Development and Validation of an Ovarian Cancer Risk Prediction Model for Family Members of Ovarian Cancer Probands With BRCA1/2 Germline Mutations

2025-5-19

Contents

I. Project Rationale

1. Research Objectives and Significance

Ovarian cancer is the gynecological malignancy with the highest mortality rate, seriously threatening the life and health of women. One of the main reasons for its high mortality rate is that approximately 70% of patients are diagnosed at an advanced stage. Fortunately, about 1/5 of ovarian cancers are associated with genetic factors, providing us with an opportunity to screen high-risk populations and thereby prevent and diagnose the disease at an early stage and reduce the disease burden. Currently, research related to hereditary ovarian cancer in China is still very scarce, and clinical practice relies on data from foreign studies. However, hereditary tumors have distinct regional and ethnic characteristics, making it urgent to conduct clinical research based on the Chinese population to guide clinical practice in China. Current research suggests that approximately 50% - 60% of hereditary ovarian cancers are closely related to the BRCA1/2 genes. Therefore, accurately assessing the risk of ovarian cancer in BRCA1/2 germline mutation carriers is of great significance for the prevention and treatment of hereditary ovarian cancer. This project intends to conduct a multicenter ambispective cohort study to describe the family characteristics and gene mutation profiles of BRCA1/2 germline mutation carriers in China, analyze the correlations between BRCA1/2 mutation characteristics, family characteristics, clinical features, lifestyles, and the incidence of ovarian cancer; and on this basis, develop and preliminarily validate an ovarian cancer risk prediction model tailored for BRCA1/2 germline mutation carriers in the Chinese population.The clinical significance of this study lies in the development of a predictive model to guide precision clinical interventions. Through risk management strategies, including

prophylactic surgery for genetically high-risk populations and hereditary cancer prevention measures, this approach will enable comprehensive lifetime risk surveillance and ultimately reduce the population-level incidence of ovarian cancer.

2. Current Research Status at Home and Abroad

The pathogenesis of tumors is complex, and the risk assessment of tumors is also relatively complex and specific [1-2]. Different from the risk of ovarian cancer in the general population, the risk assessment of hereditary ovarian cancer requires more consideration of the role of genetic factors. Currently, all international risk prediction models for ovarian cancer are established based on genetic and clinical data from foreign populations and lack data support from the Chinese population. Moreover, existing models only incorporate the presence or absence of mutations as binary variable information into the model and do not consider the characteristics of gene mutations, such as mutation types and sites.

Existing studies have shown that the characteristics of BRCA1/2 gene mutations, such as mutation types and locations, have a significant impact on the age of onset of ovarian cancer. For example, Rebbeck et al.'s study found that mutations in the c.1380 to c.4062 region of the BRCA1 gene (referred to as the ovarian cancer cluster region, OCCR) are associated with a higher risk of ovarian cancer and an earlier age of onset. Similarly, mutations in the c.3249 to c.5681 region of the BRCA2 gene are also associated with an increased risk of ovarian cancer [3]. Marchetti et al.'s study further pointed out that mutation types also have an important impact on the age of onset. For instance, nonsense mutations are associated with an earlier age of onset of high-grade serous ovarian cancer, while frameshift mutations are associated with a later age of onset. This indicates that mutation types may affect the timing of cancer occurrence [4]. Kuchenbaecker et al.'s study provided age-specific risk estimates, showing that the cumulative risk of ovarian cancer in BRCA1 mutation carriers by the age of 80 is 44%, while the risk for BRCA2 mutation carriers is 17%. Additionally, the study noted that the location of mutations within the BRCA genes can further influence these risks [5-6]. In conclusion, the type and location of BRCA1/2 mutations are key factors determining the age of onset of ovarian cancer. Certain mutation types and regions are associated with earlier or later onset ages. This information is crucial for personalized risk assessment and management of BRCA1/2 mutation carriers. However, there is currently no effective risk prediction model for BRCA1/2 mutation carriers that is applicable to the Chinese population and can guide clinical practice. Therefore, it is necessary to establish a risk prediction model for BRCA1/2 germline mutation carriers of ovarian cancer that is suitable for the Chinese population through the construction of a BRCA1/2 gene mutation carrier cohort and to conduct preliminary validation.

3. Support from the Previous Studies of the Research Team Directly Related to This Project

The research team, relying on the National Clinical Research Center for Obstetrics and Gynecology Diseases, has long been engaged in the clinical diagnosis, treatment and scientific research of gynecological malignancies. It was among the first in China to open a genetic counseling clinic for gynecological tumors and has

rich case resources of hereditary ovarian cancer and experience in molecular diagnosis and treatment. Based on the hereditary gynecological tumor diagnosis and treatment platform, the team has established a multi-center hereditary ovarian cancer registration cohort (NCT 06564428) and successfully constructed a primary screening tool for the risk of ovarian cancer in first-degree relatives. This model stratifies first-degree relatives of ovarian cancer patients into risk groups based on five easily obtainable clinical information such as the age of diagnosis of the proband and family history. It aims to screen high-risk individuals without genetic testing information, facilitating its promotion and application at the grassroots level. The results were published in the Gynecologic Oncology journal.

References

[1] National Comprehensive Cancer Network. Genetic/Familial High-Risk Assessment: Breast and Ovarian and Pancreatic, Version 1.2024.

[2] Force USPST, et al. Risk Assessment, Genetic Counseling, and Genetic Testing for BRCA-Related Cancer: US Preventive Services Task Force Recommendation Statement[J]. JAMA, 322(7): 652-665, 2019.

[3] Rebbeck TR, Mitra N, Wan F, et al. Association of Type and Location of BRCA1 and BRCA2 Mutations With Risk of Breast and Ovarian Cancer. JAMA. 2015;313(13):1347-61.

[4] Marchetti C, Ataseven B, Cassani C, et al. Ovarian cancer onset across different BRCA mutation types: a view to a more tailored approach for BRCA mutated patients. Int J Gynecol Cancer. 2023;33(2):257-262.

[5] Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers.JAMA. 2017;317(23): 2402-2416.

[6] National Comprehensive Cancer Network. Genetic/Familial High-Risk Assessment: Breast and Ovarian and Pancreatic, Version 1.2024.

II. Research Aims

1. Primary Research Aims

To develop and validate an ovarian cancer risk prediction model for family members of ovarian cancer probands with BRCA1/2 germline mutations.

2. Secondary Research Aims

① To study the correlation between the clinical efficacy of ovarian cancer in BRCA1/2 germline mutation carriers and the characteristics of BRCA1/2 mutations.

② To investigate the acceptance rate of risk-reducing salpingo-oophorectomy (RRSO) among germline BRCA1/2 mutation carriers and the influencing factors.

III. Research Design (Including the type of research to be conducted, research hypotheses and possible treatment plans involved)
Type of Research Design
This study will adopt a multicenter ambispective cohort study design to comprehensively collect and analyze the clinical characteristics, family characteristics, gene mutation characteristics, and lifestyle data of BRCA1/2 germline mutation carriers. For family members who have not yet developed ovarian cancer at the time of enrollment, telephone follow-ups will be conducted once a year to longitudinally monitor ovarian cancer incidence and to collect diagnostic evidence.

IV. Research Subjects (Including Inclusion Criteria, Exclusion Criteria, and Recruitment Strategies for research Subjects)
Research subjects: Members of ovarian cancer families carrying BRCA1/2 germline mutations
1. Ovarian cancer patients carrying BRCA1 or BRCA2 germline mutations
Inclusion criteria:
① Pathologically diagnosed with ovarian malignant tumor.
② Identified as carriers of BRCA1/2 germline pathogenic or likely pathogenic mutations through genetic testing, in accordance with the "Standards and Guidelines for the Interpretation of Sequence Variants" (2015 Edition) of the American College of Medical Genetics and Genomics (ACMG).
③ Age of 18 years or older.
④ Voluntary participation in this research and signing of the informed consent form.
Exclusion criteria:
① Patients who refuse to provide necessary information.
2. Female first-degree relatives carrying BRCA1 or BRCA2 germline mutations
Inclusion criteria:
① Female first-degree relatives (daughters, mothers, sisters) of the aforementioned ovarian cancer patients.
② Age of 18 years or older.
③ Voluntary participation in this research and signing of the informed consent form.
Exclusion criteria:
① First-degree relatives without blood ties: such as adopted daughters or stepmothers.
3. Withdrawal criteria
Participants can voluntarily withdraw from the study at any time without providing any reasons.

V. Research Outcome Measures (Including Primary Outcome Measures and Secondary Outcome Measures)

1. Primary outcome measure

Definition: Pathologically diagnosed with ovarian malignant tumors (ICD-10 C56).

2. Secondary outcome measures

2.1 Clinical efficacy of ovarian cancer in BRCA1/2 germline mutation carriers

Definition: Progression-free survival (PFS) and overall survival (OS) after standard treatment (such as surgery + chemotherapy).

2.2 Acceptance rate of risk-reducing salpingo-oophorectomy (RRSO) among BRCA1/2 mutation carriers

Definition: The proportion of BRCA1/2 germline mutation carriers in the high-risk population of ovarian cancer who undergo risk-reducing salpingo-oophorectomy (RRSO). RRSO refers to the preventive surgical removal of both fallopian tubes and ovaries to reduce the risk of ovarian cancer, fallopian tube cancer, and peritoneal cancer.

VI. Data Collection and Follow-up Plan

1. Screening of the study population: This multicenter study will enroll ovarian cancer patients with confirmed BRCA1/2 pathogenic variants identified via genetic testing since January 1, 2016. Family pedigrees will be collected, and first-degree relatives will be recruited. After providing informed consent, peripheral blood samples will be obtained for BRCA1/2 germline mutation testing to identify eligible carriers.

2. Enrollment of the study population: All BRCA1/2 germline mutation carriers, including ovarian cancer patients and female first-degree relatives, will be enrolled as research subjects and sign informed consent forms.

3. Data collection

Collection content

Personal history: Height, weight, use of oral contraceptives, use of hormone replacement therapy, smoking, drinking, history of benign gynecological diseases, long-term medication history, etc.

Menstrual and reproductive history: Age of menarche, whether menopause, age of menopause, number of pregnancies, number of deliveries, etc.

Personal history of tumors: Presence of malignant tumors, tumor type, pathological classification, tumor stage, time of onset, treatment methods, treatment outcomes, etc.

Family history of tumors (collected only from probands): Presence of new malignant tumor family history, number of first-, second-, and third-degree relatives with malignant tumors, relationship to the patient, tumor type, pathological classification, tumor stage, age of onset, genetic testing status, etc.

Other disease history: Presence of new-onset circulatory, respiratory, digestive system diseases, etc.

For patients already diagnosed with ovarian cancer, relevant clinical and pathological information of ovarian cancer and onset time should also be collected.

Genetic testing information

1. Basic information on gene mutations

Gene names: BRCA1, BRCA2.

Mutation types:

Point mutations: Missense mutations, nonsense mutations, synonymous mutations.

Insertion/deletion mutations: Frameshift mutations, non-frameshift mutations.

Large segment rearrangements: Deletions, duplications, inversions, translocations.

Mutation location:

Nucleotide position: cDNA position based on the reference sequences (e.g., BRCA1: NM_007294.4, BRCA2: NM_000059.4).

Amino acid position: Protein position based on the reference sequences (e.g., BRCA1: NP_009225.1, BRCA2: NP_000050.2).

Mutation description:

*HGVS nomenclature: Follow the nomenclature rules of the Human Genome Variation Society (HGVS) (e.g., BRCA1: c.68_69delAG, p.Glu23Valfs*17).*

Clinical significance: Classified according to the ACMG 2015 guidelines (pathogenic, likely pathogenic, variant of uncertain significance, likely benign, benign).

2. Information on detection technology

Sequencing technology: Sanger sequencing, next-generation sequencing (NGS), whole exome sequencing (WES), whole genome sequencing (WGS).

Testing scope: Whether it covers the entire gene or specific regions (e.g., hotspot mutation detection).

Testing platform: Illumina, Thermo Fisher, etc.

4. Follow-up plan

① Follow-up content

Personal history: Height, weight, use of oral contraceptives, use of hormone replacement therapy, smoking, drinking, history of benign gynecological diseases, long-term medication history, etc.

Risk-reducing salpingo-oophorectomy:

Whether RRSO was performed (yes/no), surgery time (precise to the year/month), age at surgery (calculated based on the surgery time).

Surgical indication: Whether based on BRCA1/2 mutation status. Whether there are other high-risk factors (such as family history, personal history).

Surgical details: Surgical method (laparoscopic/open), postoperative pathological results (whether there is occult cancer).

Reasons for refusing RRSO (if applicable): Personal preference (such as preserving fertility, fear of surgery), medical advice (such as surgical risk assessment), others (such as economic factors, cultural factors).

Menstrual and reproductive history: Whether menopause, age of menopause, number of pregnancies, number of deliveries, etc.

Personal history of tumors: Presence of new malignant tumors, tumor type, pathological classification, tumor stage, time of onset, treatment methods, treatment effects, etc.

Family history of tumors: Presence of new malignant tumor family history, relationship to the patient, tumor type, pathological classification, tumor stage, age of onset, genetic testing status, etc.

Other disease history: Presence of new-onset circulatory, respiratory, digestive system diseases, etc.

② Follow-up interval

Those without malignant tumors will be followed up once every 12 months, and those with malignant tumors will be followed up once every 6 months. The planned follow-up period is 5 years.

VII. Safety Considerations

1. Protection of Participant Privacy

To ensure the privacy and personal information security of participants, we will take the following measures:

- Data Anonymization: All collected data will be anonymized, removing any information that can directly identify an individual (such as name, ID number, etc.). Data will be stored and processed in the form of unique codes.

- Encrypted Storage: All electronic data will be stored on encrypted servers, and only authorized researchers can access them through secure authentication.

- Physical Security: Paper documents and biological samples will be stored in places with strict access control to ensure that only authorized personnel can access them.

2. Psychological Support

Since BRCA1/2 mutation carriers may experience higher levels of psychological stress and anxiety, especially when they learn of their increased risk of ovarian cancer, we will provide the following psychological support services:

- Professional Psychological Counseling: Free psychological counseling services will be provided to each participant to help them cope with possible emotional fluctuations and psychological burdens.

- Support Groups: Support groups will be established to give participants the opportunity to communicate with others who have similar experiences and share experience and emotional support.

3. Medical Monitoring and Intervention

To ensure the health of participants, we will implement strict medical monitoring and intervention measures:

- Regular Physical Examinations: Assist all participants in undergoing regular physical examinations, especially those related to BRCA1/2 mutation-related tumors (such as ultrasound examinations, tumor marker tests, etc.), to detect potential problems in a timely manner.

- Personalized Intervention Recommendations: Based on risk assessment results, provide personalized prevention and intervention recommendations for high-risk individuals, such as preventive surgeries and drug prevention.

4. Exit Mechanism

Respecting the autonomy of participants, allowing them to withdraw from the study at any time without any adverse effects:

- Exit Procedure: Clearly inform participants how to withdraw from the study and simplify the exit procedure to ensure they can do so easily and without burden.
- Follow-up Support: Even if participants withdraw from the study, we will continue to provide them with necessary medical and psychological support.

VIII. Data Management

Researchers will design the CRF according to the study protocol and submit it to the Academy of Advanced Medical Sciences data management team for review and modification. The study will use the REDCap electronic data collection system for data collection and management. The system is a data collection platform based on computer networks and used for data collection, data transmission, and data management in clinical trials. The system has functions such as real-time data entry, real-time data questioning, data audit trails, data validation, data verification, electronic signatures, data locking, data storage, and data export. The application of the system can ensure the data quality, authenticity and integrity of clinical trials. The entire data management process will be carried out in accordance with national laws and regulations and relevant data management SOP. The main data management processes are listed below, and other details can be found in the Data Management Plan (DMP). The DMP, as a guiding document for data management, will be written by the Data Manager (DM) and approved by the researchers. Data management work will be carried out in accordance with the timelines, content specifications, and methodologies defined in the DMP.

1.Data Collection Roles and Responsibilities

Data collection roles	Abbreviation	Responsibilities
Clinical Research Coordinator	CRC	1. Enter data; 2. Answer queries;
Sub-Investigator	Sub-I	1. Enter data; 2. Answer queries;
Principal Investigator	PI	1. Enter data; 2. Answer queries; 3. Electronic signatures;
Clinical Research Associate	CRA	1. Verify source document consistency; 2. Raise and close queries; 3. Freeze/Unfreeze data;
Project Manager	PM	1. Verify source document consistency; 2. Raise and close queries; 3. Freeze/Unfreeze data;
Data Manager	DM	1. Raise and close queries; 2. Freeze/Unfreeze data; 3. Lock data; 4. Account management.

2. Database Design and Establishment

The database (eCRF) for the project is established by the data manager at Academy of Advanced Medical Sciences, following the SOP standards and setting up edit checks in accordance with the Data Verification Plan (DVP). After passing the test and being approved by the authorized researcher, it is released for use. All role personnel, including PI, Sub-I, CRC, PM, CRA, and DM, must complete protocol-defined training before being granted system access privileges.

3. Data Entry

In accordance with the REDCap user manual, the entry personnel (PI, Sub-I, or CRC) should promptly enter the clinical trial data into the EDC system. The research data should come from original documents such as original medical records and laboratory test reports and be consistent with the original documents. All observations and examination results obtained during the trial must be documented in the eCRF promptly, accurately, completely, clearly, and truthfully in compliance with regulatory standards, and no unauthorized modifications shall be made. All items in the eCRF must be filled in, and no items should be left blank or omitted. If necessary, when making data corrections in the eCRF, the reason for the data modification should be filled as prompted by the system.

4. Data Verification

The REDCap system's automatic data logic verification function is used to implement data verification, improving data management efficiency and data quality. The logic verification program is designed based on the safety and efficacy indicators specified in the clinical protocol and the eCRF. The data logic program includes the verification of valid values and valid ranges for individual data values and the logical verification of the relationships between multiple data values or data values across modules.

5. Source Data Verification and Query

The Clinical Research Associate (CRA) is responsible for on-site monitoring at the center, verifying 100% of the eCRF entered data for correctness and completeness, and consistency with original documents such as the original medical records and laboratory test reports. If any issues are found, they can raise Queries online at any time.

The data manager manages the queries of the trial data based on the Data Verification Plan (DVP). When data are entered into the EDC system and there are any data that do not conform to logic, the system will automatically verify and raise Queries; these Queries need to be reviewed and answered by the researcher or authorized personnel. In addition to the system's automatic verification, any Queries raised through SAS programming or manual verification by the data manager can be manually added to the EDC system when requiring the researcher to clarify/verify/confirm.

6. Data Query Resolution

Researchers need to promptly answer queries. Data managers, monitors, and medical personnel provide query responses. If necessary, queries can be raised again until the data is "clean".

7. Data Review

After data cleaning is completed, the data quality needs to be reviewed. Key data (important data determined by the research protocol) should be 100% checked, and a certain range of non-key data should be spot-checked, and a final Data Verification Report (DVR) and population classification resolution should be formed.

8. Researcher's Signature

The data manager needs to ensure that all queries are cleared, and the researcher conducts an electronic signature review and confirmation. If there are any data revisions after signing, a new signature is required.

9. Data Locking and Exporting

According to the database locking procedure, the data manager, statistical analyst, clinical monitor representative, and researcher sign a written approval of the database lock document. The data manager exports and organizes the standardized database and hands it over to the statistician for statistical analysis. Once the data is locked, it cannot be edited. If there is definite evidence that unlocking is necessary after the data is locked, the researcher and relevant personnel need to sign the database unlock document.`

10. eCRF Archiving

After the study is completed, the data manager will standardize the names of all electronic documents generated during the study, categorize them into standardized folders, and convert important final draft documents into PDF format. Data should be encrypted. Documents that require signatures for archiving need to be printed and signed by the corresponding personnel, and organized into standard folders in chronological order. Project documents should be kept for at least five years.

IX. Statistical Analysis

1. Basis for Sample Size Determination

The main objective of the study is to develop an accurate model for predicting the risk of ovarian cancer in BRCA1/2 mutation carriers. To achieve this objective, it is necessary to ensure that there is a sufficient sample size to capture the complex relationships between different risk factors and to guarantee the stability and accuracy of the model.

Sample size for model development: Considering that the model will include multiple predictor variables (such as gene loci, family characteristics, lifestyle, etc.), it is expected that 10-15 variables will be included in the model. According to the rule of thumb, the number of ovarian cancer cases in the development cohort should be at least 10 times the number of expected included variables, requiring at least 150 probands with primary ovarian cancer. This study plans to include a total of 1,000 people including probands with primary ovarian cancer and their first-degree relatives from Peking University Third Hospital and Peking University First Hospital as the development cohort, with an expected inclusion of at least 500 proband cases and 500 first-degree relatives. It is expected that the development cohort will meet the sample size requirements for model development.

Sample size for model validation: According to previous literature, the ROC of the currently published models is 75%. This study's model is expected to reach 80%,

with a reliability of 95% and a validity of 80%. The ratio of probands with primary ovarian cancer to first-degree relatives is 1:1, and it is expected that 702 people will be needed, including 351 proband cases and 351 first-degree relatives. This study plans to include 1,000 people from the Cancer Hospital, Chinese Academy of Medical Sciences, Peking University People's Hospital, and other institutions, with 500 proband cases and 500 first-degree relatives. It is expected to meet the sample size requirements.

2. Statistical Analysis Methods

In this study, we will employ a series of statistical methods and machine learning techniques to ensure the scientific rigor of data processing and the reliability of the results. The detailed statistical analysis methods are as follows:

(1) Descriptive Statistical Analysis

- Basic Feature Description: Including age, mutation type, family history, menstrual history, etc. Use mean, median, standard deviation, interquartile range, etc. to describe continuous variables; use frequency and percentage to describe categorical variables.

- Chart Presentation: Data distribution was visualized using bar charts, pie charts, boxplots, and other graphical tools for intuitive representation.

(2) Univariate Analysis

- Chi-square Test: Used for the association analysis between categorical variables (such as mutation type, family history, etc.) and the incidence of ovarian cancer.

- t-test or Mann-Whitney U Test: Used for the association analysis between continuous variables (such as age) and the incidence of ovarian cancer.

(3) Multivariate Analysis

- Logistic Regression Model: Used to analyze the influencing factors of binary outcomes (such as whether to have ovarian cancer), estimate the odds ratio (OR) and its 95% confidence interval (CI) of each variable.

- Stepwise Regression Analysis: Screen out the variables that have a significant impact on the incidence of ovarian cancer through forward selection, backward elimination, or stepwise regression methods, and construct the optimal model.

(4) Machine Learning Models

- Model Selection: Select multiple machine learning algorithms (such as random forest, support vector machine, gradient boosting trees, etc.) for model construction, and evaluate the performance of different models through cross-validation techniques.

- Feature Selection: Use recursive feature elimination (RFE), LASSO regression, and other methods for feature selection to optimize model performance.

- Hyperparameter Tuning: Find the best hyperparameter combination through grid search or random search methods to enhance the predictive ability of the model.

- Model Evaluation: Evaluate the predictive performance of the model using indicators such as the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, F1 score, etc.

(5) Internal Validation and External Validation

- Internal Validation: Evaluate the stability and generalization ability of the model in

the development cohort through cross-validation (such as k-fold cross-validation) and Bootstrap resampling methods.

- External Validation: Validate the predictive performance of the model in an independent dataset to ensure generalizability across diverse populations.

(6) Sensitivity Analysis

- Missing Data Handling: Evaluate the impact of different missing data handling methods (such as multiple imputation, maximum likelihood estimation, etc.) on the results.

- Subgroup Analysis: Subgroup analyses were performed (e.g., by age strata or mutation types) to validate model performance across distinct subpopulations.

X. Quality Control Measures in Clinical Research

1. Protocol Modification

During the implementation of the study, to guarantee the quality and scientific rigor of the research, we will strictly adhere to the established protocol. However, reasonable modifications to the protocol will be permitted when necessary. The following are the specific regulations and procedures regarding protocol modification:

(1) Reasons for Modification

- Scientific basis: Based on new scientific discoveries or research outcomes, adjustments to the research methods or the addition of new variables might be necessary.

- Ethical considerations: In accordance with the suggestions or requirements of the ethics committee, modifications to the research protocol will be made to ensure compliance with ethical standards.

- Operational feasibility: During the implementation of the study, if operational difficulties or irrationalities are identified in certain steps or methods, reasonable modification suggestions can be proposed.

- Participant feedback: Based on the feedback from participants, the communication and follow-up mechanisms during the research process will be improved to enhance participant compliance and satisfaction.

(2) Modification Process

- Proposal submission: Any proposed modifications to the protocol should first be raised by the principal investigators within the research team and submitted as a written proposal. The proposal must provide a detailed explanation of the rationale for the changes, the specific modifications, and their anticipated impact.

- Internal review: After submission, the proposal will undergo preliminary evaluation by an internal review committee within the research team. The review committee consists of the project leader, statisticians, clinical doctors, and other relevant experts. The committee will assess the scientific validity, feasibility, and necessity of the proposal.

- Ethical review: For modifications involving ethical issues or having a significant impact on participant rights, they must be submitted to the ethics committee for review and approval. The ethics committee will evaluate whether the modification

conforms to ethical standards and provide corresponding opinions and suggestions.
- Formal approval: After obtaining approval from the ethics committee, the study team will make the final revision to the protocol based on the review committee's opinions and formulate a formal modified document. The modified document must be signed and confirmed by all relevant parties and archived for future reference.
- Notification to participants: If the protocol modification pertains to participants' informed consent form or other matters directly affecting participants, all recruited participants must be promptly notified and their informed consent must be re-obtained.
(3) Record-keeping and Reporting
- Detailed record-keeping: Each protocol modification must be recorded in detail, including the modification date, a detailed description of changes, the approving authority, and signatures of relevant personnel. All modification records will be properly preserved as part of the research archive.
- Regular reporting: The study team will regularly report the status of protocol modifications to the ethics committee and the funding agency to ensure that all parties are promptly informed of the progress of the research and any significant changes.
2. Training

Prior to study initiation, all site investigators will receive clinical trial quality control training conducted by clinical research associates (CRAs). The training program will cover principal investigator，project coordinators, clinical research coordinators (CRCs), laboratory personnel and other relevant site staff.

For the principal investigators and project coordinators, the training content encompasses the research protocol, inclusion and exclusion criteria, specific implementation procedures and details of the research. The principal investigators of each sub-center are responsible for selecting research subjects, while the project coordinators are responsible for supervising the progress of the trial at each center, coordinating the budget, and addressing any issues that may arise during the research to ensure its smooth implementation. For CRCs, the training content includes the filling of CRF forms and precautions, training on the data entry system, training on the informed consent process, and follow-up training. Laboratory technicians will receive training on sample collection, handling, and BRCA gene testing procedures, and standard operating procedures (SOPs) for sample collection, handling, transportation and gene testing will be established.
3. Project Monitoring

For the management of biological samples involved in this study, standardized procedures for sample collection, handling, storage, and transportation will be established, and regular reviews by institutional clinical trial departments and ethics committees across all participating sites will be conducted. Each year, research progress will be reported and monitored, including the filling and storage of CRF forms, the data quality of eCRF, the standardization of genetic testing reports, and the qualifications of laboratories.
4. Quality Control during the Research Process

The study addresses the core issues of low response rate among first-degree

relatives of probands and difficulties in collecting biological samples by implementing a three-level quality control system (self-check by researchers - quality control at sub-centers - supervision by the headquarters), and adopts the following stratified quality control strategies:

4.1 Ensuring the Compliance of Research Subjects

Through the following quality control measures, the response rate of first-degree relatives (the proportion of those who participate in genetic testing among all first-degree relatives) is expected to reach over 80%.

(1) Incentive Mechanism for Relative Participation

- Provide a health care package after blood collection (including interpretation of physical examination reports and basic disease consultation services).

- Establish a relative tracking database, recording contact information, communication time, and reasons for refusal (which need to be classified and coded: such as geographical restrictions, time conflicts, privacy concerns, etc.).

(2) Informed Consent

- Implement a "double-check system" for telephone informed consent: after CRC completes the communication, a quality control specialist will follow up within 48 hours to confirm the understanding level.

4.2 Standardization of Biological Sample Collection

(1) Construction of a Blood Collection Network

Open a "green channel for relative blood collection": relatives can have free blood collection at local cooperative hospitals with an electronic authorization code (The blood collection tubes are pre-labeled with dual QR codes: relative ID + proband linkage code).

(2) Emergency Collection Protocol

Set up alternative plan: for those who cannot have venous blood collection, provide an alternative plan of throat swab collection.

(3) No immunosuppressive drugs are allowed during blood sample and information collection, and there should be no abuse of any drugs or alcohol.

4.3 Multi-dimensional Quality Control of Data Collection

(1) Remote Data Audit

- AI monitoring of the entire telephone interview process: real-time analysis of voice emotions (such as resistance or confusion triggering the intervention of quality control specialists).

- Develop a kinship data dashboard: dynamically display the geographical distribution of response rates (with three-level drill-down: province-city-county), statistics on reasons for loss to follow-up, and sample compliance rates.

(2) Data Cleaning Rules

- Establish a Kinship Data Integrity Index (KDI) = (collected data items / target data items) × (valid biospecimens / target biospecimens)

- Set up automatic cleaning rules: for cases with KDI < 0.7, initiate a manual supplementary collection process; remaining incomplete after 30 days will be flagged as 'Partially Missing Data'.

4.4 Construction of Specialized Quality Control Teams

- Establish a kinship response rate task force: conduct weekly analyses of lost-to-follow-up cases and monthly updates to the communication script library (including specially designed dialect versions for elderly relatives).
- Develop an intelligent reminder system: automatically send care messages during special occasions like relatives' birthdays and traditional festivals to maintain study engagement.

4.5 Quantitative Indicators for Quality Assessment

Quality Control Dimension	Core Indicators	Compliance Standards
Response Rate	Genetic testing rate of first-degree relatives	≥80%
Timeliness	Interval from informed consent to blood collection completion	≤14 days
Completeness	Effective completion rate of electronic questionnaires	≥90%
Biological Samples	Proportion of hemolyzed samples	≤3%
Data Quality	Double-entry inconsistency rate	≤0.5%

XI. Protection of Research Subjects

1. Ethics Committee

The study and each sub-center must obtain approval from the Ethics Committee, which comprises medical professionals, legal experts, and lay members to ensure that the research protocol complies with ethical standards and safeguards subjects' rights and safety.

2. Obtaining Informed Consent from Research Subjects

Before the initiation of the study, researchers will provide all subjects with detailed information, including the purpose, methods, and potential risks of the study, to ensure they have sufficient knowledge to provide informed consent. For those who are unable to sign paper-based informed consent forms, detailed study information will be communicated via WeChat, followed by delivery of an electronic consent form to obtain digital signatures. If neither the paper nor the electronic consent is feasible, the research team will notify them and conduct verbal consent via telephone, with the entire communication process audio-recorded as documentation.

3. Confidentiality and Privacy

The study employs anonymous or de-identified data collection methods and manages the data uniformly through EDC to prevent data leakage and ensure data security and the confidentiality of the subjects' personal information.

4. Risks and Benefits for Research Subjects

(1) Risks for Research Subjects: Participants in this study will have 4ml of peripheral blood collected for genetic testing, which involves physiological risks. Additionally, probands and their relatives need to visit the hospital to collect peripheral blood, incurring economic and time risks.
(2) Benefits for Research Subjects: The greatest benefit for participants in this study is that if the peripheral blood test results show BRCA germline mutations, they will be proactively informed of the results and receive free professional genetic counseling and subsequent treatment plans. Moreover, the cost of sending peripheral blood will be covered by this study.

Although this study may initially impose certain financial and physical burdens on participants, in the long term, each participant will gain clarity regarding their BRCA mutation status, enabling access to more precise medical care. Furthermore, this study aims to establish a BRCA1/2 mutation carrier-based ovarian cancer risk prediction model tailored to the Chinese population. This model will predict ovarian cancer risk in individuals carrying pathogenic or likely pathogenic variants, guiding clinical intervention strategies for high-risk groups and ultimately benefiting broader society.

5. Compensation and Reimbursement

For each study participant, 4ml of peripheral blood will be collected. The costs associated with blood sample shipment and genetic testing for first-degree relatives (excluding the proband) will be covered by this study.

XII. Organization and Management

1. Research Sites and Tasks

Institution	Role and Responsibilities	Planned Enrollment Numbers
Peking University Third Hospital	Co-lead site, Patient enrollment	800 cases
Peking University First Hospital	Co-lead site, Patient enrollment	200 cases
Institute of Advanced Clinical Medicine, Peking University	Protocol design(Yangfeng Wu, Gaoqiang Xie), Data management(Wenyao Ma), Project management(Qian Xu), Statistical analysis(Yongpei Yu)	0
Cancer Hospital, Chinese	Participating site, Patient	200 cases

Academy of Medical Science	enrollment	
Yunnan Cancer Hospital	Participating site, Patient enrollment	200 cases
Tianjin Central Hospital of Gynecology Obstetrics	Participating site, Patient enrollment	200 cases
Peking University People's Hospital	Participating site, Patient enrollment	200 cases
The Second Hospital of Hebei Medical University	Participating site, Patient enrollment	100 cases
General Hospital of Ningxia Medical University	Participating site, Patient enrollment	50 cases
Henan Cancer Hospital	Participating site, Patient enrollment	50 cases

2. Preservation of Research Documents and Records

The original genetic test reports generated during the study will be retained by each participating center. Scanned electronic versions of these reports will be stored in the EDC database.

In this study, peripheral blood samples will be collected from all participants. Laboratories and sample transport personnel involved in this study must be familiar with the types and quantities of samples, as well as the types and methods of sample transport in the hospital, and comply with the relevant provisions of the Regulations of the People's Republic of China on the Administration of Human Genetic Resources for sample transport and preservation.

3. Early termination of the study

This study involves long-term follow-up with no anticipated early termination.

Subject name：_____ Subject ID：_____

Informed Consent Document

Study Title: Development and Validation of an Ovarian Cancer Risk Prediction Model for Family Members of Ovarian Cancer Probands with BRCA1/2 Germline Mutations

Institution: Peking University Third Hospital

Principal Investigator: Professor Hongyan Guo, Department of Gynecology

Dear Participant,

You are invited to participate in the research study titled "Development and Validation of an Ovarian Cancer Risk Prediction Model for Family Members of Ovarian Cancer Patients with BRCA1/2 Mutations." This study is led by Peking University Third Hospital, with Professor Hongyan Guo as the Principal Investigator, and is conducted across nine hospitals in China. The study is supported by the Peking University Century Golden Resources Tengyun Clinical Research Program. The study protocol and this informed consent form have been approved by the Institutional Review Board (IRB) of Peking University Third Hospital.

Before deciding whether to participate, please carefully review the information below. It explains the purpose, procedures, duration, potential benefits, risks, and discomforts associated with this research. You are encouraged to discuss this with family or friends. The investigator will explain the study to you, and you may ask questions at any time.

Background

Ovarian cancer is the most lethal gynecologic malignancy, posing a severe threat to women’s health. Approximately 70% of patients are diagnosed at an advanced stage, contributing to its high mortality. Notably, 20% of ovarian cancers are linked to hereditary factors, providing an opportunity to identify high-risk individuals for

prevention, early diagnosis, and reduced disease burden.

Current research on hereditary ovarian cancer in China is scarce, and clinical practice relies heavily on foreign data. However, hereditary cancers exhibit distinct geographic and ethnic variations. There is an urgent need for China-specific clinical data to guide local practice. Studies indicate that 50%–60% of hereditary ovarian cancers are closely associated with BRCA1/2 germline mutations. Accurate assessment of ovarian cancer risk in BRCA1/2 mutation carriers is critical for prevention and early intervention.

Study Purpose

This study aims to establish a bidirectional cohort of BRCA1/2 mutation carriers to comprehensively collect genetic, familial, and lifestyle data. Using statistical modeling and machine learning, we will develop a risk prediction model to:

1. accurately predict individual ovarian cancer risk in BRCA1/2 mutation carriers.
2. provide clinical decision support for optimal intervention timing.
3. reduce ovarian cancer incidence and improve quality of life.

Study Design

A cohort study design will be used to analyze germline mutation profiles, family characteristics, and lifestyle data of BRCA1/2 mutation carriers.

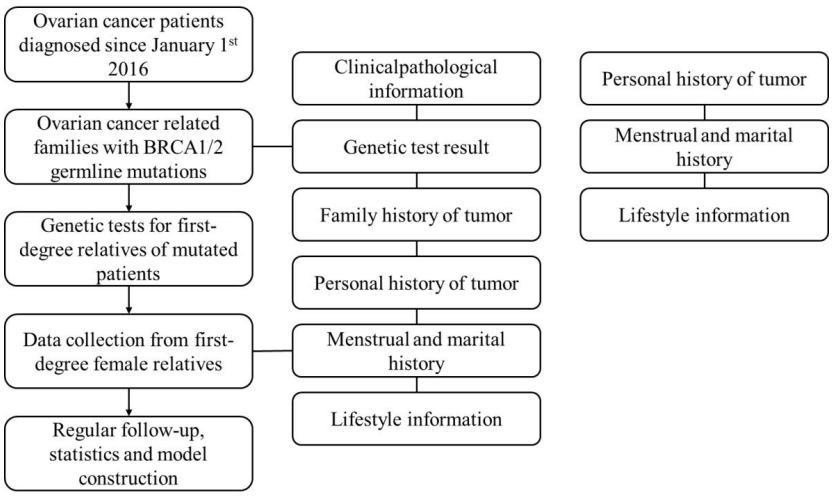

Figure 1: Study Flowchart

Study population

Family members of ovarian cancer patients with BRCA1/2 mutations.

1. Ovarian cancer patients with BRCA1/BRCA2 germline mutations

Inclusion criteria:

① Pathologically confirmed epithelial ovarian cancer.

② Confirmed BRCA1/2 germline pathogenic/likely pathogenic mutation per ACMG guidelines (2015).

③ Age ≥18 years.

④ Voluntary participation with signed informed consent.

Exclusion criteria:

① Unwillingness to provide necessary information.

2. Female first-degree relatives with BRCA1/BRCA2 germline mutations

Inclusion criteria:

① Female first-degree relative (daughter, mother, sister) of an enrolled patient.

② Age ≥18 years.

③ Voluntary participation with signed informed consent.

Exclusion criteria:

① Non-biological first-degree relatives (e.g., adopted daughters, stepmothers).

3. Withdrawal Criteria

① Participants may withdraw at any time without penalty or impact on medical care.

Sample Size & Duration

This 3-year study plans to enroll 2,000 participants (probands and female first-degree relatives).

Procedures

Blood Collection: 4 mL peripheral blood for genetic testing.

Study Timeline & Follow-up

Patients with cancer: Follow-up every 6 months (via phone/clinic).

Relatives: Annual follow-up.

Data Collected: Lifestyle, menstrual/reproductive history, personal/family cancer history.

Risks

1. Genetic Testing Limitations: Results are for research only. Clinical decisions require confirmatory clinical genetic testing.
2. Blood Draw Risks: Minor discomfort, bruising, or rare risks including infection or excessive bleeding.

Reporting Adverse Events: Notify the study physician immediately if you experience discomfort or health changes. Urgent medical issues will be communicated to you/family for prompt clinical evaluation. You will be informed of new findings affecting your participation.

Prohibited Medications

Immunosuppressants, substance abuse, or alcohol misuse during sample collection.

Costs & Compensation

Genetic testing and blood shipment costs for relatives (non-probands) are covered by the study.

Benefits

You will receive BRCA1/2 genetic results for early screening and family planning guidance. Findings may aid other families with hereditary ovarian cancer.

Note: No direct therapeutic benefit is guaranteed.

Voluntary Participation & Withdrawal

Participation is voluntary. You may withdraw anytime without affecting your medical care. The investigator may withdraw you if continuing is not in your best interest.

New Information Disclosure

New findings about your genetic results or emerging therapies will be shared with you.

Privacy & Confidentiality

Identifiers (e.g., name, gender) will be replaced with codes. Only authorized researchers will access your data. Regulatory/ethics committees may audit records per protocol. Published results will not disclose your identity.

Contact Information

Study Questions: Dr. Yuan Li, Phone: +86 15611908183

Rights Concerns: IRB Office, Peking University Third Hospital, Phone: +86 (010) 82265571 / 82265176

Investigator's Statement

"I have explained the study background, purpose, procedures, risks, and benefits to the participant. I provided time to review this form, discuss with others, and answered all questions. I informed the participant of their right to contact the investigator or IRB with concerns and to withdraw anytime without penalty. A copy of this signed consent form will be provided."

Investigator's Signature: _____

Phone: _____ Date: _____

Subject's Consent Statement

"I understand the study's background, purpose, procedures, risks, and benefits. I had sufficient time to ask questions and received satisfactory answers. I know whom to contact with concerns. I voluntarily agree to participate and confirm that all provided information is accurate. I understand I may withdraw at any time without penalty. I will receive a copy of this signed consent form."

Subject's Signature: _____

Phone: _____ Date: _____

(For subjects lacking legal capacity, complete below)

Legally Authorized Representative Signature: _____

Relationship to Subject: _____

Date:_____

(For subjects unable to read this form)

Witness Signature: _____ Date: _____

END OF DOCUMENT