



University of Science and Technology Beijing

# Shrink: A Breast Cancer Risk Assessment Model Based on Medical Social Network

Rui Wang

[wangrui@ustb.edu.cn](mailto:wangrui@ustb.edu.cn)

The 2<sup>nd</sup> China-Europe Forum on Smart Healthy  
Aging and Privacy Protection (Share2018)

# Outline

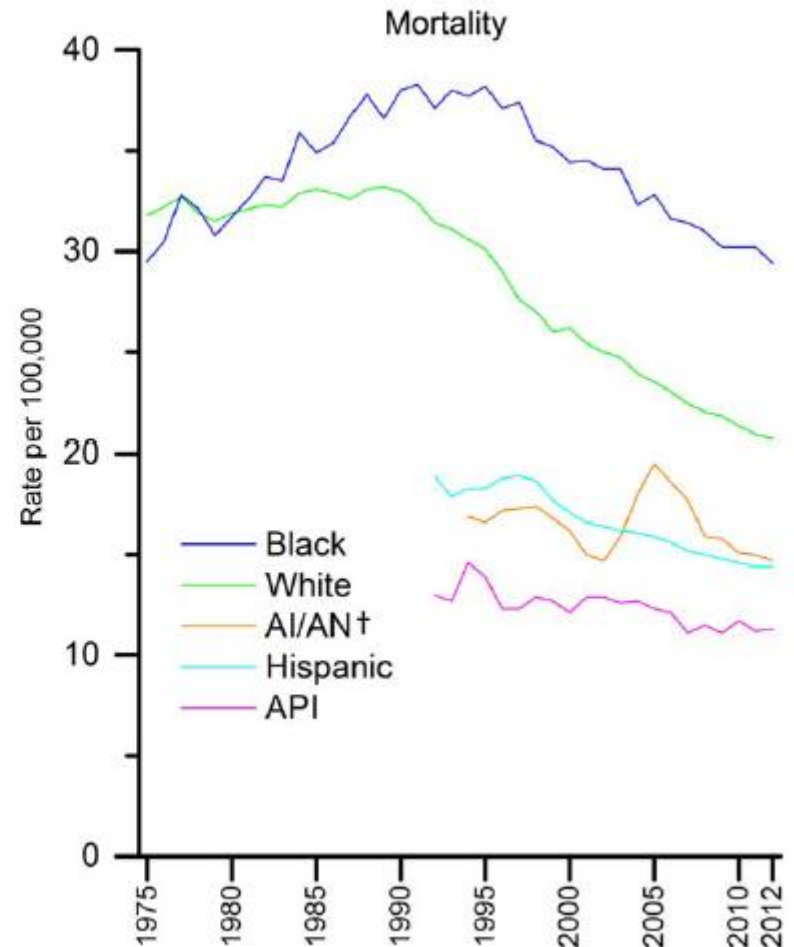
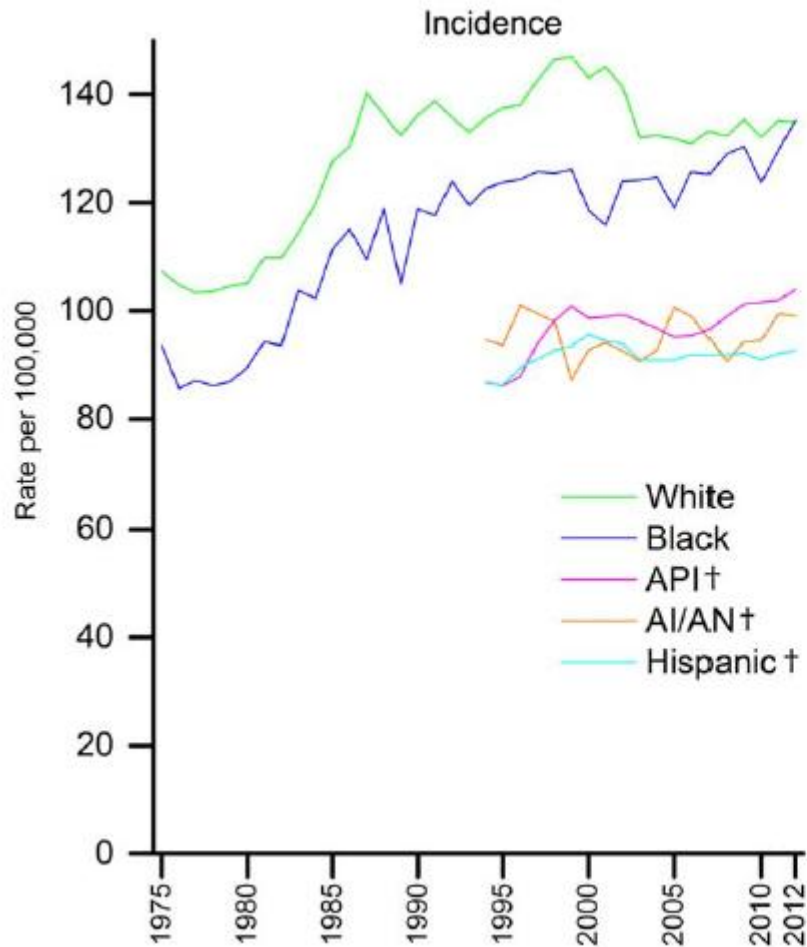
- ✓ Research background
- ✓ Shrink model
- ✓ Experiment

# Outline

- ✓ Research background
- ✓ Shrink model
- ✓ Experiment

# Research background

-The serious incidence of breast cancer



Breast cancer incidence and mortality

# Research background

-The difficulty in implementing early detection



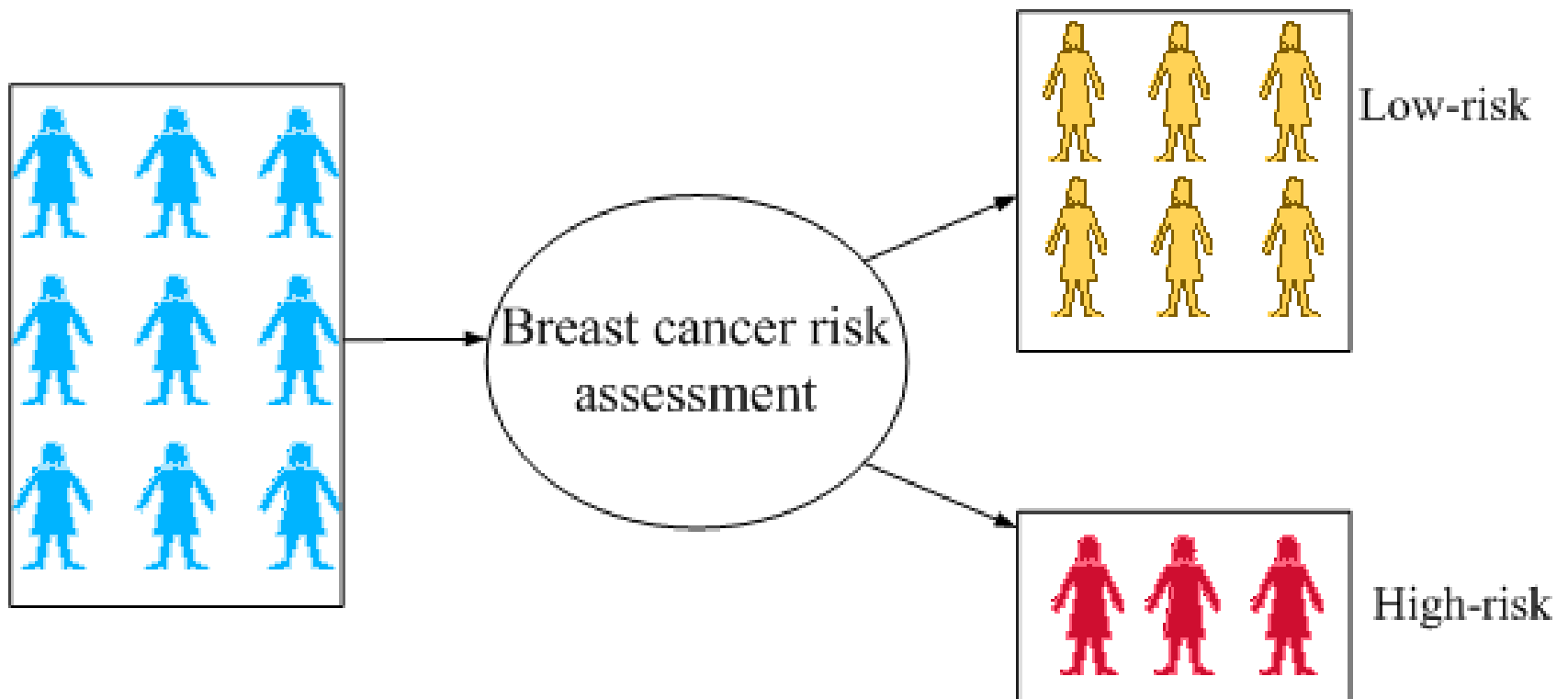
Community hospital; backward



Populous country; Cost

# Research background

-breast cancer risk assessment model



# Research background

-Existing risk assessment models

- <https://www.cancer.gov/bcrisktool/>

The screenshot shows the National Cancer Institute's Breast Cancer Risk Assessment Tool. The header includes the NCI logo and the text "National Cancer Institute at the National Institutes of Health | www.cancer.gov". The main heading is "Breast Cancer Risk Assessment Tool" with the subtitle "An interactive tool to help estimate a woman's risk of developing breast cancer". A navigation menu on the left includes "Get Started with the Risk Tool", "About the Tool", "Breast Cancer Risk Factors", and "Download Source Code". Below the menu are sections for "Page Options" (Print Page) and "Quick Links" (Breast Cancer Home Page, Breast Cancer: Prevention, Genetics, Causes, Current Clinical Trials: Breast Cancer In Situ Treatment, Current Clinical Trials: Breast Cancer Prevention, Current Clinical Trials: Breast Cancer Screening, Breast Cancer Risk in American Women). A "Need Help?" link is at the bottom left. The main content area contains an introduction, a "Risk Tool" section with a list of six questions, and a "Last modified date: 05/1" note.

NATIONAL CANCER INSTITUTE National Cancer Institute at the National Institutes of Health | www.cancer.gov

## Breast Cancer Risk Assessment Tool

An interactive tool to help estimate a woman's risk of developing breast cancer

Last modified date: 05/1

> **Get Started with the Risk Tool**

About the Tool

Breast Cancer Risk Factors

Download Source Code

**Page Options**

Print Page

**Quick Links**

[Breast Cancer Home Page](#)

[Breast Cancer: Prevention, Genetics, Causes](#)

[Current Clinical Trials: Breast Cancer In Situ Treatment](#)

[Current Clinical Trials: Breast Cancer Prevention](#)

[Current Clinical Trials: Breast Cancer Screening](#)

[Breast Cancer Risk in American Women](#)

**Need Help?**

The Breast Cancer Risk Assessment Tool is an interactive tool designed by scientists at National Cancer Institute (NCI) and the [National Surgical Adjuvant Breast and Bowel Project \(NSABP\)](#) to estimate a woman's risk of developing [invasive breast cancer](#). See [About the Tool](#) for more information.

The Breast Cancer Risk Assessment Tool may be updated periodically as new data or research becomes available.

### Risk Tool

(Click a question number for a brief explanation, or [read all explanations](#).)

1. Does the woman have a medical history of any breast cancer or of [ductal carcinoma in situ \(DCIS\)](#) or [lobular carcinoma in situ \(LCIS\)](#) or has she received previous radiation therapy to the chest for treatment of Hodgkin lymphoma?
2. Does the woman have a mutation in either the *BRCA1* or *BRCA2* gene, or a diagnosis of a genetic syndrome that may be associated with elevated risk of breast cancer?
3. What is the woman's age?  
*This tool only calculates risk for women 35 years of age or older.*
4. What was the woman's age at the time of her first [menstrual period](#)?
5. What was the woman's age at the time of her first live birth of a child?
6. How many of the woman's first-degree relatives - mother, sisters, daughters - have had breast cancer?

# Outline

- ✓ Research background
- ✓ Shrink model
- ✓ Experiment



# Shrink

-Basic idea



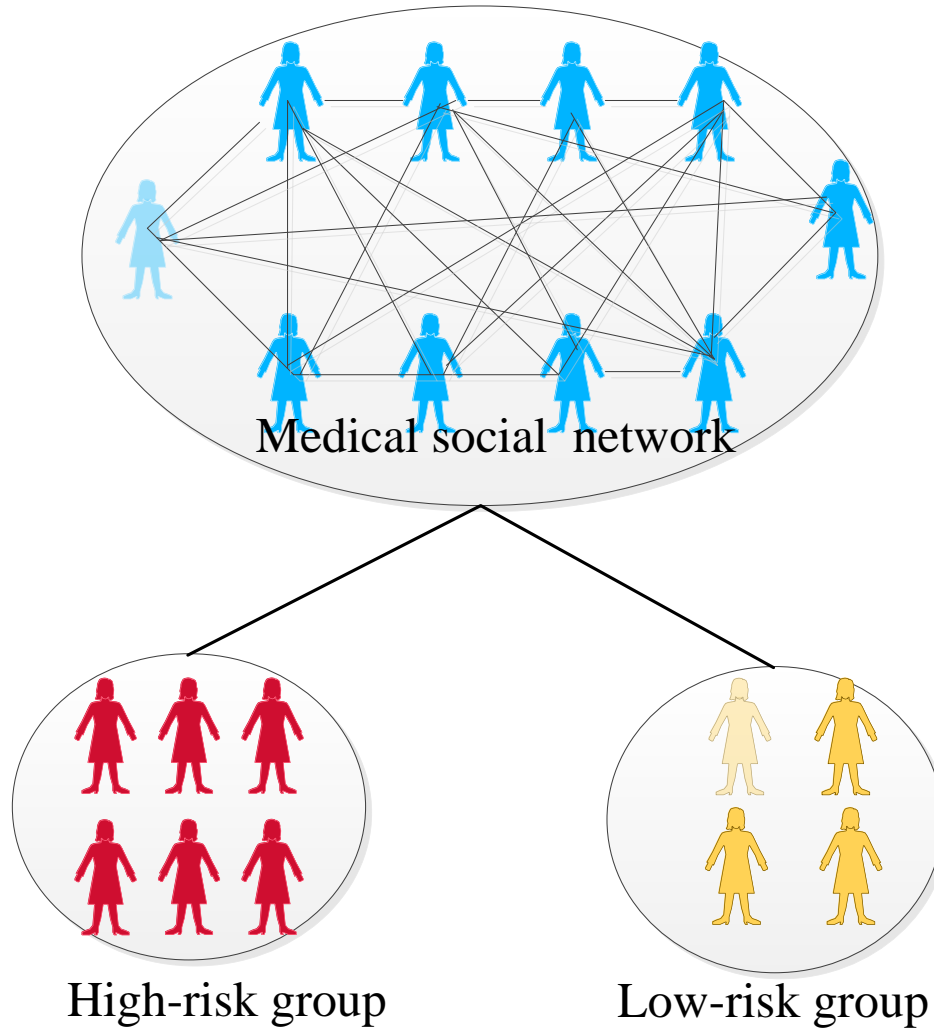
similar habits and customs

probability of suffering  
the same kind of disease

High

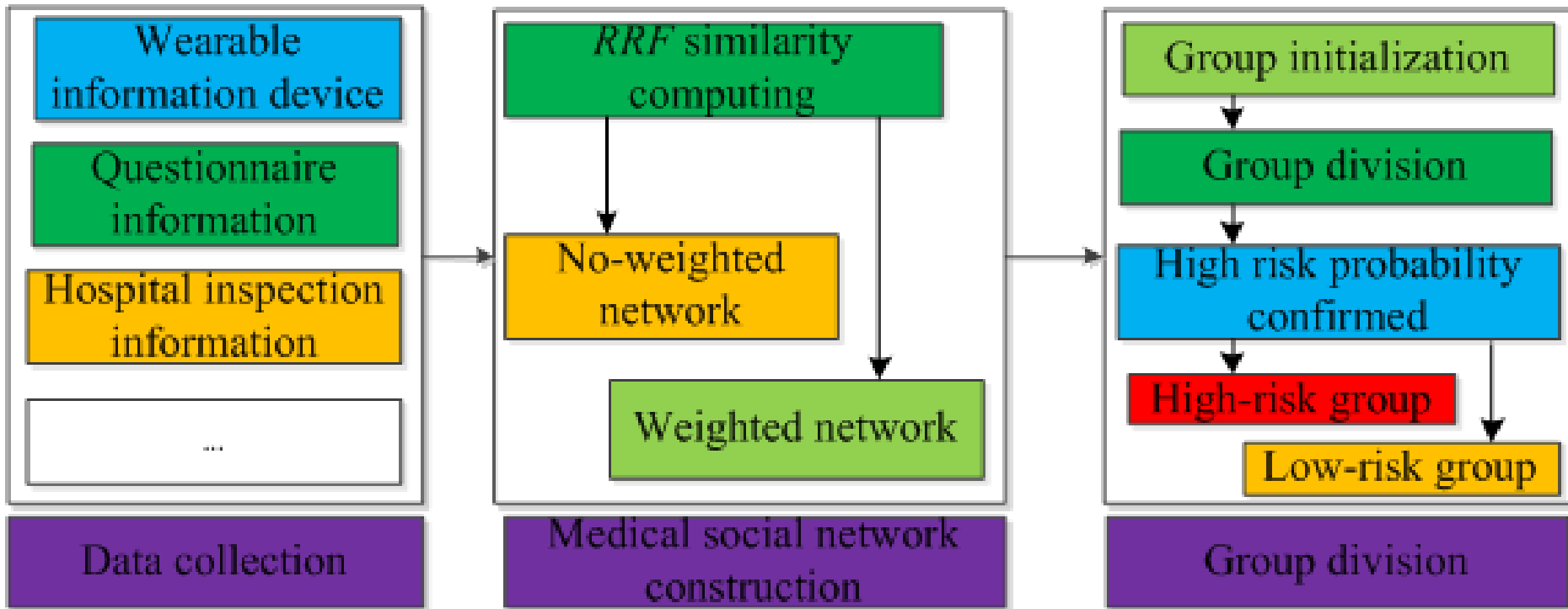
# Shrink

-The purpose



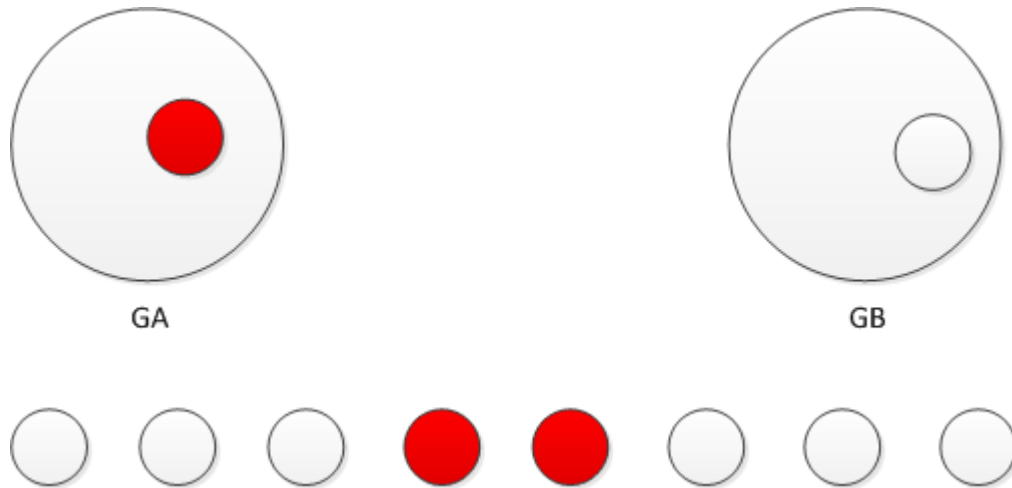
# Shrink

-The module



# Shrink

-Group initialization

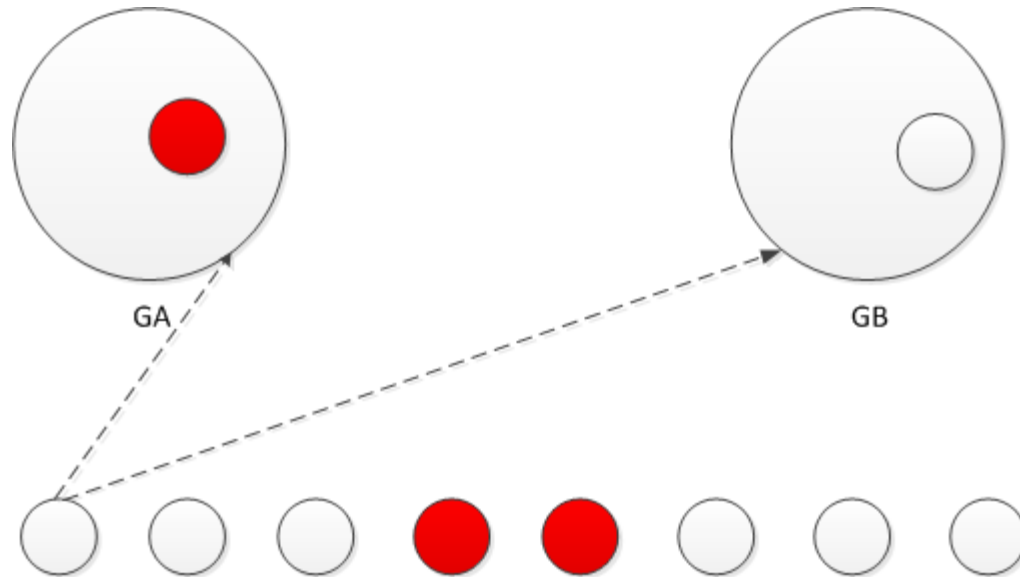


$GA = \{x \mid \text{only one } x \text{ is selected from } P \text{ according to medical knowledge}\},$

$GB = \{y \mid s_{mf_i}(x, y) \text{ is minimum, } x \in GA, y \in N\} \quad (2)$

# Group division

- Group division : Moving node



*The change of **modularity** is computed.*

# Shrink

- Group division :Modularity

- **Modularity (Q)** is the fraction of the edges that fall within the given groups **minus** the expected such fraction if edges were distributed at random.

- $Q = \sum_{i=1}^c (e_{ii} - a_i^2)$

- $E_{ij}$  means that the number of edges that connect group i and j.
- $a_i$  means that the number of edges that connect group i.

$$\Delta Q = \left[ \frac{\sum in + 2K_{i,in}}{2m} - \left( \frac{\sum_{tot} + K_i}{2m} \right)^2 \right] - \left[ \frac{\sum in}{2m} - \left( \frac{\sum_{tot}}{2m} \right)^2 - \left( \frac{K_i}{2m} \right)^2 \right]$$

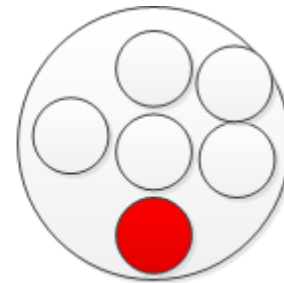
# Shrink

-Node join in GA or GB

- If  $\Delta Q_{GA} > \Delta Q_{GB}$ , the node joins in *GA* group, else it joins in *GB* group.
- Lastly, each node will join in *GA* or *GB*.



GA



GB

# Outline

- ✓ Research background
- ✓ Shrink model
- ✓ Experiment



# EXPERIMENT

## —DATA

- The survey data is collected through **questionnaire**. (1) demographic characteristics (2) female physiological and reproductive factors: (3) medical and family history: (4) dietary habits: frequency of the intake of various types of food; (5) lifestyle habits: (6) breast-cancer-related knowledge:
- More than **120,000** people accepts breast cancer epidemiological survey. In the end, **103,679** hale people and **301** breast cancer patients are confirmed as test cases.

# Experiment

## —Setting

- **Test 1:** Firstly, the test will prove the assessment value of Shrink with different *RRF* in order to discover suitable *RRF* to be used by Shrink.
- *RRF4* = {family history of breast cancer, personal history of benign breast disease, current life satisfactory degree, miscarriages times},
- *RRF5* = *RRF4* + {age at first birth}
- *RRF6* = *RRF5* + {BMI}
- *RRF7* = *RRF6* + {age at menarche}
- *RRF8* = *RRF7* + {age}
- *RRF9* = *RRF8* + {diabetes}
- *RRF10* = *RRF9* + {urban or rural}

# Experiment

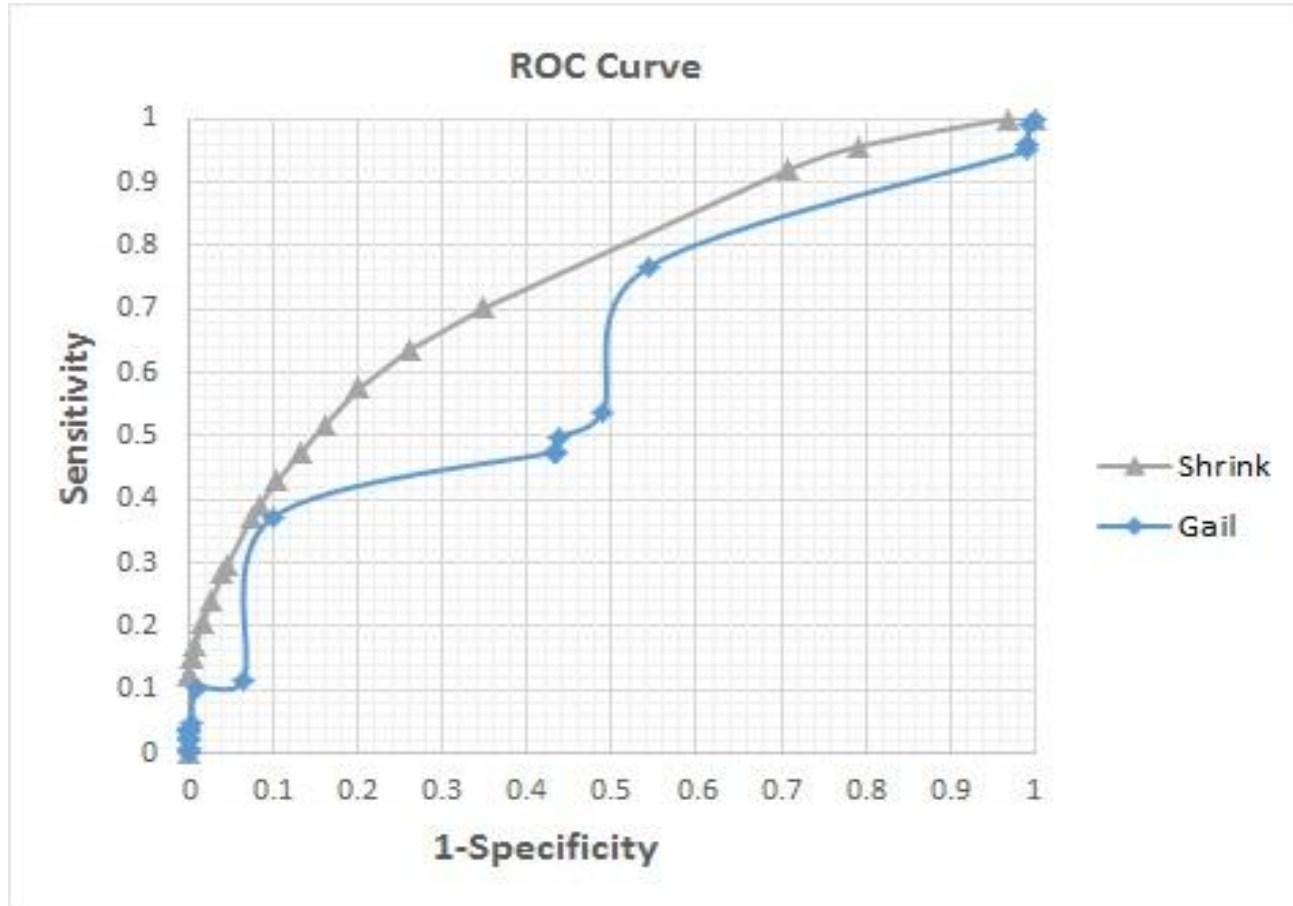
## —Setting

- **Test 2:** The test2 will compare the assessment value between **Gail and Shrink**. The *RRF* will be selected according to the result of test 1.



# Experiment

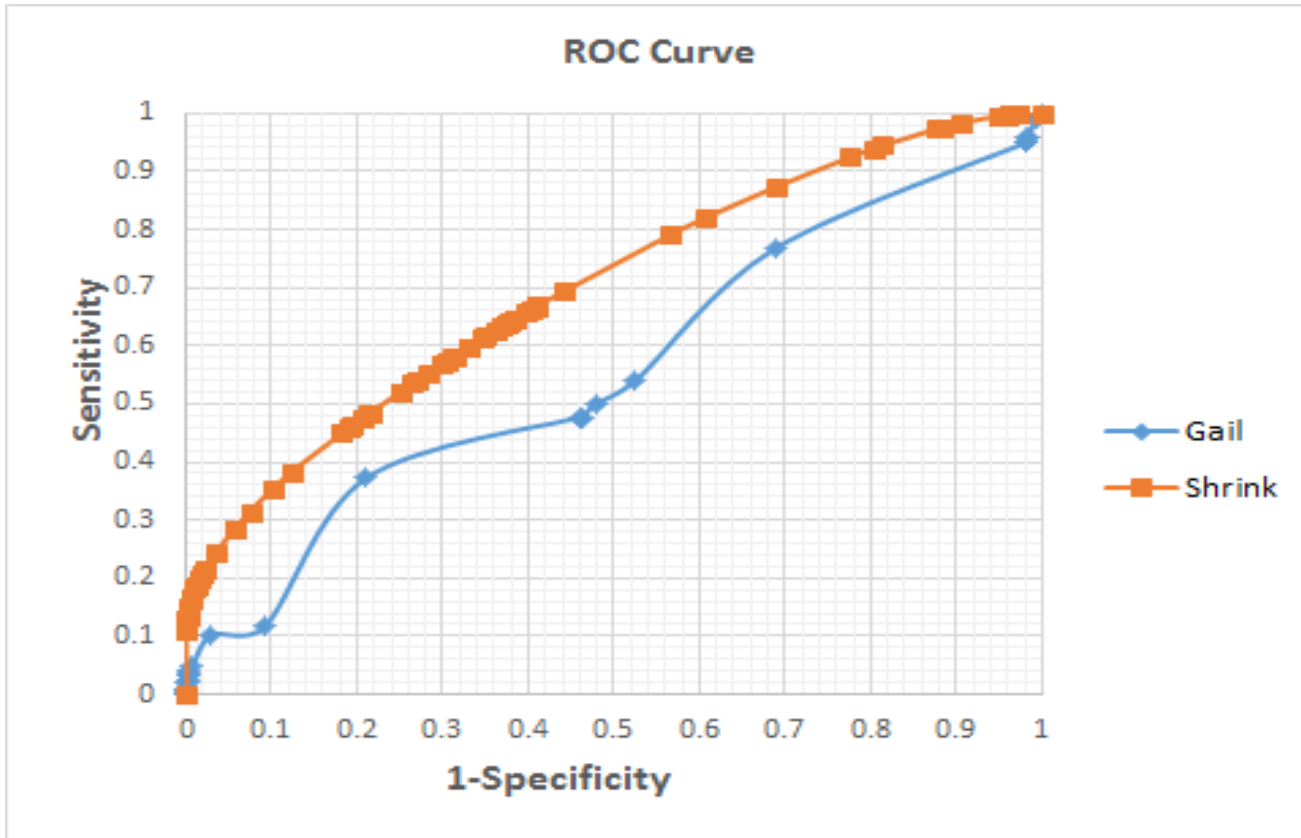
—Test 2 result



The ROC Curve of Shrink and Gail based on 5,000 data

# Experiment

—Test 2 result



The ROC Curve of Shrink and Gail based on 103,679 data

# Conclusion

- We put forward a new model named Shrink that depends on epidemiological factors, which is more adaptive to the populous country like China than those models. The model utilizes the similarities among epidemiological factors to construct a breast cancer high-risk group, the members of which have high similarity with breast cancer patients.
- The model is meaningful for the prevention and control of breast cancer. And it can be generalized to other countries and regions.

**THANK YOU!**