STUDY REPORT SYNOPSIS

HER2 Retrospective Epidemiology Study

Prevalence and Clinicopathologic feature of different HER2 level in Chinese breastcancer patients (HER2 PATH)

Milestones:	Proposal Approval	20 Oct 2020
	Final Clinical Study protocol	18 Mar 2021
	Study Start Up: contracts in place, regulatory submissions, initiation visits	18 Feb 2022
	First slice assess	22 Feb 2022
	Last slice assess	5 Jul 2023
	Database Lock	25 Jul 2023
	Final Clinical Study Report	16 Nov 2023
Phase of development:	<<Observational Study >>	
Sponsor:	AstraZeneca	
Author:	[redacted]	

This study was performed in compliance with Good Clinical Practice (GCP) and Good Pharmacoepidemiology Practice (GPP), including the archiving of essential documents.

This submission/document contains trade secrets and confidential commercial information, disclosure of which is prohibited without providing advance notice to AstraZeneca (AZ) and opportunity to object.

Background/rationale:

Breast cancer is the most common malignancy affecting women worldwide, with approximately 1.4 million new cases diagnosed each year. The overexpression of the HER2 gene, found in around 15-20% of metastatic breast cancer (mBC) patients, is associated with a higher risk of recurrence and poorer prognosis. Traditionally, the determination of HER2 status has been a binary classification, dividing patients into HER2-positive and HER2-negative categories, which plays a crucial role in guiding therapeutic strategies and is widely adopted in clinical practice.

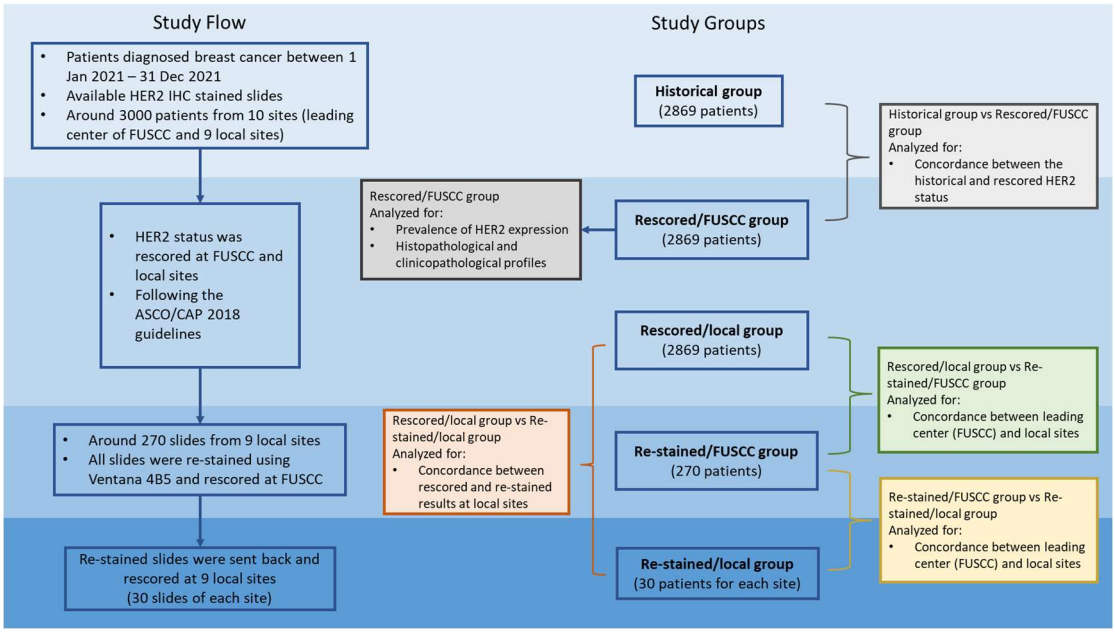
The findings from the DESTINY-Breast04 trial have brought forth a new perspective on the treatment landscape of mBC with low levels of HER2 expression. The trial demonstrated that this specific subset of patients can benefit from a novel anti-HER2 antibody-drug conjugate (ADC) called Trastuzumab Deruxtecan. This discovery has led to the proposal of a three-tiered classification system for HER2 expression in breast cancer: HER2-positive (IHC 3+ or IHC 2+/ISH+), HER2-low (IHC 1+ or IHC 2+/ISH-), and HER2 IHC 0, which deviates from the traditional binary categorization. The lower threshold for HER2 expression that can benefit from HER2-directed antibody-drug conjugates (ADCs) is still being investigated, such as HER2 immunohistochemistry (IHC) >0 to <1+ (defined as IHC 0 with incomplete and faint staining in ≤10% of tumor cells) in the DESTINY-Breast06 trial.

Estimates suggest that up to 55% of breast cancer patients may exhibit HER2-low expression without gene amplification. Limited previous studies conducted in China have provided data on the prevalence of HER2-low expression. Most of the available information is based on historical results, which primarily identified the HER2-positive populations (IHC3+ or IHC2+/ISH+). As a result, high quality data on the prevalence of HER2-low expression and IHC >0 to <1+ in breast cancer patients remains largely unexplored in Chinese cohorts.

With HER2-low breast cancer patients emerged as a new targetable population, ensuring accurate and consistent diagnosis of HER2-low status is paramount for guiding HER2-targeted treatment decisions. While IHC and FISH will continue to be the main methods for selecting patients with HER2-low status, it is important to acknowledge that the historical scoring methods have predominantly concentrated on identifying HER2-positive populations. This emphasis on HER2-positive cases has raised concerns regarding the accuracy of assessing HER2-low expression. Furthermore, there is a significant variation in the interpretation of HER2 IHC results among clinicians, leading to inconsistent diagnoses. Recent research by Fernandez et al. has highlighted the poor scoring accuracy for HER2 IHC in the low range (0 and 1+) based on the CAP survey data set. Their study revealed only a 26.0% concordance between HER2 0 and 1+ compared to a 58.0% concordance between 2+ and 3+ across 18 pathologists and 170 cases. These findings underscore the urgent need for improved standardization and consistency in HER2 IHC interpretation.

On another note, the integration of artificial intelligence (AI) in medical imaging and biomarker interpretation is gaining momentum. AI, especially deep learning, has shown promise in enhancing the accuracy and consistency of HER2 status determination. Several studies have illustrated the effectiveness of AI in interpreting biomarkers, suggesting its potential to standardize practices and ultimately improve treatment outcomes for breast cancer patients.

This retrospective study aimed to estimate the prevalence and describe the clinical manifestations of different HER2 level BC by accurate reassessing of HER2 expression in archived HER2 IHC slides, and analyzing its clinicopathologic feature and outcomes, and explore the concordance between manual and AI assisted HER2 interpretation. The study was divided into two parts. PART1 aimed to include 200 subjects diagnosed at Fudan University Shanghai Cancer Center (FUSCC) between January 2015 and December 2015. This part planned to involve re-staining the archived slides and re-assessing HER2 expression levels. PART2 aimed to include 3000 patients from 10 study sites across China, diagnosed between July 2021 and July 2022. We also planned to re-stain the archived slides and re-assess HER2 expression level in this part.



Methods:

Study design:

This multicenter retrospective study aimed to estimate the prevalence of different HER2 expression levels (HER2+, HER2-low, HER2 IHC 0) in approximately 200 breast cancer patients at Fudan University Shanghai Cancer Center (FUSCC) from January 2015 to December 2015 (PART1) and 3000 breast cancer patients from 10 medical centers in China between July 2021 and July 2022 (PART2).

The first part of the study (PART1) involved collecting historical HER2 expression level results and re-staining and rescoring HER2 IHC slides from 200 subjects who were pathologically diagnosed with breast cancer at FUSCC during the specified time period. However, due to limited slide quality and missing data in PART1, it was not included in the final analysis. Therefore, all the results presented in the study were derived solely from the data obtained in PART2.

Patient Selection and Disposition:

The PART2 included individuals aged 18 years or older with a confirmed diagnosis of breast cancer between July 2021 and July 2022. Eligible patients needed to have at least one archived HER2 IHC slide in good condition for rescoring and available fluorescence in situ hybridization (FISH) results for HER2 IHC2+. Detailed exclusion criteria were specified in the study.

Patients from 10 medical centers in China who underwent breast cancer surgery between July 2021 and July 2022 were included in the study. A total of 300 patients per site were collected chronologically, with Fudan University Shanghai Cancer Center (FUSCC) being the leading study center. Relevant information, including patients' general details, demographic data, diagnosis, clinicopathological features, and historical IHC scores, were extracted from medical records. Archived HER2 IHC slides from these patients were subjected to rescoring by a review committee, blind to the historical results.

Re-staining and rescoring at FUSCC:

A total of 270 patient samples from 9 local sites, with 30 samples from each local site, were carefully chosen for HER2 IHC re-staining. Within each site, a random selection process was used to choose 8 patients with a HER2 IHC score of 0, 9 patients with a score of 1+, 9 patients with a score of 2+, and 4 patients with a score of 3+. These selected samples were then re-sectioned and sent to Fudan University Shanghai Cancer Center (FUSCC) for re-staining and rescoring. Subsequently, the re-stained slides were returned to their respective centers for scoring by the local review committee.

AI-assisted HER2 assessment

In Part 2 of the study, out of the 3,000 patients, approximately 800 HER2 IHC samples were specifically identified for validation using an AI-assisted HER2 assessment system. To ensure a representative sample, approximately 80 slides from each site were chosen randomly, maintaining a proportional distribution of historical IHC scores (0, 1+, and 2+) in a ratio of 4:4:2.

These selected 80 slides from each site were then subjected to re-scoring using the AI-assisted HER2 assessment system after a wash period of 2-4 weeks. It is important to note that the same pathologists who performed the initial assessment were also involved in the re-assessment and AI-assisted re-scoring processes. This approach was adopted to validate the clinical utility of the HER2 AI system. The results from each individual observer were recorded and analyzed as part of the study.

Review Committee

The pathologists committee consisted of 3 professional pathologists in each site. The reading mode was "2 readers + 1 adjudicator". Each of the two readers would complete the evaluation of the same slide independently. If results matched between the two reader, then it would be recorded as the final result. Otherwise the adjudicator would make the final judgement where the more “accurate” result would be selected.

All slides were stained using Ventana 4B5 and scored following the ASCO/CAP 2018 guidelines, including the addition of the IHC >0<1+ cut-off as defined in the DESTINY-Breast 06 trial.

Outcomes:

Primary outcomes: The Prevalence of different HER2 expression levels in PART2

Secondary outcomes:

- Histopathological and clinicopathological characteristics by different HER2 expression level (HER2 IHC 0/Low/Positive) in PART2 including: tumor size, positive lymph node number, TNM stage, pathological type, histological grade, multi-focal, ER status, ERpercentage, PR status, PR percentage, Ki-67 percentage, HER2 IHC score, FISH status.
- HER2 IHC score concordance between historical scoring and reassess scoring in PART2:

- The kappa coefficient on overall agreement between historical and reassessmet scoring;
- The detailed distribution of historical IHC scoring within each reassessed IHC score category in terms of: percentage of each possible shifting circumstance (i.e., percentage of patients with historical score as 0/1+/2+/3+ in reassessed score as 0/1+/2+/3+).

- HER2 IHC score concordance between leading site (FUSCC) and other centers in PART2: percentage of each possible shifting circumstance between the scores (0/1+/2+/3+) from the leading site and other sites in China.
- HER2 IHC score concordance between manual and AI assisted HER2 interpretation in PART2:

- The kappa coefficient on overall agreement between manual and AI assisted HER2 interpretation.

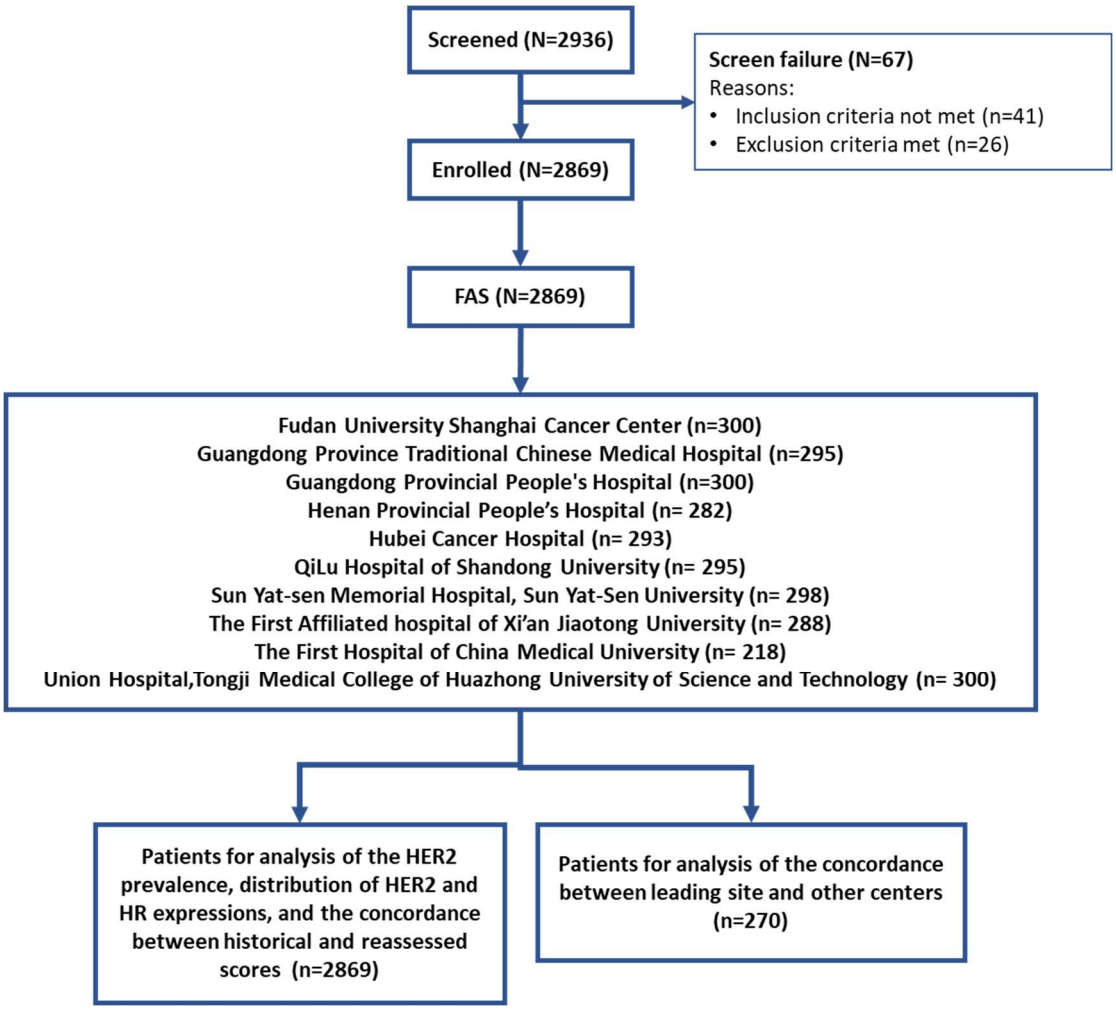
Exploratory outcomes:

- Prevalence of patients with IHC staining score as HER2 IHC 0 and HER2 IHC >0<1+ in PART2.

Results:

In the study, PART1 was not subjected to analysis due to significant factors such as the poor quality of the tissue sections and the severe lack of data. These issues rendered the PART1 segment unsuitable for reliable evaluation, so we did not report the PART1 results.

A total of 2936 patients were screened, among which 2869 were included in this study (Figure 2).



Distribution of HER2 IHC scores and expression levels based on rescores

Based on the rescores conducted by the pathologist committee, a total of 682 patients (23.8%) were classified as IHC 0, 871 patients (30.4%) as 1+, 801 patients (27.9%) as 2+, and 515 patients (18.0%) as 3+. When the FISH results were combined, the prevalence of HER2-low was determined to be 54.5% (95% confidence interval: 52.7%, 56.3%). Additionally, among all patients, the prevalence of HER2 IHC >0 to <1+ and HER2 null was found to be 10.6% (N=379) and 13.2% (N=303), respectively.

Items	Total patients (N=2869)
HER2 IHC reassessed scores	
0	682 (23.8%)
IHC null	379 (13.2%)
IHC >0<1+	303 (10.6%)
1+	871 (30.4%)
2+	801 (27.9%)
3+	515 (18.0%)
Total	2869
FISH results for HER2 IHC reassessed scores with 2+	
FISH-	692 (86.6%)
FISH+	107 (13.4%)
Unavailable	2
Total	801
HER2 expression level based on the reassessed HER2 status	
HER2 IHC 0	682 (23.8%)
HER2-low	1563 (54.5%)
HER2-positive	622 (21.7%)
Unavailable	2
Total	2869

Concordance between historical results and rescores

The concordance rate between historical results and rescores for HER2 IHC score was found to be 83.1% (2383/2869). Overall, there was substantial agreement between the two groups regarding IHC scores, with a kappa value of 0.77 (95% CI, 0.75-0.79). The HER2 IHC 1+ group had a lower concordance rate of 74.5% compared to the IHC 0 (85.2%), IHC 2+ (81.3%), and IHC 3+ (98.6%) groups. However, it is important to mention that 12.1% (n=106) of the cases initially categorized as IHC 1+ were rescored as IHC 2+, but this had limited impact on the diagnosis of HER2-low.

Regarding the expression levels of HER2, the overall concordance rate between historical results and rescores was 91.7% (2632/2869). The concordance rates for HER2 IHC 0, low, and positive were 85.2%, 91.7%, and 99.2% respectively. The kappa value of 0.86 (95% CI, 0.85-0.88) indicated an almost perfect agreement between the two groups in terms of HER2 expression levels.

Interobserver variation based on FUSCC re-stained slides

There was an almost perfect agreement observed between the results assessed by the review committee from FUSCC and the local sites, both in terms of IHC scores (concordance rate of 82.2% and kappa value of 0.77) and HER2 expression levels (concordance rate of 91.5% and kappa value of 0.88). It is noteworthy that all nine local sites demonstrated either perfect or substantial agreement with FUSCC, as indicated by the Simple Kappa coefficient values ranging from 0.66 to 0.86.

Concordance on AI results and manual rescores

The overall concordance rate was 82.6% (659/798). Of the 316 patients scored as HER2 IHC 0 by human, AI confirmed this for 249 (79.6%), while 63 (20.1%) were re-assessed as HER2-low, and 1 (0.3%) as HER2-positive. For the 455 patients classified as HER2-low by pathologists, 391 (86.9%) were confirmed by AI, 52 (11.6%) were re-assessed as HER2 IHC 0, and 7 (1.6%) as HER2-positive.

Conclusion:

This study has provided an insightful analysis of the distribution across different HER2 groups in breast cancer patients. We found that 23.8% (95% CI: 22.2%, 25.3%) of the subjects were HER2 IHC 0, 54.5% (95% CI: 52.7%, 56.3%) were categorized as HER2-low, and 21.7% (95% CI: 20.2%, 23.2%) were HER2 positive. Notably, among patients which could be previously diagnosed as HER2-negative (HER IHC 0 and HER2-low), 69.6% were found to be HER2-low. This underscores the significant value of shifting from a binary HER2 positive/negative classification to a more nuanced tripartite categorization. Such refined diagnostic classification allows clinicians to select more appropriate treatment strategies.

Our consistency analysis indicates that the concordance between historical interpretations and re-assessment of the slides was generally good. However, when re-reading slides based on re-staining, both performed by the leading centre and the other centres, the concordance for IHC 2+ was relatively low. Additionally, the overall consistency was not satisfactory when the other centres re-assess the slides based on both historical and re-staining. These findings highlight the importance of standardized staining procedures and interpretation criteria in the future HER2 status diagnosis .

AI-assisted HER2 assessment demonstrated high accuracy in our study. As we move towards the widespread adoption of precise diagnostics and treatment strategies for HER2, standardized operational procedures and assessing criteria training become imperative. Furthermore, the role of AI-assisted interpretation is likely to be increasingly significant, offering valuable support in achieving more accurate and consistent diagnostic outcomes.