Study Protocol

Protocol Title: Clinical Characteristics, Treatment Patterns, and Prognosis of Breast Phyllodes Tumors: A Retrospective Study Based on 20-Year Real-World Data from a National Multicenter Cohort

Sponsor/Institution: Sun Yat-sen Memorial Hospital, Sun Yat-sen University

Version: 1.0

Date: May 13, 2026

Principal Investigator: Nie Yan

1. Abstract (Brief Summary)

This is a national, multicenter, retrospective cohort study based on real-world data from approximately 3,500 patients with breast phyllodes tumor (PT) diagnosed between 2001 and 2023 across 8 tertiary hospitals in China. The study aims to systematically describe the clinicopathological characteristics, treatment patterns, and long-term prognosis of Chinese PT patients. The primary outcomes are local recurrence (LR), distant metastasis (DM), and overall survival (OS). Secondary outcomes include disease characteristics and current treatment practices. This study will provide Chinese population evidence to support personalized treatment and guideline updates.

2. Background

2.1 Overview of Breast Phyllodes Tumor

Phyllodes tumor (PT) of the breast is a rare fibroepithelial neoplasm composed of both stromal and epithelial components, accounting for less than 1% of all breast tumors [1]. Clinically, it typically presents as a rapidly growing, painless mass. According to the World Health Organization (WHO), PT is classified into three grades based on histological features: benign, borderline, and malignant, with reported recurrence rates of 7.1%, 16.7%, and 25.1%, respectively [2]. The median/mean age at diagnosis ranges from 36.4 to 55 years, and the mean tumor size ranges from 3.7 to 11.3 cm. Malignant PT has the highest rate of distant metastasis, ranging from 6.5% to 62.5%, primarily via hematogenous spread, with the lungs and bones being the most common sites [3].

Because PT is rare, large-scale studies on its etiology, diagnosis, treatment, and prognosis are very limited. PT can grow very rapidly; delayed diagnosis and treatment may allow the tumor to become so large that surgery is no longer an option. Current imaging modalities (ultrasound, mammography, MRI) have limited ability to differentiate PT from other breast lesions and to accurately predict histological grade. Moreover, due to the often large tumor size and significant intratumoral heterogeneity, small needle biopsy samples may not fully represent the entire lesion, so accurate histological grading still relies on complete pathological evaluation of the excised tumor. PT is not sensitive to radiotherapy or chemotherapy; the mainstay of treatment is complete surgical excision [4]. The efficacy of adjuvant therapies such as targeted therapy and immunotherapy in PT requires further investigation. Although histological grade can, to some extent, predict patient outcomes, some benign PT may recur repeatedly or even undergo malignant transformation, whereas some malignant PT may remain stable for long periods. The current grading system does not fully explain individual differences in prognosis.

In recent years, most prediction models for PT have been built using Western population data from public databases such as SEER, with little validation in Asian populations. Most are single-center, small-sample retrospective case series, typically including dozens to a few hundred cases, lacking sufficient statistical power [4-9]. Many models have only been validated internally within the same database, without external, especially multicenter, validation, so their real-world predictive performance is

uncertain. Research content has mostly focused on clinical and imaging characteristics, differential diagnosis between benign and malignant PT, and comparisons of different surgical approaches on prognosis. Few studies have further analyzed postoperative recurrence and metastasis patterns. Larger sample sizes are still needed to explore risk factors for rare events such as multiple recurrences, distant metastasis, and death, and to reflect real-world clinical practice.

2.2 Current State of Research – Domestic and International

2.2.1 Clinicopathological Features

International research on breast PT began early. As early as 1774, this rare tumor was described as a "giant fibroadenoma" [10]. In 1838, Johannes Muller first used the term "cystosarcoma phyllodes" to describe PT, characterizing it as a sarcoma-like lesion with a cystic, leaf-like cut surface, rapid clinical growth, and large size. In 1981, the WHO adopted the term "phyllodes tumor" and defined it as "a rare fibroepithelial neoplasm with a leaf-like structure, composed of breast fibrous connective tissue and epithelial components" [11]. With deeper understanding of PT features, the 2019 WHO three-tier classification (benign, borderline, malignant based on tumor borders, stromal cellularity, stromal atypia, mitotic count, stromal overgrowth, and malignant heterologous elements) has become the internationally accepted diagnostic standard [12].

2.2.2 Epidemiology

PT accounts for less than 1% of all breast tumors, and its epidemiological characteristics show distinct racial, regional, and age differences. As early as 1987, Chua et al. noted that PT incidence is higher in non-white populations, particularly in Asian women [13]. Asian PT patients are often younger (peak age 25-30 years), while Western patients are typically in their 40s. In 1993, Bernstein et al. reported that Latinas have a higher risk of PT compared to other racial groups. Age distribution varies by race, with Asian and Latina patients being significantly younger on average than non-Latina white patients [14]. Moten and Goldberg, in a large analysis of 1,202 patients with malignant PT from the SEER database, found that minority groups (Black, Hispanic, Asian) were diagnosed at significantly younger ages and had a higher risk of large tumors (>5 cm) and tumor extension beyond breast tissue [15].

Multiple studies show that the incidence of PT is approximately 2.1 per million women, accounting for 0.3%-1.0% of all breast tumors, with age at diagnosis concentrated between 35 and 55 years (median 45 years). Histologically, about 60%-75% of PT are benign, but both benign and malignant PT have the potential for local recurrence [3,5,6]. A systematic review and meta-analysis by Yu et al. of 52 real-world studies found an overall recurrence rate of approximately 12.6%, with risk increasing significantly with higher pathological grade: recurrence rates for benign, borderline, and malignant PT were approximately 7.1%, 16.7%, and 25.1%, respectively [10].

2.2.3 Treatment Patterns

Surgical resection is the only curative treatment for PT. The basic principle is wide local excision, but the optimal margin width remains controversial. The NCCN guidelines recommend negative margins of ≥1 cm, but several recent studies have questioned this. Rosenberger et al. (2021), in a multicenter cohort of 550 cases, suggested that margin width requirements may be individualized, and that some borderline PT with margins

<1 cm may be safe [16]. A Swedish multicenter real-world study of 191 cases (1999-2018) found that, in multivariate analysis, margin status was not significantly associated with local recurrence or overall survival [17]. PT is not sensitive to radiotherapy or chemotherapy, and the efficacy of targeted therapy, immunotherapy, and other adjuvant treatments requires further study. Regarding lymph node management, meta-analyses and consensus suggest that axillary lymph node metastasis rates are <5%, so routine dissection is not recommended [10].

In 2025, the Breast Surgery Group of the Chinese Society of Surgery, Chinese Medical Association, published the "Clinical Practice Guidelines for the Diagnosis and Treatment of Breast Phyllodes Tumors (2025 Edition)" [18], marking the beginning of standardized diagnosis and treatment of PT in China. The guidelines provide recommendations on surgical margins, lymph node management, and adjuvant therapy based on available evidence, but the guideline developers also noted that due to the lack of high-quality Chinese population data, some recommendations are still mainly based on foreign studies.

2.2.4 Prognostic Factors

Domestic and international studies consistently agree that WHO histological grade (benign, borderline, malignant) is the most important prognostic factor. Higher grade is associated with greater risks of recurrence and metastasis. In addition, margin status, stromal overgrowth, tumor size, and patient age have also been identified as key factors affecting local recurrence [18-20]. Sun Yat-sen Memorial Hospital, Sun Yat-sen University, previously developed and validated a nomogram based on clinical and pathological factors from 334 PT cases using Cox regression to assist in evaluating PT prognosis, with external validation [21]. Benign PT does not significantly affect overall survival, so survival-related prognostic models mainly focus on borderline and malignant PT [9,22].

2.2.5 Gaps in Current Research – Domestic and International

Domestic research on PT started relatively late. Xu et al. (2022) noted that Chinese studies have focused on clinical presentation, pathological diagnosis, and surgical treatment, but sample sizes are generally small, making multivariable adjustment difficult [25]. Most published studies are single-center retrospective analyses of dozens to a few hundred cases, lacking large-scale, national multicenter studies. The largest Asian retrospective study to date, from the electronic database of the First Affiliated Hospital of Xi'an Jiaotong University (Li et al., 2025) [24], is a dual-center study of 829 cases, but due to the rarity of the disease and limited geographic representation, it still cannot fully capture the characteristics of the entire Chinese population.

Although international studies have relatively larger sample sizes, they are mostly based on Western populations, and Chinese patient data are lacking. Moreover, most are single-center retrospective analyses, and sample sizes are still insufficient for adequate subgroup analyses. Currently, there are few systematic studies on recurrence patterns (time to recurrence, number of recurrences, grade progression after recurrence), and the dynamic pattern of PT progression remains to be further elucidated.

3. Study Objectives and Endpoints

3.1 Objectives

3.1.1 Primary Objective

To systematically elucidate the survival and prognostic characteristics of Chinese patients with breast phyllodes tumor (PT) using 20-year real-world data from a national multicenter cohort.

3.1.2 Secondary Objectives

(1) Describe distribution characteristics: To determine the proportion of benign, borderline, and malignant PT in the Chinese population, as well as the distribution patterns of demographic and clinical features.
(2) Evaluate current treatment patterns: To depict the landscape and temporal trends of surgical procedures, margin status, lymph node surgery, and adjuvant radiotherapy/chemotherapy for PT in China over the past two decades.
(3) Quantify prognostic outcomes: To calculate local recurrence rates, distant metastasis rates, and tumor-specific mortality rates for the overall cohort and by histological grade, and to describe the patterns of PT outcomes.

3.1.3 Exploratory Objectives

(1) Analyze recurrence event patterns: To analyze the time distribution of recurrences, number of recurrences, and probability of pathological grade progression after recurrence, providing clinical insights into tumor progression mechanisms.
(2) Identify independent prognostic factors: To screen independent risk factors for local recurrence, distant metastasis, and survival using multivariable regression methods, and to compare differences in prognostic factors across pathological grades, providing evidence for personalized follow-up and treatment decisions.
(3) Lay the foundation for prospective studies: The database generated from this study will provide hypotheses and a theoretical framework for future prospective, interventional clinical research.

3.2 Study Endpoints

3.2.1 Primary Endpoints and Definitions

(1) **Local Recurrence (LR)**

Local recurrence is defined as a pathologically confirmed reappearance of PT in the ipsilateral breast, chest wall, or axillary lymph nodes after initial surgery. Diagnosis is confirmed by postoperative follow-up records (outpatient visits, hospital admissions, imaging findings) and pathology reports.

(2) **Local Recurrence Rate (%)**

The percentage of patients with local recurrence among the total study population.

(3) **Recurrence-Free Survival (RFS)**

The time interval (in months) from the date of initial surgery to the date of first pathologically confirmed local recurrence. Patients without recurrence are censored at the date of last follow-up.

(4) **Cumulative Recurrence Rate**

The cumulative probability of local recurrence at 1, 3, 5, and 10 years after surgery, estimated by the Kaplan-Meier method or competing risk method.

(5) **Distant Metastasis (DM)**

Distant metastasis is defined as radiologically (CT, PET-CT, bone scan, etc.) and/or pathologically confirmed spread to organs outside the breast (lung, liver, bone, brain, etc.).

(6) **Distant Metastasis Rate (%)**

The percentage of patients with distant metastasis among the total study population.

(7) **Distant Metastasis-Free Survival (DMFS)**

The time interval (in months) from the date of initial surgery to the date of first distant metastasis. Patients without metastasis are censored at the date of last follow-up.

(8) **Cumulative Metastasis Rate**

The cumulative probability of distant metastasis at 1, 3, 5, and 10 years after surgery.

(9) **Overall Survival (OS)**

The time interval (in months) from the date of initial surgery to the date of death from any cause. Surviving patients are censored at the date of last follow-up.

(10) **Overall Survival Rate**

The cumulative probability of survival at 1, 3, 5, and 10 years after surgery, estimated by the Kaplan-Meier method.

(11) **Median Survival Time**

The time point (in months) after surgery at which 50% of patients are still alive, estimated from the Kaplan-Meier survival curve.

3.2.2 Secondary Endpoints and Definitions

(1) **Age at diagnosis** – Age (in years) when the patient was first diagnosed with breast PT.

(2) **Menstrual status** – Premenopausal or postmenopausal.

(3) **Childbirth history** – Yes or no.

(4) **Family history of malignant tumors and PT** – Includes family history of breast cancer, other malignant tumors, and PT. Categorized as yes or no.

(5) **Previous fibroadenoma surgery** – History of breast surgery for fibroadenoma. Categorized as yes or no.

(6) **Maximum tumor diameter** – Largest tumor dimension (in cm) measured by preoperative imaging (ultrasound, etc.).

(7) **Multifocality** – Solitary or multiple (≥2 lesions in the same breast).

(8) **Tumor location** – Left breast, right breast, or bilateral.

(9) **Imaging findings** – Descriptions from breast and axillary ultrasound, mammography, breast MRI, chest CT, PET-CT, etc. (e.g., lobulated shape, margins, vascularity).

(10) **Histological grade** – Benign, borderline, or malignant according to WHO criteria.

(11) **Surgical procedure** – Type of initial surgery: wide local excision, total mastectomy, mastectomy with reconstruction, breast-conserving surgery, etc.

(12) **Margin status** – Whether tumor cells are present at the surgical margin on postoperative pathology: negative (no residual tumor) or positive (residual tumor present). Margin distance (in cm) is also recorded.

(13) **Adjuvant chemotherapy** – Whether adjuvant chemotherapy was given after surgery (yes/no), with details of regimen if available.

(14) **Adjuvant radiotherapy** – Whether adjuvant radiotherapy was given after surgery (yes/no), with details of regimen if available.

(15) **Axillary biopsy report** – Whether axillary lymph node biopsy (sentinel lymph node biopsy or axillary lymph node dissection) was performed, and the total number of lymph nodes examined and number of positive nodes recorded.

3.2.3 Secondary Endpoints – Categorized

(1) **Demographic characteristics** – Age at diagnosis, menstrual status, childbirth history, family history of malignant tumors/PT, previous fibroadenoma surgery, history of other malignant tumors (excluding breast cancer).

(2) **Tumor characteristics** – Maximum tumor diameter, multifocality, tumor location, imaging findings, histological grade.

(3) **Treatment modalities** – Surgical procedure, margin status, chemotherapy (yes/no), radiotherapy (yes/no), axillary biopsy report.

4. Study Population

4.1 Inclusion Criteria

(1) Patients with a postoperative paraffin-embedded pathology diagnosis of breast phyllodes tumor (benign, borderline, or malignant) according to WHO criteria.
(2) Patients who underwent surgical treatment at any participating center during the designated time period.
(3) Patients with complete or reasonably complete clinical records, pathology reports, and follow-up data sufficient to extract the core variables required by the study protocol.

4.2 Exclusion Criteria

(1) Patients with a postoperative pathological diagnosis of "phyllodes tumor, uncertain behavior" or with an unclassified grade.
(2) Patients with a history of prior malignant tumor (excluding breast cancer) for whom the time from curative treatment to study enrollment is <5 years, or the expected disease-free survival is <5 years, or who have other concurrent diseases that may affect the assessment of study endpoints.
(3) Patients lost to follow-up or with missing critical data (e.g., pathological grade, surgical procedure, follow-up outcomes).

5. Study Design

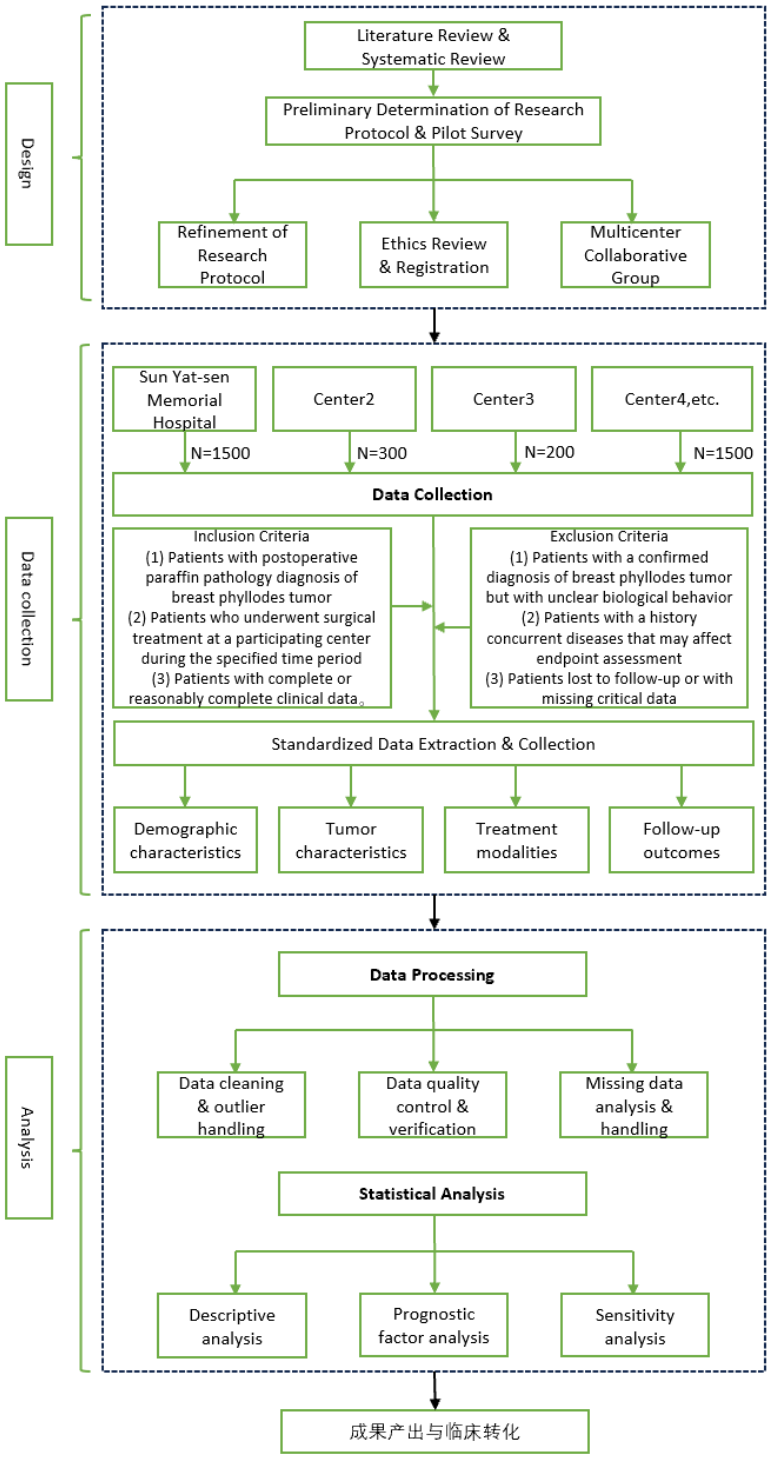
5.1 Overall Design

This is a multicenter, retrospective cohort study. Data will be collected from patients who underwent breast surgery at the Departments of Breast Surgery of the following eight participating centers between January 1, 2001, and December 31, 2023:

- Sun Yat-sen Memorial Hospital, Sun Yat-sen University
- The First People's Hospital of Foshan
- Sun Yat-sen University Cancer Center
- Peking University Shenzhen Hospital
- Foshan Women and Children's Hospital
- The First Affiliated Hospital of Guangzhou Medical University
- Zhongshan People's Hospital
- Shanwei People's Hospital

The expected sample size is approximately 3,500 patients. Using real-world data over nearly 20 years, the study will systematically describe the clinicopathological characteristics, treatment patterns, and long-term prognosis of breast PT in China, and identify independent risk factors through descriptive analysis, propensity score methods, competing risk models, and sensitivity analyses.

5.2 Study Flow Diagram



5.3 Methods to Reduce Bias

(1) **Standardized data collection** – All data extractors at participating centers will receive uniform training. Standardized Case Report Forms (CRFs) will clearly define each variable (e.g., "positive margin" defined as tumor cells present at the margin on pathology; "local recurrence" requires pathological confirmation) to reduce measurement and information bias.

(2) **Double independent extraction of key variables** – For primary outcome variables (recurrence, metastasis, death) and key exposure variables (pathological grade, surgical procedure, margin status), two researchers will independently extract data and cross-validate. Discrepancies will be adjudicated by a third senior researcher.

(3) **Clear inclusion/exclusion criteria** – Strict definitions will avoid selection bias. All eligible consecutive cases will be enrolled; no sampling based on outcome or exposure.

(4) **Handling of missing data** – Cases with >20% missing key variables will be excluded. For variables with ≤20% missingness, multiple imputation (m=5) will be used. Results from complete case analysis and imputed analysis will be compared to assess missingness bias.

(5) **Confounding control** – Multivariable regression models will adjust for known prognostic factors (age, tumor size, pathological grade, margin status, etc.). For non-randomized treatment comparisons (e.g., different surgical procedures), propensity score matching (PSM, caliper = 0.2 × SD of PS) or inverse probability of treatment weighting (IPTW) will be used to balance covariates. E-values will be calculated to evaluate the strength of unmeasured confounding.

(6) **Control of multicenter heterogeneity** – Center effects will be adjusted using Cox shared frailty models or generalized estimating equations (GEE) with center as a random effect.

5.4 Statistical Analysis

5.4.1 Sample Size and Justification

This is a retrospective cohort study; sample size is not based on a pre-specified effect size. All eligible patients with breast PT from the eight centers between January 1, 2001, and December 31, 2023, will be included. Based on each center's historical annual

caseload and pathology database searches, approximately **3,500 patients** are expected (benign:borderline:malignant ≈ 7:2:1).

This sample size is the largest multicenter real-world study in the PT field to date. According to the EPV (events per variable) rule, with local recurrence as the primary endpoint (estimated overall recurrence rate ~11%, yielding ~380 events), the sample size supports multivariable analysis of 30–40 independent variables, fully meeting the requirements for prognostic factor screening.

5.4.2 Statistical Analysis Plan

Statistical analyses will be performed using R version 4.2.0 or higher (key packages: tidyverse, survival, rms, tableone, MatchIt, mice, tidycmprsk, mstate, etc.). The significance level $\alpha = 0.05$ (two-sided), and 95% confidence intervals will be reported.

(1) Descriptive analysis

- **Demographic and clinicopathological characteristics**: Continuous variables will first be tested for normality (Shapiro-Wilk test). Normally distributed variables will be presented as mean ± SD, with group comparisons using t-test or ANOVA; non-normally distributed variables as median (IQR), with group comparisons using Mann-Whitney U or Kruskal-Wallis test. Categorical variables as frequency (%), with group comparisons using Pearson χ^2 test or Fisher's exact test (when expected frequency <5).
- **Treatment patterns**: Distributions and temporal trends (e.g., by 5-year intervals) of surgical procedures, margin status, lymph node surgery, and adjuvant radiotherapy/chemotherapy will be described, stratified by pathological grade and time period.
- **Outcome events**: Local recurrence rate, distant metastasis rate, and tumor-specific mortality rate will be calculated for the overall cohort and by pathological grade. Time to recurrence, number of recurrences, and proportion/probability of grade progression after recurrence will be described.

(2) Analysis of primary endpoints

- **Local recurrence (LR)**: Recurrence-free survival (RFS) curves will be plotted using the Kaplan-Meier method, and cumulative recurrence-free rates at 1, 3, 5,

and 10 years will be calculated. Univariate analysis using log-rank test; multivariable analysis using Cox proportional hazards regression (hazard ratios and 95% CI). When death is a competing risk for recurrence, cumulative incidence functions (CIF) and Fine-Gray competing risk models will be used.

- **Distant metastasis (DM)**: Similarly, distant metastasis-free survival (DMFS) will be analyzed using Kaplan-Meier and Cox regression, with death as competing event.
- **Overall survival (OS)**: Kaplan-Meier method will estimate median OS and 1-, 3-, 5-, 10-year survival rates. Cox regression will identify influencing factors.

(3) Analysis of secondary endpoints

Secondary endpoints (demographic, tumor, treatment characteristics) will mainly be analyzed using descriptive statistics and between-group comparisons as described in (1). Additional exploratory analyses will include these variables as covariates in Cox or logistic regression to assess associations with primary outcomes. Latent class analysis may be used to identify subtypes based on tumor characteristics.

(4) Statistical adjustments

- Multivariable model variable selection: backward stepwise regression (based on AIC) or LASSO penalized regression (when many variables).
- Proportional hazards assumption: tested using Schoenfeld residuals. If violated, stratified Cox, time-dependent covariates, or accelerated failure time models will be used.
- Multiple comparisons: For three primary endpoints (LR, DM, OS), each will be analyzed separately without multiple correction. For multiple exposure comparisons within the same endpoint, Bonferroni correction will be applied.

(5) Bias control (as described in 5.3) – PSM, IPTW, E-value.

(6) Stratified, subgroup, and sensitivity analyses

- **Stratified analysis**: All primary analyses will be stratified by pathological grade (benign/borderline/malignant) to compare prognostic factors across grades.
- **Pre-specified subgroups**: Age (<35 vs. ≥35 years), tumor size (<5 cm vs. ≥5 cm), surgical procedure (segmental resection vs. minimally invasive rotary

excision, wide local excision vs. total mastectomy). Forest plots will display subgroup HRs with interaction P-values.

- **Sensitivity analyses**:
    - Complete case analysis vs. multiple imputation.
    - Varying PSM parameters (1:1 vs. 1:2 matching, caliper 0.1 vs. 0.2).
    - Separate analysis of recurrence definitions (ipsilateral chest wall/axilla only) to exclude misclassification.
    - Excluding patients with follow-up <6 months.
    - For OS, competing risk analysis with death from PT vs. other causes.

6. Data Collection and Management

6.1 Case Report Form / Electronic Data Recording

6.1.1 Case Report Form Design

This study will use an electronic Case Report Form (eCRF) specifically designed according to the study protocol. The eCRF will clearly define all data items to be collected, including but not limited to patient demographics, medical history, diagnostic information, laboratory results, imaging assessments, pathology reports, treatment regimens, and outcome measures.

6.1.2 Data Sources and Recording

All study data will be obtained from the authorized medical information systems of each participating center, primarily including the Hospital Information System (HIS), Laboratory Information System (LIS), Pathology Information System (PIS), and Picture Archiving and Communication System (PACS). Data recorders will extract information directly from these source systems and accurately and completely enter it into the corresponding fields of the eCRF.

6.1.3 Electronic Data Recording System

The electronic data capture system used in this study will have comprehensive functions including user permission management, data audit trails, and electronic signatures, ensuring the standardization and traceability of the data entry process. All modifications to entered data will be automatically recorded by the system, including the modifier,

modification time, and reason for modification.

6.2 Data Management

6.2.1 Data Anonymization

All extracted data will be stripped of direct identifiers such as name and ID number, and will only be identified by a study number. Data access will be restricted to the principal investigator and designated data analysts.

6.2.2 Data Storage

Data will be stored electronically. All final confirmed study data will be stored on a securely encrypted server managed by the hospital information center, within a research-dedicated database.

6.2.3 Data Integrity Checks

- **Source Data Verification (SDV)**: The project team will randomly select 10% of cases. A quality control officer will compare the data in the eCRF with the original hospital electronic medical records item by item to ensure transcription accuracy.
- **Logical Checks**: Logic validation rules will be embedded at the database design stage (e.g., surgery date cannot be later than recurrence date; tumor size cannot be negative). The system will automatically flag data points that violate logic, requiring mandatory review.
- **Missing Value Management**: The system will generate regular missing value reports. For critical variables (e.g., pathological grade, surgical margin status, follow-up outcomes), every effort will be made to trace back to medical records for supplementation. If supplementation is truly impossible, the value will be clearly recorded as "missing" and its impact will be assessed in subsequent statistical analyses.

6.2.4 Database Establishment and Management Methods

The database will contain multiple interrelated tables, such as: patient basic information table, diagnosis information table, surgical treatment table, pathological characteristics table, follow-up information table, and molecular biomarker results table. All tables will be linked through a unique medical record number.

- **Variable Definition**: A detailed Data Dictionary will be developed to clearly define each variable in the eCRF (including variable name, type, value meaning, unit, source of collection, etc.), ensuring consistent understanding among all research personnel.
- **Blinding**: Data entry personnel will be blinded to patients' final follow-up outcomes (whether recurrence/metastasis occurred) during the entry of clinical and pathological data, to avoid introducing subjective bias at the data entry stage.
- **Quality Monitoring**: Regular centralized monitoring of data quality will be performed, with reports on outliers, logical errors, data consistency, and data entry progress.
- **Database Lock**: After final data cleaning, verification, and confirmation, the database will be formally locked, and any unauthorized modifications will be prohibited. Any request to modify locked data must be approved by the principal investigator, with detailed documentation of the reason, content, and date of modification, and the locking process will need to be repeated. The data lock for this study is planned to occur once, just before the formal statistical analysis begins.

7. Ethical Requirements

This study will be conducted in compliance with the Good Clinical Practice (GCP) for Drug Clinical Trials, the Measures for the Administration of Investigator-Initiated Clinical Research in Medical and Health Institutions (for trial implementation), and the Declaration of Helsinki. The study will not begin until the protocol has been approved by the Ethics Committee of our hospital. During the research process, if protocol amendments are necessary, the revised protocol must be resubmitted to the Ethics Committee for review, and the investigator must wait for the Ethics Committee's approval before implementing the new protocol.

This study will collect clinical data and personal information from research subjects for scientific research purposes, which involves patients' privacy rights. All study personnel and data analysts will sign confidentiality agreements and will not disclose

patients' personal information or disease-related information to any individuals or institutions unrelated to this study. All collected patient data will be uniformly managed to prevent any leakage of personal privacy.

8. References

[1] Rowell MD, Perry RR, Hsiu JG, et al. Phyllodes tumors. Am J Surg. 1993;165(3):376-379.

[2] Yu CY, Huang TW, Tam KW. Management of phyllodes tumor: A systematic review and meta-analysis of real-world evidence. Int J Surg. 2022;107:106969.

[3] Zhang Y, Kleer CG. Phyllodes tumor of the breast: Histopathologic features, differential diagnosis, and molecular/genetic updates. Arch Pathol Lab Med. 2016;140(7):665-671.

[4] Chaney AW, Pollack A, McNeese MD, et al. Primary treatment of cystosarcoma phyllodes of the breast. Cancer. 2000;89(7):1502-1511.

[5] Ben Hassouna J, et al. Phyllodes tumors of the breast: a case series of 106 patients. Am J Surg. 2006;192(2):141-147.

[6] Tan PH, et al. Predicting clinical behaviour of breast phyllodes tumours: a nomogram based on histological criteria and surgical margins. J Clin Pathol. 2012;65(1):69-76.

[7] Wang H, Wang X, Wang CF. Comparison of clinical characteristics between benign borderline and malignant phyllodes tumors of the breast. Asian Pac J Cancer Prev. 2014;15(24):10791-10795.

[8] Rosenberger LH, Thomas SM, Nimbkar SN, et al. Contemporary multi-institutional cohort of 550 cases of phyllodes tumors (2007-2017) demonstrates a need for more individualized margin guidelines. J Clin Oncol. 2021;39(3):178-189.

[9] Bartels SAL, van Olmen JP, Scholten AN, et al. Real-world data on malignant and borderline phyllodes tumors of the breast: A population-based study of all 921 cases in the Netherlands (1989-2020). Eur J Cancer. 2024;201:113924.

[10] Fiks A. Cystosarcoma phyllodes of the mammary gland—Muller's tumor. Virchows Arch. 1981;392(1):1-6.

[11] World Health Organization. Histologic typing of breast tumors. 2nd ed. Geneva: WHO; 1981:2.

[12] WHO Classification of Tumours Editorial Board. Breast tumours. In: WHO Classification of Tumours. 5th ed. Lyon: International Agency for Research on Cancer; 2019:2.

[13] Chua CL, Thomas A, Ng BK, et al. Cystosarcoma phyllodes—Asian variations. Surgery. 1988;58:301-305.

[14] Bernstein L, Deapen D, Ross RK. The descriptive epidemiology of malignant cystosarcoma phyllodes tumors of the breast. Cancer. 1993;71(10):3020-3024.

[15] Nash AL, Thomas SM, Nimbkar SN, et al. Racial-ethnic variations in phyllodes tumors among a multicenter United States cohort. J Surg Oncol. 2023;127(3):369-373.

[16] Sars C, Frisell J, Dickman PW, et al. Phyllodes tumors of the breast: Real world data from a multi-institution cohort. Breast. 2025;82:104491.

[17] Chinese Society of Breast Surgery, Chinese Society of Surgery, Chinese Medical Association. Clinical practice guidelines for the diagnosis and treatment of breast phyllodes tumors (2025 edition). Chinese Journal of Practical Surgery. 2025;45(12):1383-1386. (Note: This is a Chinese journal article; for English version, provide translation of title only.)

[18] Liu SL, Ran R, Tu G. Analysis of recurrence, metastasis and prognostic factors in 96 cases of breast phyllodes tumors. Chinese Journal of Bases and Clinics in General Surgery. 2020;27(1):58-62.

[19] Li Y, Song Y, Lang R, et al. Retrospective study of malignant phyllodes tumors of the breast: Younger age, prior fibroadenoma surgery, malignant heterologous elements and surgical margins may predict recurrence. Breast. 2021;57:62-70.

[20] Zhou ZR, Wang CC, Sun XJ, et al. Prognostic factors in breast phyllodes tumors: a nomogram based on a retrospective cohort study of 404 patients. Cancer Med. 2018;7(4):1030-1042.

[21] Chao X, Jin X, Tan C, et al. Re-excision or “wait and watch”—a prediction model in breast phyllodes tumors after surgery. Ann Transl Med. 2020;8(6):371.

[22] Pan C, Han R, Xia S, et al. Construction of a clinical prediction model for overall survival and cancer-specific survival in malignant phyllode tumor of the breast based on the SEER database. Discov Oncol. 2025;16:1200.

[23] Xu JW, Li XK, Yu TH, et al. Current status and research progress in diagnosis and treatment of breast phyllodes tumors. Chinese Journal of General Surgery. 2022;31(11):1527-1534.

[24] Li Y, Chen H, Qu F, et al. Clinicopathological features and risk factors for local recurrence of breast phyllodes tumors: a dual-center, 12-year, retrospective study. Front Oncol. 2025;15:1538523.