification of HER2-low status using FDA-cleared methods, saving resources for new assay development. This streamlines analysis, sets data-driven quality cutoffs, and improves patient selection for anti-HER2 ADCs, advancing both clinical practice and research. This methodological advancement represents a significant step forward, addressing an unmet clinical need and providing actionable diagnostic tools for HER2-targeted therapies.

MATERIALS AND METHODS

Human Breast Tissue

The present study analyzes published datasets from 515 invasive breast cancer cores with HER2 IHC status and HER2 ISH results obtained by pathologists who participated in the College of American Pathologists proficiency surveys for HER2 expression in breast cancer (2014 HER2-B–2021 HER2-B) and the Taiwan Society of Pathology HER2 IHC Proficiency Test (2013–2021). A board-certified pathologist evaluated 52 peritumoral benign glands of 52 breast cancers as reference ISH values, as described in [14]. The raw data are from Supplemental Table 1, which includes HER2 ISH and IHC data for the study cohort, as published in [14]. Data are available in the online supplemental table 1 from J Clin Pathol 76 (11):753–756. doi:10.1136/jcp-2022-208444. The *Journal of Clinical Pathology* (JCP) is a subscription-based journal published by BMJ in partnership with the Association of Clinical Pathologists. Because free public access may be delayed, the full copy of the paper obtained through the institutional subscription is provided in the Supplementary Materials of this paper. Clinical-pathologic data were unavailable.

Ethical approval and consent to participate

Not applicable. The study is a research paper that utilizes published data and does not involve any additional ethical issues.

IHC and FISH assay data

The HER2 IHC status of breast cancer on tissue microarrays was assessed using the FDA-approved Ventana PATHWAY anti-HER2/neu (4B5) rabbit monoclonal antibody assay (Ventana Medical Systems, Tucson, Arizona, USA) on a BenchMark ULTRA system (Ventana) according to the manufacturer's instructions. All HER2 IHC staining results in this study were reviewed by a board-certified pathologist following the 2018 American Society for Clinical Oncology and the College of American Pathologists (ASCO/CAP) HER2 guidelines. Of the 515 cancer cores tested for HER2 IHC by PT, 145 were scored 3+, 11 were scored 2+, 27 were scored 1+, and 332 were scored 0. A dual-probe fluorescence ISH (FISH) test for detecting HER2 gene amplification was performed on tissue microarrays using a PathVysion HER2 DNA probe kit (Abbott Laboratories, Des Plaines, Illinois, USA) following the manufacturer's guidelines. The pathologists provided the HER2 FISH results for the tissue cores. Determination of HER2 status by FISH was done by assessing the mean number of HER2 copies or by calculating the mean HER2 to chromosome 17 centromere (CEP17) ratio in a population of tumor nuclei, where CEP17 represents the number of chromosomes 17. The HER2 amplification status, determined by FISH, was interpreted according to the manufacturer's guidelines for a dual-probe PathVision FISH kit and the 2018 ASCO/CAP clinical practice guideline focused update. A HER2:CEP17 ratio of ≥ 2.0 was used as the primary threshold for amplification, with average HER2 copy numbers of 4 and 6 serving as additional cutoffs [17]. "Non-classical" HER2 FISH results were defined according to the publication cited in reference [27].

Statistical analyses

We performed advanced software-assisted statistical analysis of FISH assay results, including the ratio of average HER2 to chromosome 17 centromere (CEP17) signals per nucleus (ratio) and the average HER2 copy number per nucleus (HER2 signal). We employed the estimation-maximization (EM) algorithm-based *g3mclass* software, specifically developed and validated for multi-class classification of clinical samples on biomarkers [18, 19]. **The g3mclass** software does not make assumptions about HER2 status; instead, it models a probability distribution based on the entire spectrum of available FISH assay results (either ratio or HER2 signal) in cancer (test) and benign (reference) samples. From the models, the software calculated data-driven cutoffs and assigned each sample to the group to which it was most likely to belong, based on the principle of maximum likelihood. In this article, dataset modeling and analysis were performed using **g3mclass** v.1.2 on macOS Mojave v.10.14.6, with Python v.3.9.5 and wxPython v.4.1.1. Additionally, we used GraphPad Prism 9.3.0 (GraphPad Software, LLC, La Jolla, CA, USA) to visualize group comparisons with dot plots. We used a single-factor ANOVA to compare the group means. A *P*-value of less than 0.01 was considered significant.

Declarations

Disclosure of potential conflict of interest

No potential financial and non-financial conflicts of interest were disclosed.

Funding

The authors received no external funding to support this research.

Author Contribution

Author's ContributionsConception and analysis: M.A.G. Manuscript writing, review, and revision: M.A.G.

Acknowledgement

Acknowledgments We thank Dr. Yen-Ying Chen, a board-certified pathologist from Taipei Medical University, Taipei, Taiwan, for sharing published data and engaging in valuable discussions. We thank Dr. Serguei Sokol for maintaining the pypi.org g3mclass site.

Data Availability

Data sharing does not apply to this article as no datasets were generated during the current study. The g3mclass is a standalone software. It is distributed under the GPL-2 Open-Source license and is freely accessible at https://pypi.org/project/g3mclass. It can be installed using the standard Python pip tool.

References

1. Cheng, X. A Comprehensive Review of HER2 in Cancer Biology and Therapeutics. *Genes (Basel)*. **15** (7). 10.3390/genes15070903 (2024).
2. Lee, J. & Park, Y. H. Trastuzumab deruxtecan for HER2 + advanced breast cancer. *Future Oncol.* **18** (1), 7–19. 10.2217/fon-2021-0550 (2022).
3. Bardia, A. et al. Trastuzumab Deruxtecan after Endocrine Therapy in Metastatic Breast Cancer. *N Engl. J. Med.* **391** (22), 2110–2122. 10.1056/NEJMoa2407086 (2024).
4. Dilawari, A. et al. US Food and Drug Administration Approval Summary: Trastuzumab Deruxtecan for the Treatment of Adult Patients With Hormone Receptor-Positive, Unresectable or Metastatic Human Epidermal Growth Factor Receptor 2-Low or Human Epidermal Growth Factor Receptor 2-Ultralow Breast Cancer. *J. Clin. Oncol.* **43** (26), 2942–2951. 10.1200/JCO-25-00812 (2025).
5. Franchina, M. et al. Low and Ultra-Low HER2 in Human Breast Cancer: An Effort to Define New Neoplastic Subtypes. *Int. J. Mol. Sci.* **24** (16). 10.3390/ijms241612795 (2023).
6. Chen, Z. et al. Is HER2 ultra-low breast cancer different from HER2 null or HER2 low breast cancer? A study of 1363 patients. *Breast Cancer Res. Treat.* **202** (2), 313–323. 10.1007/s10549-023-07079-8 (2023).
7. Camp, R. L., Dolled-Filhart, M., King, B. L. & Rimm, D. L. Quantitative analysis of breast cancer tissue microarrays shows that both high and normal levels of HER2 expression are associated with poor outcome. *Cancer Res.* **63** (7), 1445–1448 (2003).
8. Allison, K. H. & Wolff, A. C. ERBB2-Low Breast Cancer-Is It a Fact or Fiction, and Do We Have the Right Assay? *JAMA Oncol.* **8** (4), 610–611. 10.1001/jamaoncol.2021.7082 (2022).
9. Fernandez, A. I. et al. Examination of Low ERBB2 Protein Expression in Breast Cancer Tissue. *JAMA Oncol.* **8** (4), 1–4. 10.1001/jamaoncol.2021.7239 (2022).
10. Zaakouk, M. et al. Concordance of HER2-low scoring in breast carcinoma among expert pathologists in the United Kingdom and the republic of Ireland -on behalf of the UK national coordinating committee for breast pathology. *Breast* **70**, 82–91. 10.1016/j.breast.2023.06.005 (2023).
11. Lv, H. et al. Prevalence and concordance of HER2-low and HER2-ultralow status between historical and rescored results in a multicentre study of breast cancer patients in China. *Breast Cancer Res.* **27** (1), 45. 10.1186/s13058-025-02001-0 (2025).
12. Wrobel, A. et al. Accuracy of human epidermal growth factor receptor 2 (HER2) immunohistochemistry scoring by pathologists in breast cancer, including the HER2-low cutoff: HER2 IHC scoring concordance in breast cancer. *Diagn. Pathol.* **20** (1), 35. 10.1186/s13000-025-01624-3 (2025).
13. Wu, S. et al. Interobserver consistency and diagnostic challenges in HER2-ultralow breast cancer: a multicenter study. *ESMO Open.* **10** (2), 104127. 10.1016/j.esmoop.2024.104127 (2025).

14. Chen, Y. Y., Yang, C. F. & Hsu, C. Y. Proposal to distinguish HER2-low from HER2-normal breast cancer by in situ hybridisation. *J. Clin. Pathol.* **76** (11), 753–756. 10.1136/jcp-2022-208444 (2023).
15. Sato, A., Maeda, Y., Matsumoto, A., Ikeda, T. & Jinno, H. Clinicopathological features and prognosis of breast cancer with low HER2 status evaluated by fluorescence in situ hybridization. *Oncol. Lett.* **29** (2), 92. 10.3892/ol.2024.14838 (2025).
16. Mascarello, J. T. et al. Section E9 of the American College of Medical Genetics technical standards and guidelines: fluorescence in situ hybridization. *Genet. Med.* **13** (7), 667–675. 10.1097/GIM.0b013e3182227295 (2011).
17. Wolff, A. C. 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. *J. Clin. Oncol.* **36** (20), 2105–2122. 10.1200/JCO.2018.77.8738 (2018).
18. Guvakova, M. A. & Sokol, S. The g3mclass is a practical software for multiclass classification on biomarkers. *Sci. Rep.* **12** (1), 18742. 10.1038/s41598-022-23438-9 (2022).
19. Guvakova, M. A. Automated Classification of Breast Cancer Across the Spectrum of ERBB2 Expression Focusing on Heterogeneous Tumors With Low Human Epidermal Growth Factor Receptor 2 Expression. *JCO Clin. Cancer Inf.* **7**, e2300013. 10.1200/CCI.23.00013 (2023).
20. Seo, J. et al. HER2 amplification level by in situ hybridization predicts survival outcome in advanced HER2-positive breast cancer treated with pertuzumab, trastuzumab, and docetaxel regardless of HER2 IHC results. *Breast Cancer Res.* **25** (1), 154. 10.1186/s13058-023-01746-w (2023).
21. Wolff, A. C. et al. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: ASCO-College of American Pathologists Guideline Update. *J. Clin. Oncol.* **41** (22), 3867–3872. 10.1200/JCO.22.02864 (2023).
22. Palm, C. et al. Determining HER2 Status by Artificial Intelligence: An Investigation of Primary, Metastatic, and HER2 Low Breast Tumors. *Diagnostics (Basel)*. **13** (1). 10.3390/diagnostics13010168 (2023).
23. Du, J. et al. Machine learning prediction of HER2-low expression in breast cancers based on hematoxylin-eosin-stained slides. *Breast Cancer Res.* **27** (1), 57. 10.1186/s13058-025-01998-8 (2025).
24. Ruschoff, J. et al. Global Study on the Accuracy of Human Epidermal Growth Factor Receptor 2-Low Diagnosis in Breast Cancer. *Arch. Pathol. Lab. Med.* **149** (5), 431–438. 10.5858/arpa.2024-0052-OA (2025).
25. Dickerson, J. C. et al. Cost and Cost-Effectiveness of Treating Human Epidermal Growth Factor Receptor 2-Low Metastatic Breast Cancer. *J. Clin. Oncol.* **43** (19), 2208–2217. 10.1200/JCO-24-01960 (2025).
26. Han, J. et al. Pharmacovigilance study of adverse reactions of anti-HER-2 drugs for the treatment of HER-2-positive breast cancer based on the FAERS database. *Breast Cancer Res.* **27** (1), 54. 10.1186/s13058-025-02013-w (2025).

27. Ballard, M. et al. Non-classical' HER2 FISH results in breast cancer: a multi-institutional study. *Mod. Pathol.* **30** (2), 227–235. 10.1038/modpathol.2016.175 (2017).
28. Biserni, G. B., Engstrom, M. J. & Bofin, A. M. HER2 gene copy number and breast cancer-specific survival. *Histopathology* **69** (5), 871–879. 10.1111/his.13010 (2016).
29. Tarantino, P. et al. Comprehensive genomic characterization of HER2-low and HER2-0 breast cancer. *Nat. Commun.* **14** (1), 7496. 10.1038/s41467-023-43324-w (2023).

Figures

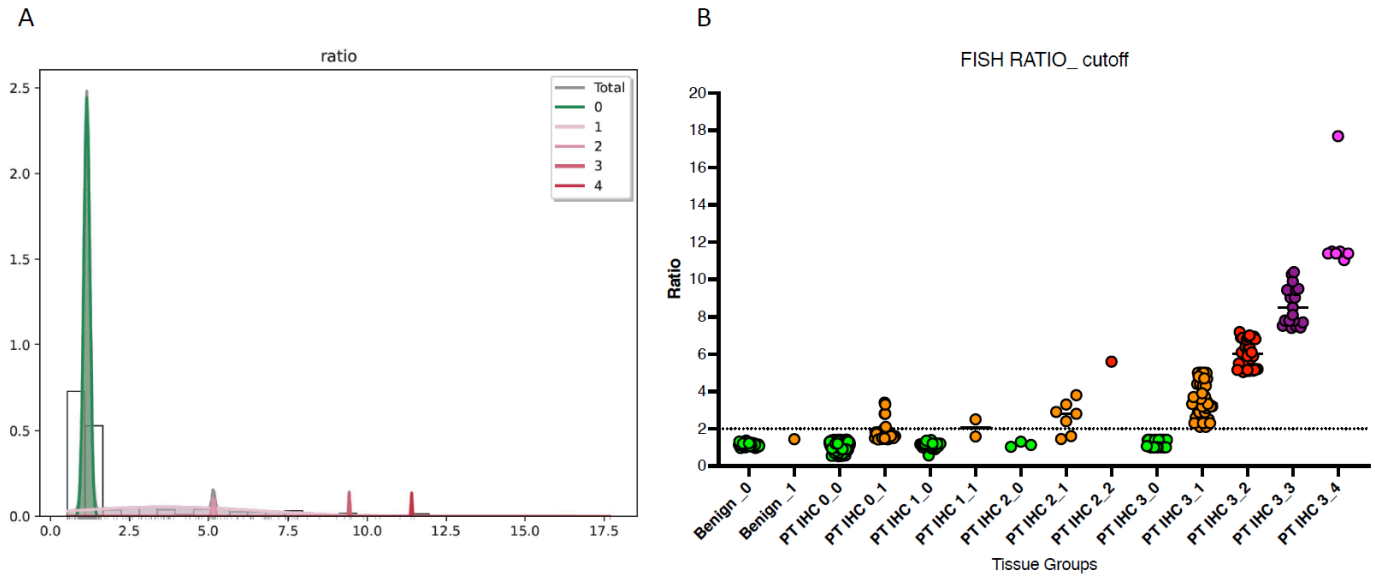


Figure 1

A. The *g3mclass*-generated Gaussian Mixture **models for the average HER2:CEP17 data set**. The *g3mclass* estimated cutoff is based on modeling the HER2/CEP 17 ratio results from the test (cancer) and reference (benign) samples. The probability density function for the Gaussian Mixture Model (gray, total) and individual group distributions (colored) are overlaid on a histogram. Y axis, probability density; X axis, **the ratio of average HER2:CEP17 signals per cell. B. Dot plot,** ratios in IHC scored tissue groups. The scatter plot depicts the predicted cancer classification based on the model output. **Green dots represent group 0**: Benign_0, benign tissue sample; PT IHC 0_0, tumor sample IHC 0, PT IHC 1_0, tumor sample IHC 1+, PT IHC 2_0, tumor sample IHC 2+; PT IHC 3_0, tumor sample IHC 3+. **Orange dots represent group 1:** Benign_1, benign tissue sample; PT IHC 0_1, tumor sample IHC 0; PT IHC 1_1, tumor sample IHC 1+; PT IHC 2_1, tumor sample IHC 2+. **Red dots represent group 2:** PT IHC 2_2, tumor sample IHC 2+. PT IHC 3_2, tumor sample IHC 3+. **Purple dots represent group 3**: PT IHC 3_3, tumor sample IHC 3+. **Magenta dots represent group 4**: PT IHC 3_4, tumor sample IHC 3+. Dotted horizontal line, the ratio of average HER2:CEP17 signals per cell = 2.

A B

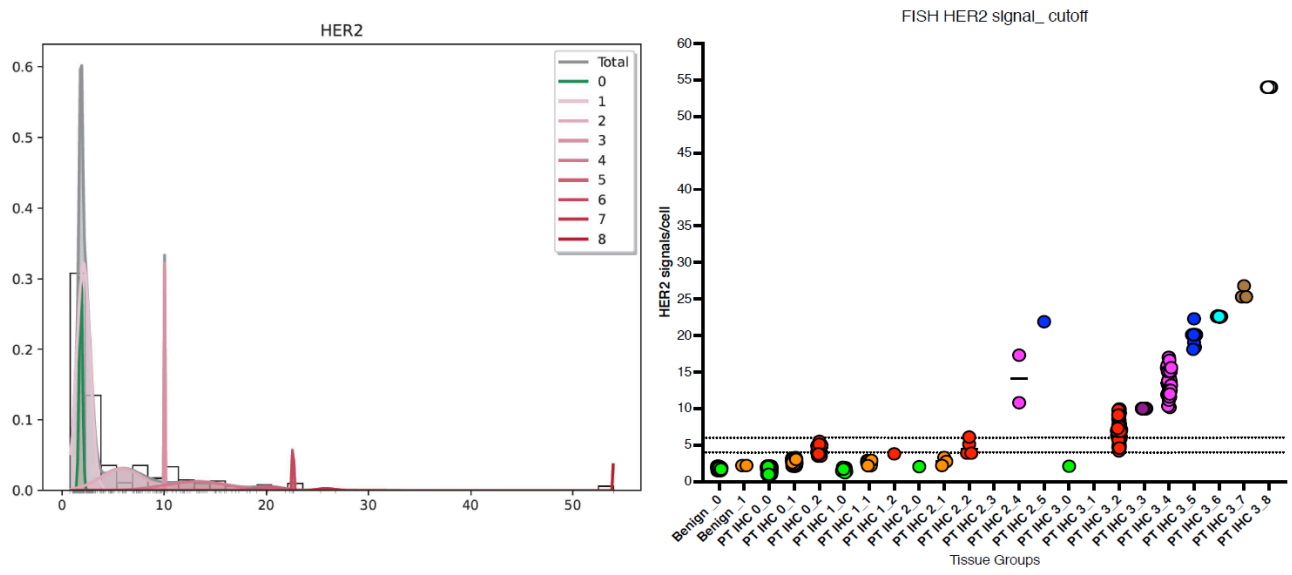


Figure 2

A. The *g3mclass*-generated Gaussian Mixture **models for the average HER2 signal**. The probability density functions for the mixture model (gray, total) and individual class distributions (colored) are overlaid on a histogram. Y axis: probability density; X axis: **the average HER2 copy numbers. B**. **Dot plot,** HER2 signal in IHC-scored tissue groups. **Green dots represent group 0**: Benign_0, benign tissue sample; PT IHC 0_0, tumor sample IHC 0, PT IHC 1_0, tumor sample IHC 1+, PT IHC 2_0, tumor sample IHC 2+; PT IHC 3_0, tumor sample IHC 3+. **Orange dots represent group 1:** Benign_1, benign tissue sample; PT IHC 0_1, tumor sample IHC 0; PT IHC 1_1, tumor sample IHC 1+; PT IHC 2_1, tumor sample IHC 2+. **Red dots represent group 2:** PT IHC 0_2, tumor sample IHC 0; PT IHC 1_2, tumor sample IHC 1+; PT IHC 2_2, tumor sample IHC 2+. PT IHC 3_2, tumor sample IHC 3+.

Purple dots represent group 3: PT IHC 3_3, tumor sample IHC 3+. **Magenta dots represent group 4**: PT IHC 2_4, tumor sample IHC 2+, PT IHC 3_4, tumor sample IHC 3+. **Blue dots represent group 5**: PT IHC 2_5, tumor sample IHC 2+; PT IHC 3_5, tumor sample IHC 3+.

Turquoise dots represent group 6: PT IHC 3_6, tumor sample IHC 3+. **Brown dots represent group 7**: PT IHC 3_7, tumor sample IHC 3+. **White dots represent group 8**: PT IHC 3_8, tumor sample IHC 3+. Dotted horizontal lines indicate the average HER2 copy numbers of 4 and 6.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GUVAKOVASUPPLEMENTARYMATERIALS.pdf