風險評估的時機決定多基因風險評分在乳癌篩檢中的增額價值

Timing of Risk Assessment Determines the Incremental Value of Polygenic Risk Scores in Breast Cancer Screening

Internship Institution: Taipei Veterans General Hospital

Division of Breast Surgery and Comprehensive Breast Health Center

Students: Ting-Chen Cho, Chang-Wei Huang Advisors: Tzu-Pin Lu, Ph.D., Ping-Hsing Tsai, Ph.D.

Institutional Supervisor: Chi-Cheng Huang, MD, Ph.D.

1 Background

The cohort was screened under Taiwan's earlier national criteria: biennial mammography for women aged 45–69, and for women aged 40–44 with a first- or second-degree family history of breast cancer. Coverage was extended in January 2025 to all women aged 40–74 regardless of family history. Mammography remains the primary screening tool, but its diagnostic impression becomes available only after the image has been read by radiologists; identifying high-risk women earlier would augment breast cancer prevention by further incorporating genetic factors such as PRS.

Polygenic risk scores (PRS) aggregate weighted risk effects across many genetic variants to quantify individual susceptibility. All four scores used here were developed in European-ancestry populations and lose performance in other groups: the 313-variant score achieves an AUROC of only 0.571 in women of African ancestry (Du et al., JNCI 2021), and its per-SD odds ratio falls from 1.46 in European-ancestry women to 1.19 in African-ancestry women (Liu et al., JAMA Netw Open 2021). Their value in a Taiwanese population requires local validation.

2 Objective

This study evaluated whether four international breast cancer PRS add predictive value to an existing questionnaire-based risk model from the mammography screening service, distinguishing two clinical settings:

- Pre-report model — questionnaire, demographic data, and mammographic breast density
- Post-report model — the above plus the mammographic diagnostic impression (BI-RADS classification)

The only difference between them is whether the mammography reading is already known; the contrast evaluates the value of PRS within the pre-report workflow.

3 Materials and Methods

Data and sample. Genotyping used the Taiwan Precision Medicine SNP array (TWB2.0). Screening questionnaires were linked to genome-wide imputed genotypes (6,937 individuals) by unique identifier; after intersection and cleaning, 5,569 women remained (421 cases, 5,148 controls; 7.56%).

Polygenic risk scores. Four scores from the PGS Catalog (Table 1) span p-value thresholding and PRS-CS (continuous shrinkage) construction, ranging from 77 to 1.12 million variants. Each was normalized by the number of alleles counted to make scores of different scale comparable (Figure 1).

Analysis. Models were fitted by logistic regression with class weighting, since cases were only 7.56%. All models used 5-fold cross-validation, with per-fold standardization to prevent leakage. Discrimination was compared using DeLong's paired AUC test.

Table 1. The four polygenic risk scores

Score	Variants	Construction method	Coverage
PGS000001	77	Significance threshold	96.1%
PGS000873	143	Significance threshold	90.9%
PGS000004	313	Threshold + stepwise	81.8%
PGS000508	1,120,410	PRS-CS	99.99%

Coverage: matched score variants as a share of those available after liftover to GRCh38.

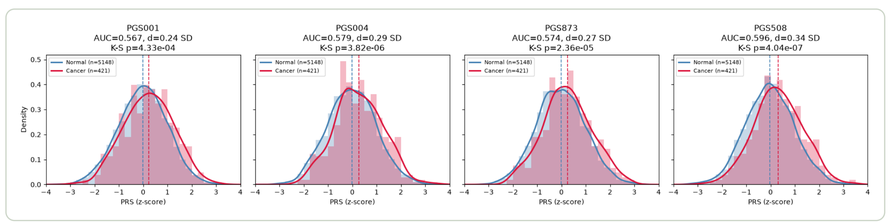

Figure 1. Distribution of the four polygenic risk scores in cases and controls. Labels are shortened PGS Catalog IDs.

4 Results

Table 2. Incremental predictive value in two settings

Model	Pre AUC	ΔAUC	p	Post AUC	ΔAUC	p
Questionnaire only	0.6677	—	—	0.7958	—	—
+ PGS000001	0.6801	+0.0124	0.027	0.7987	+0.0029	0.164
+ PGS000873	0.6831	+0.0154	0.015	0.7988	+0.0030	0.262
+ PGS000004	0.6914	+0.0237	0.001	0.8010	+0.0052	0.108
+ PGS000508	0.6924	+0.0247	<0.001	0.7993	+0.0035	0.254

Incremental value is confined to the pre-report setting. All four scores gave statistically significant gains in discrimination before the reading was known; once the diagnostic impression entered the model, none of these gains remained significant (Table 2). At the variable level, the adjusted per-SD OR for PRS fell from 1.23–1.33 before reporting to 1.15–1.21 after, the latter still nominally significant.

Risk stratification is clinically meaningful. Comparing highest and lowest deciles, breast cancer prevalence rose from 4.5–5.4% to 11.7–12.6%, with odds ratios of 2.32–3.06; all four scores were significant (Figure 2).

Variant coverage does not explain the decline. Coverage ranged from 81.8% to 99.99% but did not track performance — the lowest-coverage score ranked second, ruling out unmatched variants. All four weight sets were derived in European-ancestry data (PGS000508 with a European LD reference panel), so the linkage disequilibrium they assume may not hold here; this was not tested.

The estimate agrees with an external cohort. PGS000508's effect size here (OR 1.415 per SD) is close to the 1.498 reported in the Taichung Veterans General Hospital TPMI cohort — an independent cross-institutional replication.

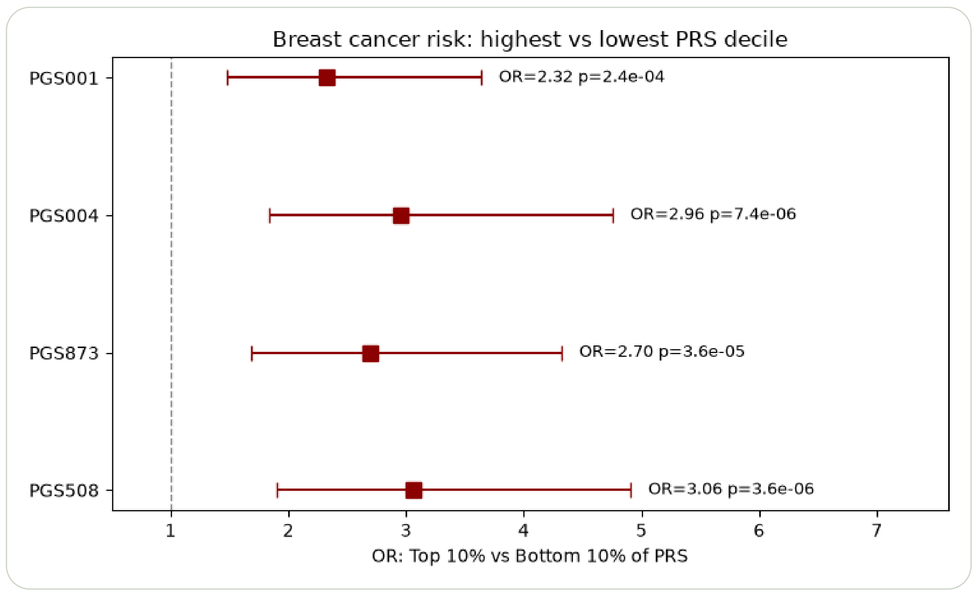

Figure 2. Breast cancer risk, highest versus lowest polygenic risk score decile. Same label convention as Figure 1.

5 Conclusion

All four PRS add significant predictive value before the mammographic reading is available; once the reading enters the model, that increment disappears. The value of PRS lies clearly before the mammography report — suitable for risk stratification when imaging is complete but the report is pending (for example, whether to add ultrasound or shorten the follow-up interval), rather than replacing image interpretation itself.

Limitations. The cohort included 133 women with a personal history of breast cancer, among whom prevalence reached 57.9%; this may inflate the baseline AUC of both models. All results come from internal cross-validation in a single cohort, with no external validation, calibration, or net-benefit assessment. The cohort also predates the 2025 expansion of the program to ages 40–74.

References: Mavaddat N, J Natl Cancer Inst 2015 · Mavaddat N, Am J Hum Genet 2019 · Brentnall AR, Int J Cancer 2019 · Fritsche LG, Am J Hum Genet 2020 · Du Z, J Natl Cancer Inst 2021;113:1168 · Liu C, JAMA Netw Open 2021;4:e2119084 · Lambert SA, Nat Genet 2021 (PGS Catalog)