scientific reports

Check for updates

OPEN

Development and validation of a postoperative prognostic model for hormone receptor positive early stage breast cancer recurrence

Ruixin Pan[1,7], Haoting Shi[1,7], Yiqing Shen[2,7], Xue Wang[3], Shi Zhao[4], Nan Zhang[4], Xueyan Zhang[5], Shuwen Dong[1], Chao Hu[1], Jiayi Wu[1], Weimin Chai[6], Xiaosong Chen[1]✉ & Kunwei Shen[1]✉

Predicting recurrence among early-stage hormone receptor-positive human epidermal growth factor receptor-negative breast cancer (HR+/HER2− BC) is crucial for guiding adjuvant therapy. However, studies are limited for patients with low recurrence risk. HR+/HER2− early-stage (T1-2N0-1) invasive BC patients who received definitive surgery and followed by endocrine therapy from four independent medical centers were included in this retrospective study. Patients from center 1 were used as derivation cohort, while those from other centers were combined as an external test cohort. A deep learning prognostic model, HERPAI, was developed based on Transformer to predict risk of invasive disease-free survival (iDFS) utilizing clinical and pathological predictors. The model performance was evaluated using C-index for the overall population and subgroups. Threshold for selecting 5-year recurrence risk > 10% was determined. Hazard ratio (HR) was estimated between risk groups for iDFS. A total of 6340 patients were included, of whom 5424 were assigned to the derivation cohort (training and validation [N = 4882] and internal test cohort [N = 542]), while 916 patients were utilized as external cohort. HERPAI yielded a C-index of 0.73 (95% CI 0.65–0.81), 0.73 (95% CI 0.62–0.85), and 0.68 (95% CI 0.60–0.77), in the validation, internal, and external test cohort, respectively. Consistent performances were observed for pre-specified subgroups. High-risk patients were associated with an increased risk of recurrence for validation (HR, 2.56 [95% CI 1.25–5.22], *P* = 0.01), internal test (HR, 2.52 [95% CI 0.97–6.57], *P* = 0.06) and external test (HR, 1.94 [95% CI 1.00–3.74], *P* = 0.049) cohort, respectively. HERPAI was a promising tool for selecting vulnerable early-stage HR+/HER2− BC patients who were at high-risk of recurrence. It could facilitate the prioritization of patients who may benefit more from escalating adjuvant treatment.

Keywords Hormone receptor-positive, Breast cancer, Artificial intelligence, Prognostic model, Endocrine therapy

Abbreviations

BC	Breast cancer
HR	Hormone receptor
HER2	Human epidermal growth factor receptor
ET	Endocrine therapy
iDFS	Invasive disease-free survival
RS	Recurrence score
OS	Overall survival
STEEP	Standardized definitions for efficacy end points

[1]Department of General Surgery, Comprehensive Breast Health Center, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 20025, China. [2]Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA. [3]Department of Pathology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. [4]School of Public Health, Tianjin Medical University, Tianjin, China. [5]Changchun Institute of Biological Products, Changchun, China. [6]Department of Radiology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. [7]Ruixin Pan, Haoting Shi and Yiqing Shen contributed equally to this work. ✉email: chenxiaosong0156@hotmail.com; kwshen@medmail.com.cn

DDFS	Distant disease-free survival
ER	Estrogen receptor
PR	Progesterone receptor
SD	Standardized deviation
IQR	Interquartile
MICE	Multiple imputation with chained equations
BMI	Body mass index
CV	Cross validation
AUC	Area under the curve
DCA	Decision curve analysis
HR	Hazard ratio
CDK4/6	Cyclin-dependent kinases 4 and 6

Breast cancer (BC) is one of the most prevalent malignancies among women worldwide[1], of which hormone receptor-positive and human epidermal growth factor receptor-negative (HR+/HER2−) subtype accounts for the majority[2]. Although the prognosis of stage I-II HR+/HER2− BC was improved by adjuvant endocrine therapy, the 20-year distant recurrence rate was still greater than 10%[3]. Escalating endocrine treatment (ET) with CDK4/6 inhibitors has demonstrated improved invasive disease-free survival (iDFS) for stage II-III HR+/HER2− BC[4–6], though with conflict results[7]. These trials excluded stage I patients partly due to low recurrence risk, however, unmet treatment needs persist among them with a substantial recurrence rate after long-term follow-up[3]. Thus, there is a burning need to identify HR+/HER2− BC patients for escalated adjuvant treatment.

Several prognostic and predictive models were proposed and validated for HR+/HER2− patients[8–13]. However, these models were mostly developed based on all stage I–III BC, which may not be applicable for patients with earlier disease stages. Genetic-based clinical decision tools, such as Recurrence Score (RS) from Oncotype Dx[8] and MammaPrint[9], demonstrated less predictive performance among certain populations, including the postmenopausal women and the elderly, compared with their counterparts. Additionally, these models utilized distant recurrence and overall survival (OS) as their primary outcome[11,12], which is less encountered in patients with earlier stage and may impose the clinical application in the adjuvant setting. Overall, a significant gap persists for developing a robust prognostic model which focuses on early-stage HR+/HER2− BC to predict the comprehensive profile of disease recurrence, and to further guide adjuvant treatment.

Here, we sought to develop and validate a prognostic model to predict iDFS among early-stage (T1-2N0-1) HR+/HER2− BC with conventional clinical and pathological data.

Methods

Study design and participants

In this retrospective cohort study, female HR-positive BC cancer patients who received definitive surgery and were pathologically diagnosed as stage T1-2N0-1M0 per AJCC 8th guideline[14], from January 1, 2012, to December 31, 2023, were screened in four independent medical centers. The overall study design was shown in Fig. 1. The inclusion and exclusion workflow were demonstrated in Supplementary Fig. 1. Patients with HER2 positive BC, multiple lesions with unavailable IHC results, and HR negative or HER2 positive subtype, were excluded. Also, those who received any anti-tumor therapy before surgery, with concurrent other malignancies (defined as time from diagnosis less than 5 years), and those who died within 30 days after surgery were excluded.

This study was approved by the Ethical Review Boards of Ruijin Hospital (No. 2023126). Informed consent was obtained from all participants in this study. The study was reported consistent with the TRIPOD (transparent reporting of a multivariable prediction model for individual prognosis or diagnosis) guidelines[15] and was performed in accordance with the Declaration of Helsinki.

Outcome definitions

The primary outcome for the prognostic model was invasive disease-free survival (iDFS) per the Standardized Definitions for Efficacy End Points (STEEP) system[16], defined as the date from surgery to the first occurrence of one of the following events: invasive local recurrence, regional recurrence, distant metastasis, contralateral breast cancer, second primary invasive cancer, or death from any cause. Recurrence and metastases were centrally ascertained via radiological method or biopsy and determined by a radiologist with more than 20 years' experience (W.C) or a pathologist with more than 10 years' experience (X.W) who majored in BC diagnosis. (Supplementary Methods).

Distant disease-free survival (DDFS) and overall survival (OS) were the secondary outcomes. DDFS was defined as the date from surgery to the date of distant metastasis of breast cancer at any other site of the body (except for the contralateral breast). OS was defined as the date from surgery to the date of death of any cause. Patients were followed from surgery completed, until death, loss to follow-up, or study date cutoff (31st March 2024), whichever came first.

Candidate predictors

Candidate predictors were screened and selected based on clinical practice and previous literature. The candidate clinical and pathological predictors were detailed in the Supplementary Methods. Clinical variables included age, body mass index (BMI), menopause status, family history, breast surgery modality, and lymph node surgery modality. Pathological variables consisted of histological type, histological grade, estrogen receptor (ER) expression, progesterone receptor (PR) expression, HER2 expression, Ki67 expression, pathological T stage and N stage. (Supplementary Methods).

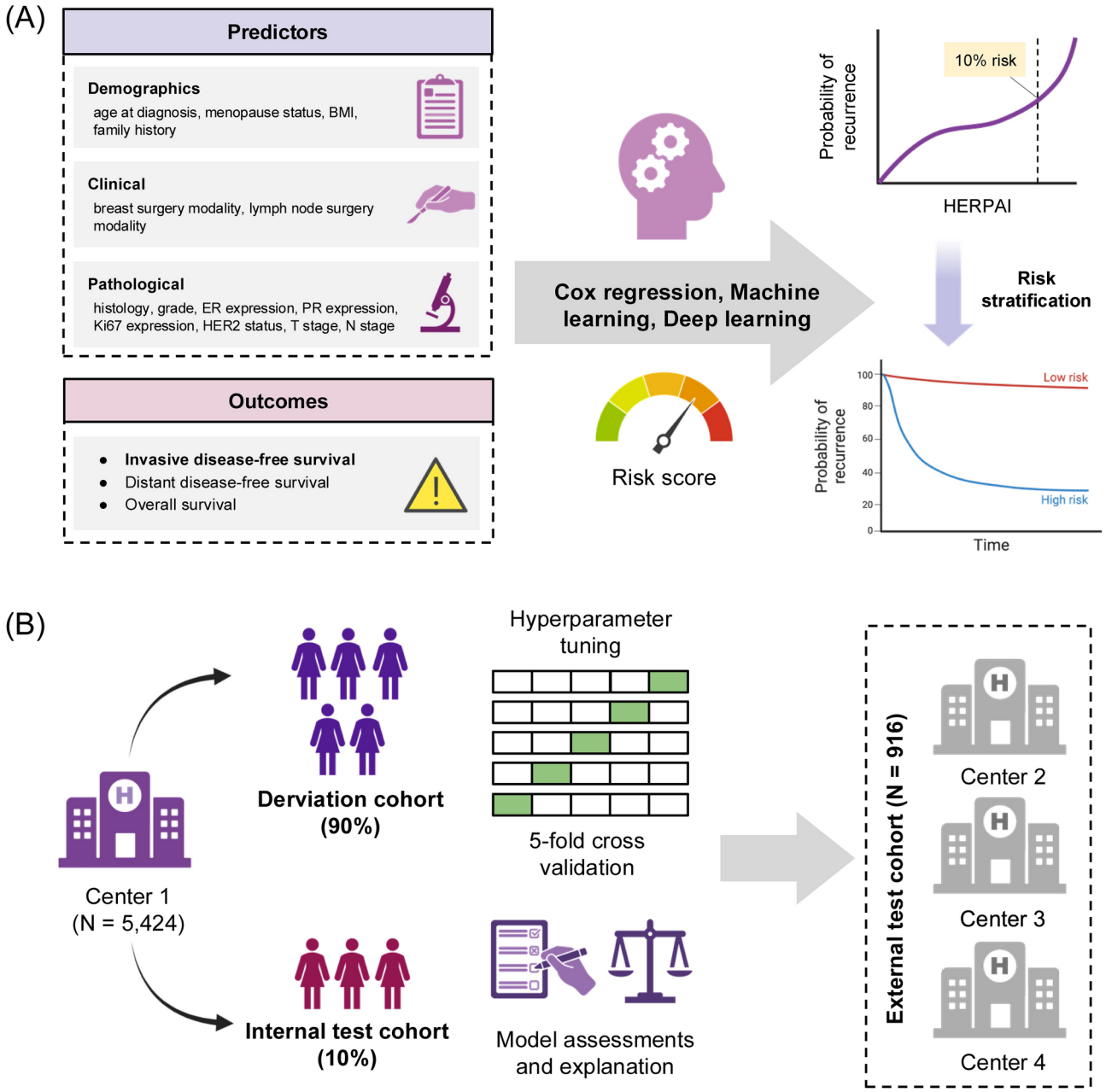

Fig. 1. Overall design of this study. (**A**) Predictors, outcomes, and methodology used in this study. (**B**) Study workflow.

Statistical analysis

Continuous variables were summarized using mean (standardized deviation [SD]) and median (interquartile [IQR]), as appropriate. Categorical variables were presented as count and percentage. Patients from center 1 were used as a derivation cohort. The derivation cohort was randomly divided into a training and validation cohort (90%) and an internal test cohort (10%). The external test cohort consisted of patients from other three independent medical centers.

Multiple imputation with chained equations (MICE) was employed to address missing BMI, tumor size, and grade data, under the assumption of missing at random[17]. Imputation was performed repeatedly to create 20 datasets for deviation, internal test, and external test cohort, respectively. All estimates were combined using Rubin's rule[18] (Supplementary Methods).

The deep learning model, HERPAI, was developed in the training and validation cohort using a fivefold cross validation (CV) strategy and was evaluated in the internal test and external test cohort (Supplementary Fig. 2). HERPAI was constructed based upon the FT-Transformer model[19]. Each Transformer layer included a multi-head self-attention mechanism and incorporated a position-wise feed-forward network applied to the output of the self-attention mechanism. The final outputs from the Transformer layers were passed through

fully connected layers to predict iDFS (Supplementary Methods, Supplementary Table 1 and Supplementary Fig. 2). The Cox regression model and XGBoost were also constructed for comparison with a similar strategy (Supplementary Methods, Supplementary Table 1, and Supplementary Fig. 3).

The performance of HERPAI was assessed using C-index for the overall population and pre-specified subgroups defined by menopausal status, node involvement status, grade, Ki67 expression, and adjuvant chemotherapy and radiation therapy reception. The time-dependent area under the curve (AUC) at 3 and 5 years was also reported. The calibration curve was depicted to characterize the calibration ability[20]. Decision curve analysis (DCA) was performed to demonstrate the clinical usefulness[21]. Similar evaluations were performed for the Cox regression model and XGBoost model. We further determined the threshold of HERPAI score to select high-risk patient who has a predicted risk of recurrence greater than 10% at 5 years in the validation cohort[22]. The hazard ratio (HR) between high- and low-risk groups was estimated in overall population and pre-specified subgroups, respectively. The iDFS rate was calculated for 3- and 5-years in the overall population. The predictive value of HERPAI risk group in treatment decisions was explored. The association between HERPAI score and clinical and pathological variables were determined using linear regression in the derivation and internal test cohort. Also, we examined the association between HERPAI with selected genes in patients with genetic data from the derivation and internal test cohort. The associations were determined using linear regression models with and without adjusted for age, menopausal status, and family history of breast cancer, to provide insights into the biological foundation of HERPAI (Supplementary Methods).

All of the statistical analyses were performed using R statistical software (version 4.3.1). A two-tailed *P*-value less than 0.05 was considered as a sign of statistical significance unless otherwise specified.

Results

Characteristics of the participants

A total of 5424 and 916 patients was included in derivation cohort and external test cohort, respectively. The demographic and clinicopathologic characteristic were demonstrated in Table 1. In the derivation cohort, 4882 (90%) were used for development (median age 57 [IQR, 47–66] years, 3319 [68.0%] were T1 and 3740 [76.6%] were N0), while 542 (10%) were utilized for internal test (median age 57 [IQR, 47–66] years, 367 [67.7%] were T1 and 402 [74.2%] were N0). External test cohort were consisting of 916 patients (median age 52 [IQR, 45–62] years, 604 [65.9%] were T1 and 656 [71.6%] were N0) from three independent medical centers. After a median follow-up duration of 58.6 (IQR, 30.8–94.5), 56.1 (IQR, 29.5–92.3), and 53.2 (IQR, 38.0–73.5) months, a total of 305, 32, and 78 iDFS events were observed for training and validation, internal test, and external test, respectively (Supplementary Table 2).

Model performance evaluation

HERPAI yield a C-index of 0.73 (95% CI 0.65–0.81), 0.73 (95% CI 0.62–0.85), and 0.68 (95% CI 0.60–0.77) in validation, internal test, and external test cohort, respectively (Fig. 2A). A 5-years iDFS AUC of 0.75 (95% CI 0.56–0.93), 0.75 (95% CI 0.43–1.00), and 0.71 (95% CI 0.47–0.94) for HERPAI was observed in validation, internal test, and external test cohort, respectively (Fig. 2B). HERPAI demonstrated a good performance with good calibration and DCA in all cohorts (Supplementary Fig. 4–5). The majority of the DCA curve was above the treat-all and treat-none curve, indicating that HERPAI offered desirable trade-off.

Consistent performances were observed among pre-specified key subgroups (Supplementary Table 3). Of note, HERPAI reached C-indices of 0.69 (95% CI 0.52–0.86) and 0.68 (95% CI 0.55–0.81) for patients with and without lymph node involvement, respectively; and among patients with or without adjuvant chemotherapy (0.69 [95% CI 0.59–0.78] and 0.67 [95% CI 0.41–0.93]) or radiation therapy (0.67 [95% CI 0.53–0.82] and 0.69 [95% CI 0.59–0.78]) among external test cohort. The 5-year iDFS AUC among subgroups were shown in Supplementary Table 6.

Moreover, HERPAI outperformed the Cox regression model, XGBoost model, and other known risk factors (including age, T stage, N stage, Ki67 expression, and grade) in predicting iDFS (ranged 0.51–0.61, Supplementary Table 3). HERPAI could also predict DDFS and OS with C-indices of 0.68 (95% CI 0.49–0.86) and 0.71 (95% CI 0.47–0.94), respectively, among external test cohort (Supplementary Table 4–5 and 7–8). Further, HERPAI score was significantly associated with iDFS after adjusted for chemotherapy and radiation therapy (Supplementary Table 9). Among patients with genetic data (N = 2795 [51.5%], Supplementary Table 10), RS were less predictive in comparison with HERPAI (C index for RS, 0.620 [95% CI 0.577–0.662]), especially among those for postmenopausal women (C index 0.594 [95% CI 0.540–0.648]) and ER expression < 50% (C index 0.607 [95% CI 0.472–0.742]), compared with HERPAI (Supplementary Table 11). Sensitivity analysis demonstrated the robustness of HERPAI's performance in predicting iDFS when accounting for the potential effect of COVID pandemic on follow-up and bias due to imputation (Supplementary Table 12). The genetic association between HERPAI score and selected genes among patients with genetic sequencing data were demonstrated in Supplementary Table 13. Higher HERPAI score were associated with significantly expression of *BAG1*, *PR*, and *STMY3* (coefficient [95% CI], 0.015 [0.002–0.028], 0.020 [0.007–0.033], and 0.018 [0.004–0.032], respectively), after adjusted for age, menopausal status, and family history.

Reclassification according to HERPAI score

Approximately 25% of the participants were reclassified as high-risk patients. Elevated recurrence risk was observed for high-risk groups among validation (HR 2.56, 95% CI 1.25–5.22, $P = 0.01$), internal test (HR 2.52, 95% CI 0.97–6.57, $P = 0.06$), and external test cohort (HR 1.94, 95% CI 1.00–3.74, $P = 0.049$) (Fig. 3). Decreased 5-year survival rate were observed for high-risk groups in validation (88.7% vs 95.3%), internal test (89.3% vs 95.4%), and external test cohort (88.6% vs 93.9%) (Supplementary Table 14). The shifting between HERPAI risk group, RS risk group, and N stage was shown in Supplementary Fig. 6. Reclassification ability among subgroups

	Development cohort (N = 4882)	Internal test cohort (N = 542)	External test cohort (N = 916)
Age at diagnosis, median (IQR), years	57 (47–66)	57 (47–66)	52 (45–62)
BMI, median (IQR), kg/m^{2}[a]	23.1 (21.2–25.4)	23.0 (21.5–25.3)	23.7 (21.6–25.9)
Menopausal status, No. (%)			
Pre and peri menopausal	1888 (38.7)	207 (38.2)	445 (48.6)
Postmenopausal	2994 (61.3)	335 (61.8)	471 (51.4)
Breast cancer family history, No. (%)			
No	4526 (92.7)	501 (92.4)	888 (96.9)
Yes	356 (7.3)	41 (7.6)	28 (3.1)
Breast surgery modality, No. (%)			
BCS	2057 (42.1)	220 (40.6)	493 (54.5)
Mastectomy	2825 (57.9)	322 (59.4)	411 (45.5)
Lymph node surgery modality, No. (%)			
SLNB	2201 (45.1)	236 (43.5)	554 (61.8)
ALND	2681 (54.9)	306 (56.5)	342 (38.2)
Histology, No. (%)			
IDC	4246 (87.0)	480 (88.6)	793 (86.6)
ILC	210 (4.3)	24 (4.4)	25 (2.7)
Others[b]	426 (8.7)	38 (7.0)	98 (10.7)
Grade, No. (%)[c]			
I	551 (13.2%)	61 (13.1%)	67 (8.46%)
II	2859 (68.7%)	316 (67.8%)	539 (68.1%)
III	751 (18.0%)	89 (19.1%)	186 (23.5%)
ER expression, median (IQR)	95 (90–95)	95 (90–95)	90 (80–90)
PR expression, median (IQR)	70 (15–90)	70 (15–90)	70 (20–90)
HER2 status			
IHC 0+	1090 (22.3)	118 (21.8)	454 (49.6)
IHC 1+ or 2+	3792 (77.7)	424 (78.2)	462 (50.4)
Ki67 expression, median (IQR)[d]	15 (5–30)	15 (5–30)	20 (10.0–40)
Pathological T stage, No. (%)			
T1mic	41 (1.0)	5 (1.0)	2 (0)
T1a	283 (5.8)	30 (5.5)	44 (4.8)
T1b	726 (14.9)	73 (13.5)	144 (15.7)
T1c	2269 (46.5)	259 (47.8)	414 (45.2)
T2	1563 (32.0)	175 (32.3)	312 (34.1)
Lymph node involvement, No. (%)			
N0	3740 (76.6)	402 (74.2)	656 (71.6)
N1mic	139 (2.9)	17 (3.1)	20 (2.2)
N1	1003 (20.5)	123 (22.7)	240 (26.2)
Adjuvant chemotherapy received, No. (%)			
No	2506 (51.3)	277 (51.1)	238 (26.0)
Yes	2376 (48.7)	265 (48.9)	678 (74.0)
EC/TC	1553 (31.8)	165 (30.4)	300 (32.8)
EC-T	673 (13.8)	86 (15.9)	285 (31.1)
Others[e]	150 (3.1)	14 (2.6)	93 (10.2)
Adjuvant radiation therapy received, No. (%)			
No	2263 (46.4)	258 (47.6)	431 (47.1)
Yes	2619 (53.6)	284 (52.4)	485 (52.9)

Table 1. Characteristics of the participants. *IQR* interquartile range, *SLNB* sentinel lymph node biopsy, *ALND* axillary lymph node dissection, *IDC* invasive ductal carcinoma, *ILC* invasive lobular carcinoma, *ER* estrogen receptor, *PR* progesterone receptor, *HER2* human epidermal growth factor receptor 2, *IHC* immunohistochemistry, *EC* epirubicin plus cyclophosphamide, *TC* docetaxel plus cyclophosphamide, *EC-T* epirubicin plus cyclophosphamide followed by docetaxel. [a]A total of 200 (4.1%), 22 (4.1%), and 15 (1.6%) patients were missing for BMI in the derivation, internal test, and external cohort, respectively. [b]Other histological types include papillary carcinoma, mucinous carcinoma, medullary carcinoma, metaplastic carcinoma, tubular carcinoma, occult breast cancer, and other histology. [c]A total of 721 (14.8%), 76 (14.0%), and 124 (13.5%) patients were missing for grade in the derivation, internal test, and external test cohort, respectively. [d]Three patients were missing for Ki67 expression in the external cohort. [e]Other chemotherapy regimen includes CMF (Cyclophosphamide, Methotrexate, and 5-Fluorouracil), CEF (Cyclophosphamide, Epirubicin, plus 5-Fluorouracil), A-P-C (Adriamycin, Paclitaxel plus Cyclophosphamide).

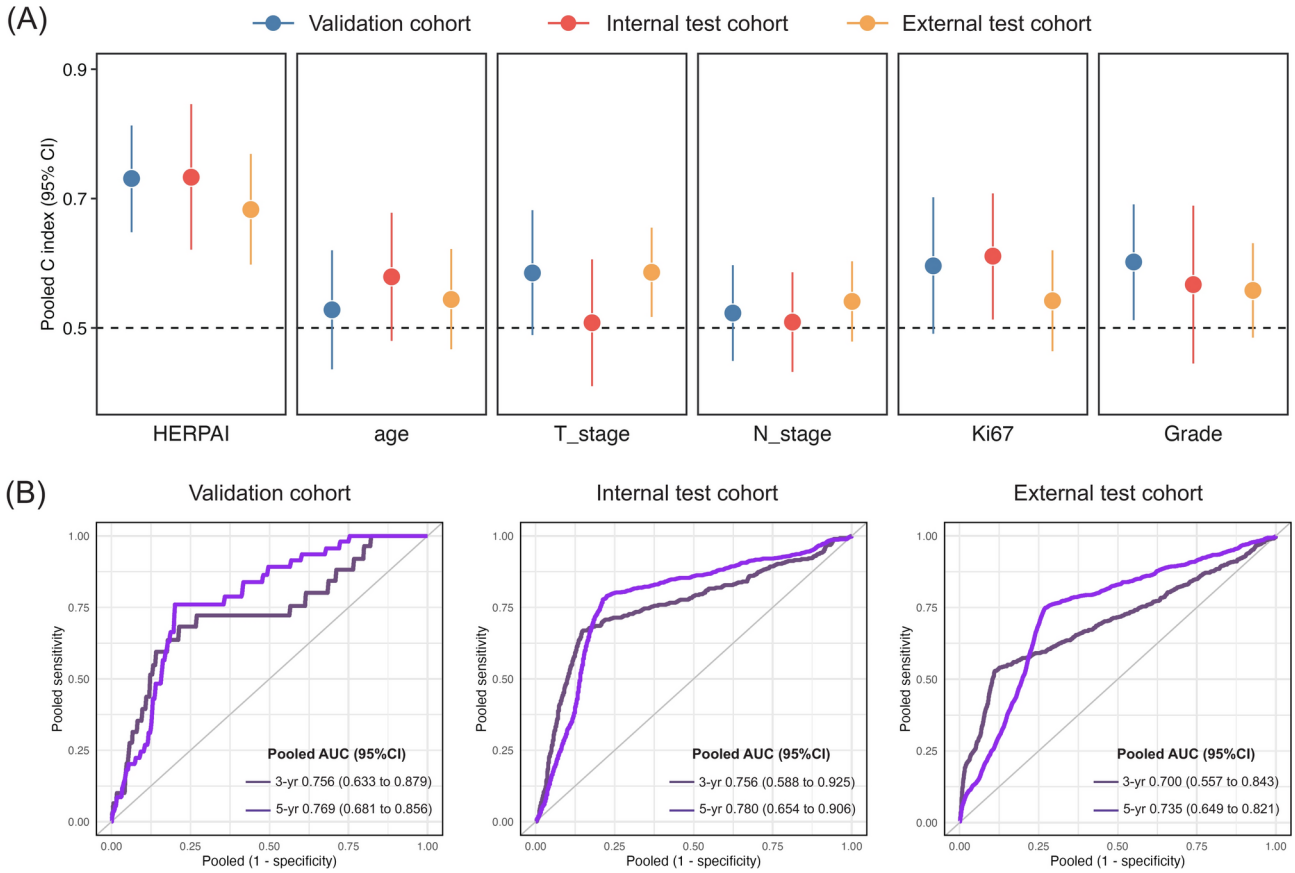

Fig. 2. Model performance in validation, internal test, and external test cohort. (**A**) C index for HERPAI, age, T stage, N stage, grade, and Ki67 in predicting iDFS. (**B**) 3- and 5-year area under receiver operating characteristics curve in predicting iDFS.

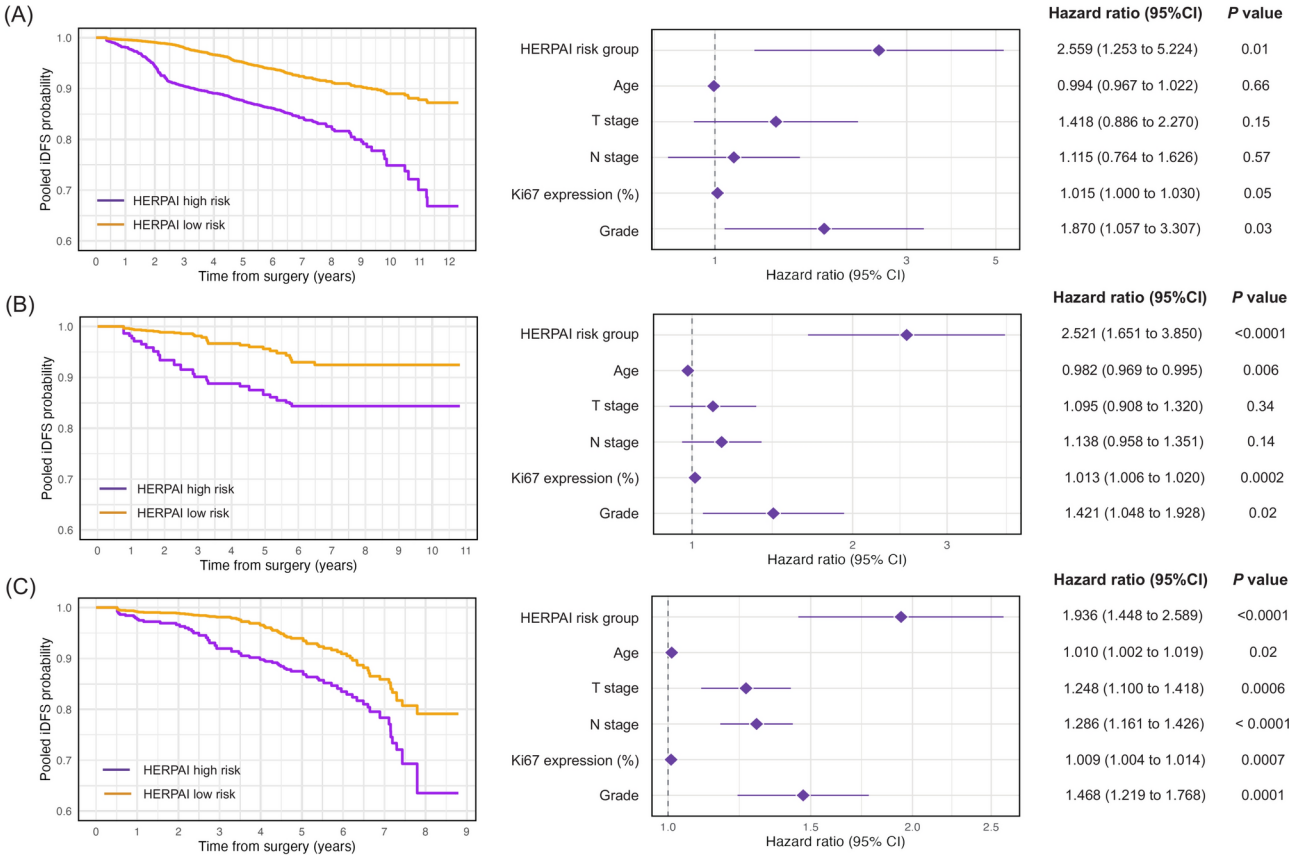

Fig. 3. Reclassification ability of HERPAI among (**A**) validation, (**B**) internal test, and (**C**) external test cohort.

was demonstrated in Supplementary Table 15. No significant interaction effects were observed for HERPAI and adjuvant chemotherapy or radiation therapy receipt (both $P > 0.05$).

Discussion

Leveraging conventional clinical and pathological data, we developed and validated a deep learning-based model, HERPAI, to predict recurrence and risk stratification among early-stage HR+/HER2− BC patients. Reliable performance was observed for HERPAI in both internal and external test cohorts, and key subgroups, including node negative population. Furthermore, HERPAI also had the potential in predicting distant metastases and overall survival. Notably, HERPAI outperformed other prognostic models including the Cox regression and XGBoost based model, and other known risk factors, with a higher C index observed in all cohorts.

Several prognostic and predictive models were proposed for HR+/HER2− BC in the past decades[8–13]. Some of which had been approved by the Food and Drug Administration and widely used, such as Oncotype Dx and MammaPrint. However, among elderly population, of whom the carcinogenesis was less contributed to genetic predisposition[23,24] and was of more heterogeneity[23,25], less predictive ability in guiding adjuvant chemotherapy was observed for genetic tools like Oncotype Dx[8]. In comparison with the prognostic value of RS in predicting recurrence, our model, HERPAI, utilized routine pathological characteristics instead of gene expression, yielded equivalent, and even better performance among most of the subgroups. Further validation in predictive value of HERPAI and comparison with standard-of-care decision tools in the setting of escalating adjuvant treatment are warranted. Other models, such as CTS5 and PREDICT, focused on distant recurrence and overall survival[11,12]. However, early-stage HR+/HER2− BC experienced a certain number of loco-regional recurrence events[3,26,27], which would further translate to distant metastases or death[28]. Here, the STEEP 2.0 criteria were utilized for HERPAI's outcome ascertainment, i.e., iDFS, which consisted of local or regional recurrence, distant metastases, secondary primary cancer, and death[16]. iDFS was widely adopted in clinical trials with adjuvant setting, which was more comprehensive in profiling disease progression and was more suitable for those patients with relatively low risk of recurrence, compared with stage III disease. These inherent advantages made HERPAI more generalizable and more fit for early-stage HR+/HER2− populations.

Moreover, HERPAI showed its potential in identifying high recurrence risk patients even in subgroups without pathological risk factors (such as more advanced N stage and higher grade)[29–32]. For example, in patients without lymph node involvement, the C index for predicting iDFS was greater than 0.68 in all three cohorts. A substantial proportion of these patients were classified as the high-risk group. These high-risk N0 patients featured with almost twice to triple elevated recurrence risk, compared with those low-risk N0 patients determined by HERPAI (HR 2.57, 3.47, and 1.91 among validation, internal test, and external test cohort,

respectively). Escalation and de-escalation adjuvant treatment among HR+/HER2− BC is a debated topic for stage I diseases or tumors without aggressive biological behavior[33–35]. By leveraging HERPAI, physicians might select patients with unmet treatment needs and optimize treatment strategies, even in cases where classical clinical-pathological risk factors were absent.

In addition, HERPAI was developed based on Transformer, a deep learning method. Compared with the linear survival models, deep learning algorithm could explore the interaction between predictors without a pre-specified setting[36,37]. We also employed continuous variables rather than categorical form for predictors including age, BMI, ER, PR, and Ki67 expression, to yield an individual resolution survival prediction. Further, a software is developing and will further validate in a prospective, multicenter cohort for the clinical implementation of HERPAI. The software enables physicians a visualized comparison of the survival benefit from escalated adjuvant therapy, such as escalated ET with CDK4/6 inhibitors, at various time points (Supplementary Fig. 7). The absolute risk reduction contributed to escalated adjuvant therapy was estimated by multiplying the recurrence risk by the risk reduction reported by previous prospective trials[4,5]. It facilitates physicians and patients to discuss the treatment benefit, as well as the potential toxicity risks together.

This research has some strength. First, to the best of our knowledge, this is the largest study to develop a prognostic model in predicting iDFS using multi-center settings. Second, several model strategies were employed, including statistical, machine learning, and deep learning methods, which ensured a rigorous methodology choice[38]. Thirdly, only conventional variables from clinical routine and pathology reports were considered when developing HERPAI. This minimized the cost of risk stratification and maximized the generalizability of HERPAI.

This study also has some limitations. First, the nature of retrospective design may introduce bias. Secondly, relative low mortality in this population requires longer follow-up to further test the robustness of HERPAI in predicting OS. Third, the current follow-up may underestimate the late recurrence in low-risk population. Further update of follow-up is needed to confirm the predictive ability for HERPAI in late recurrence. Besides, the adjuvant regimen depended on physician's preference and may differ between patients though every patient underwent multidisciplinary team discussion[39,40]. The clinical value of HERPAI in guiding adjuvant treatment should be further prospectively validated in a homogeneous cohort which received the same treatment. Further studies targeting certain subgroups, particularly patients without classical clinical-pathological risk factors, were needed to perform a systemic evaluation of HERPAI's value.

Taken together, we developed and validated a deep-learning based model, HERPAI, that utilizes conventional clinical and pathological data to predict recurrence and perform risk reclassification for early-stage HR+/HER2− BC patients. HERPAI was a promising tool for selecting vulnerable patients who were at high-risk of recurrence and may benefit from escalating adjuvant treatment. Well-designed prospective validations are warranted before clinical implementation.

Data availability

The data utilized in this study are available from the corresponding author upon reasonable request. Codes used in this study were available on https://github.com/Haoting-shi/HERPAI.

Received: 19 November 2024; Accepted: 3 March 2025
Published online: 22 March 2025

References

1. Siegel, R. L., Giaquinto, A. N. & Jemal, A. Cancer statistics, 2024. *CA A Cancer J. Clin.* **74**(1), 12–49. https://doi.org/10.3322/caac.21820 (2024).
2. Giaquinto, A. N. et al. Breast cancer statistics, 2022. *CA Cancer J. Clin.* **72**(6), 524–541. https://doi.org/10.3322/caac.21754 (2022).
3. Pan, H. et al. 20-Year risks of breast-cancer recurrence after stopping endocrine therapy at 5 years. *N. Engl. J. Med.* **377**(19), 1836–1846. https://doi.org/10.1056/NEJMoa1701830 (2017).
4. Slamon, D. et al. Ribociclib plus endocrine therapy in early breast cancer. *N. Engl. J. Med.* **390**(12), 1080–1091. https://doi.org/10.1056/NEJMoa2305488 (2024).
5. Johnston, S. R. D. et al. Abemaciclib plus endocrine therapy for hormone receptor-positive, HER2-negative, node-positive, high-risk early breast cancer (monarchE): Results from a preplanned interim analysis of a randomised, open-label, phase 3 trial. *Lancet Oncol.* **24**(1), 77–90. https://doi.org/10.1016/S1470-2045(22)00694-5 (2023).
6. Rastogi, P. et al. adjuvant abemaciclib plus endocrine therapy for hormone receptor-positive, human epidermal growth factor receptor 2-negative, high-risk early breast cancer: Results from a preplanned monarche overall survival interim analysis, including 5-year efficacy outcomes. *J. Clin. Oncol.* **42**(9), 987–993. https://doi.org/10.1200/JCO.23.01994 (2024).
7. Gnant, M. et al. Adjuvant palbociclib for early breast cancer: The PALLAS Trial Results (ABCSG-42/AFT-05/BIG-14-03). *J. Clin. Oncol.* **40**(3), 282–293. https://doi.org/10.1200/JCO.21.02554 (2022).
8. Kalinsky, K. et al. 21-Gene assay to inform chemotherapy benefit in node-positive breast cancer. *N. Engl. J. Med.* **385**(25), 2336–2347. https://doi.org/10.1056/NEJMoa2108873 (2021).
9. Piccart, M. et al. 70-gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. *Lancet Oncol.* **22**(4), 476–488. https://doi.org/10.1016/S1470-2045(21)00007-3 (2021).
10. Cardoso, F. et al. 70-Gene signature as an aid to treatment decisions in early-stage breast cancer. *N. Engl. J. Med.* **375**(8), 717–729. https://doi.org/10.1056/NEJMoa1602253 (2016).
11. Dowsett, M. et al. Integration of clinical variables for the prediction of late distant recurrence in patients with estrogen receptor-positive breast cancer treated with 5 years of endocrine therapy: CTS5. *J. Clin. Oncol.* **36**(19), 1941–1948. https://doi.org/10.1200/JCO.2017.76.4258 (2018).
12. Wishart, G. C. et al. PREDICT: A new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res.* **12**(1), R1. https://doi.org/10.1186/bcr2464 (2010).
13. Mook, S. et al. Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: A hospital-based retrospective cohort study. *Lancet Oncol.* **10**(11), 1070–1076. https://doi.org/10.1016/S1470-2045(09)70254-2 (2009).

14. Giuliano, A. E. et al. Breast cancer-major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *Cancer J. Clin.* **67**(4), 290–303. https://doi.org/10.3322/caac.21393 (2017).
15. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. M. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMJ* **350**, g7594. https://doi.org/10.1136/bmj.g7594 (2015).
16. Tolaney, S. M. et al. Updated standardized definitions for efficacy end points (STEEP) in adjuvant breast cancer clinical trials: STEEP Version 2.0. *JCO* **39**(24), 2720–2731. https://doi.org/10.1200/JCO.20.03613 (2021).
17. White, I. R., Royston, P. & Wood, A. M. Multiple imputation using chained equations: Issues and guidance for practice. *Stat. Med.* **30**(4), 377–399. https://doi.org/10.1002/sim.4067 (2011).
18. Rubin, D. B. *Multiple Imputation for Nonreponse in Surveys* (Wiley, Hoboken, 1987).
19. Gorishniy, Y., Rubachev, I., Khrulkov, V. & Babenko, A. Revisiting Deep Learning Models for Tabular Data. (2023); (arXiv:2106.11959). Accessed September 23, 2024. http://arxiv.org/abs/2106.11959.
20. Gerds, T. A., Andersen, P. K. & Kattan, M. W. Calibration plots for risk prediction models in the presence of competing risks. *Stat. Med.* **33**(18), 3191–3203. https://doi.org/10.1002/sim.6152 (2014).
21. Vickers, A. J. & Elkin, E. B. Decision curve analysis: A novel method for evaluating prediction models. *Med. Decis. Mak.* **26**(6), 565–574. https://doi.org/10.1177/0272989X06295361 (2006).
22. Hippisley-Cox, J. et al. Development and validation of a new algorithm for improved cardiovascular risk prediction. *Nat. Med.* **30**(5), 1440–1447. https://doi.org/10.1038/s41591-024-02905-y (2024).
23. Van Herck, Y. et al. Is cancer biology different in older patients?. *Lancet Healthy Longev.* **2**(10), e663–e677. https://doi.org/10.1016/S2666-7568(21)00179-3 (2021).
24. Mealey, N. E. et al. Mutational landscape differences between young-onset and older-onset breast cancer patients. *BMC Cancer* **20**(1), 212. https://doi.org/10.1186/s12885-020-6684-z (2020).
25. Romain, S., Chinot, O., Guirou, O., Soullière, M. & Martin, P. Biological heterogeneity of er-positive breast cancers in the post-menopausal population. *Int. J. Cancer* **59**(1), 17–19. https://doi.org/10.1002/ijc.2910590105 (1994).
26. Salvo, E. M. et al. Risk of recurrence among patients with HR-positive, HER2-negative, early breast cancer receiving adjuvant endocrine therapy: A systematic review and meta-analysis. *Breast* **57**, 5–17. https://doi.org/10.1016/j.breast.2021.02.009 (2021).
27. Early Breast Cancer Trialists' Collaborative Group. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* **386**(10001), 1341–1352. https://doi.org/10.1016/S0140-6736(15)61074-1 (2015).
28. Sopik, V., Nofech-Mozes, S., Sun, P. & Narod, S. A. The relationship between local recurrence and death in early-stage breast cancer. *Breast Cancer Res. Treat.* **155**(1), 175–185. https://doi.org/10.1007/s10549-015-3666-y (2016).
29. Rakha, E. A. et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res.* **12**(4), 207. https://doi.org/10.1186/bcr2607 (2010).
30. Cardoso, F. et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **30**(10), 1674. https://doi.org/10.1093/annonc/mdz189 (2019).
31. Tonellotto, F. et al. Impact of number of positive lymph nodes and lymph node ratio on survival of women with node-positive breast cancer. *Eur. J. Breast Health* **15**(2), 76–84. https://doi.org/10.5152/ejbh.2019.4414 (2019).
32. Mauriac, L. et al. Predictors of early relapse in postmenopausal women with hormone receptor-positive breast cancer in the BIG 1–98 trial. *Ann. Oncol.* **18**(5), 859–867. https://doi.org/10.1093/annonc/mdm001 (2007).
33. Curigliano, G. et al. De-escalating and escalating treatments for early-stage breast cancer: The St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. *Ann. Oncol.* **28**(8), 1700–1712. https://doi.org/10.1093/annonc/mdx308 (2017).
34. Weiser, R. et al. De-escalation of endocrine therapy in early hormone receptor-positive breast cancer: When is local treatment enough?. *Ann. Surg.* **274**(4), 654–663. https://doi.org/10.1097/SLA.0000000000005064 (2021).
35. Pistilli, B., Lohrisch, C., Sheade, J. & Fleming, G. F. Personalizing adjuvant endocrine therapy for early-stage hormone receptor-positive breast cancer. *Am. Soc. Clin. Oncol. Educ. Book* **42**, 60–72. https://doi.org/10.1200/EDBK_350358 (2022).
36. Howard, F. M. & Pearson, A. T. Prognosis and Treatment of non-small cell lung cancer in the age of deep learning. *JAMA Netw. Open* **3**(6), e206368. https://doi.org/10.1001/jamanetworkopen.2020.6368 (2020).
37. Katzman, J. L. et al. DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med. Res. Methodol.* **18**(1), 24. https://doi.org/10.1186/s12874-018-0482-1 (2018).
38. Clift, A. K. et al. Development and internal-external validation of statistical and machine learning models for breast cancer prognostication: cohort study. *BMJ* **381**, e073800. https://doi.org/10.1136/bmj-2022-073800 (2023).
39. Taylor, C., Shewbridge, A., Harris, J. & Green, J. S. Benefits of multidisciplinary teamwork in the management of breast cancer. *Breast Cancer (Dove Med Press)* **5**, 79–85. https://doi.org/10.2147/BCTT.S35581 (2013).
40. Harbeck, N. Breast cancer is a systemic disease optimally treated by a multidisciplinary team. *Nat. Rev. Dis. Primers* **6**(1), 30. https://doi.org/10.1038/s41572-020-0167-z (2020).

Author contributions

RP: Conceptualization, Methodology, Investigation, Formal Analysis, Visualization, Writing—Original Draft, Writing—Review & Editing; HS: Conceptualization, Methodology, Data Curation, Formal Analysis, Validation, Writing—Original Draft, Writing—Review & Editing; YS: Methodology, Validation, Software, Formal Analysis, Writing—Review & Editing; XW: Formal analysis, Writing—Review & Editing; SZ: Methodology, Formal analysis, Validation, Writing—Review & Editing; NZ: Methodology, Formal analysis, Validation, Writing—Review & Editing; XZ: Methodology, Formal analysis, Validation, Writing—Review & Editing; SD: Data curation, Investigation, Writing—Review & Editing; CH: Data curation, Investigation, Writing—Review & Editing; JW: Investigation, Writing—Review & Editing; WC: Formal analysis, Writing—Review & Editing; XC: Conceptualization, Funding Acquisition, Project administration, Supervision, Writing—Review & Editing; KS: Conceptualization, Funding Acquisition, Resources, Project administration, Supervision, Writing—Review & Editing.

Funding

This work was supported by Natural Science Foundation of Shanghai Science and Technology Committee (Grant number: 23ZR1439500), Innovative Research Team of High-Level Local Universities in Shanghai (SHSMU-ZDCX20212200), and Science and Technology Commission of Shanghai Municipality Shanghai Sailing Program (22YF1426500). The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Declarations

Competing interests
The authors declare no competing interests.

Ethics approval and consent to participate
This study was approved by the Ethical Review Boards of Ruijin Hospital (No. 2023126). All participants included in this study signed the informed consent form.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-92872-2.

Correspondence and requests for materials should be addressed to X.C. or K.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.