

Genome-based benchmarking of 177 different breast cancer polygenic risk scores

Ertugrul Eyol

Universitätsklinikum Tübingen

Ulrike Faust

Universitätsklinikum Tübingen

Marc Sturm

Universitätsklinikum Tübingen

Antje Stäbler

Universitätsklinikum Tübingen

Silja Gauß

Universitätsklinikum Tübingen

Kristin Bosse

Universitätsklinikum Tübingen

Ines Gruber

Universitätsklinikum Tübingen

Andreas Hartkopf

Universitätsklinikum Tübingen

German Demidov

Heidelberg University

Nicolas Casadei

Universitätsklinikum Tübingen

Kira Mehnert

Universitätsklinikum Tübingen

Janna Witt

Universitätsklinikum Tübingen

Stephan Ossowski

Universitätsklinikum Tübingen

Tobias Haack

Universitätsklinikum Tübingen

Olaf Riess

Universitätsklinikum Tübingen

Christopher Schroeder

Universitätsklinikum Tübingen

Dennis Witt

dennis.witt@med.uni-tuebingen.de

Universitätsklinikum Tübingen

Research Article

Keywords: polygenic risk score, breast cancer, short-read genome sequencing, HBOC, PGS Catalog, risk stratification, benchmarking, case-control study, clinical genomics

Posted Date: October 11th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10655789/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. GD has received research funding from the European Union's Horizon 2020 research and innovation programme (grant agreement 779257) and consulting fees from Diagenius BV, Amsterdam, the Netherlands, both unrelated to the present work. NC has received research funding from the Deutsche Forschungsgemeinschaft (INST 37/1049-1) and payment for lectures at the University of Reutlingen. SO has received research funding from Illumina Inc. and honoraria from Illumina Inc. for a presentation at the GfH conference 2023, as well as support for attending meetings from Illumina Inc. and ONT in 2023. TBH has received research funding from Illumina Inc. OR has received research funding from Illumina Inc., the German Research Foundation, the European Union, and the German Ministry of Health, and is the representative of the German Ministry of Health in the European +MG project (unpaid). CS has received research funding from Illumina Inc. and research grants from BMS Stiftung Immunonkologie and Westdeutsche Studiengruppe GmbH, outside the submitted work. The remaining authors declare that they have no competing interests.

Abstract

Background More than 170 breast cancer polygenic risk scores (PRS) are publicly available, yet no systematic comparison guides score selection for clinical use. Guidance is especially needed in hereditary breast and ovarian cancer (HBOC) diagnostics, where many patients meeting clinical criteria carry no pathogenic variant in established risk genes, therefore lacking actionable risk stratification.

Methods Scores, including the genome sequencing-optimized BRIDGES PRS, were benchmarked in a monocentric, clinically annotated HBOC cohort of 775 female breast cancer patients (680 non-informative, without a (likely) pathogenic variant in 13 established HBOC genes; 95 informative carriers) and 1,313 cancer-free female controls of European ancestry, all sequenced by clinical-grade GS. PRS were calculated with standardized pipelines and assessed by area under the receiver operating characteristic curve (AUC) and Cohen's d. Benchmarks were stratified by informative status, age at diagnosis, molecular subtype, and breast density. SNP overlap among the top-performing scores was quantified using pairwise Jaccard similarity.

Results The best-performing PRS yielded an AUC of 0.656 (95% CI 0.631–0.681; Cohen's d = 0.576; OR 1.77 per standard deviation) in HBOC patients without informative testing. 40.1% of these patients fell in the top PRS quintile versus 20% of controls, and 7.2% in the bottom quintile (top versus bottom OR 5.57, 95% CI 3.93–7.90). The top 20 scores formed a performance plateau (AUC 0.643–0.656) despite ranging from 3,820 to 3.7 million SNPs, and were dominated by near-identical variant composition, though compact scores matched them from far fewer or largely distinct variants. Performance varied by molecular subtype, strongest for hormone receptor-positive tumors (Luminal B-like HER2-positive, AUC 0.689; Luminal A-like, 0.677) and markedly reduced for triple-negative breast cancer (0.568).

Conclusions Independently developed breast cancer PRS have converged on comparable performance and a largely shared set of variants. For laboratories already performing GS, computing a PRS adds negligible marginal cost while separating patients without a monogenic finding into groups whose odds of breast cancer differ more than five-fold. PRS thus helps close a major diagnostic gap and supports integration into HBOC counseling, showing that the choice of score is not a barrier to clinical implementation.

Background

Breast cancer is the most diagnosed cancer in women worldwide and a leading cause of cancer-related mortality [1]. Age-based population screening has improved detection and reduced breast cancer mortality [2], and risk-based screening strategies are being evaluated as a means of targeting screening more efficiently [3]. Screening entry is currently determined mainly by age and other clinical information. Such information should be complemented by approaches that integrate genetic and other individual risk factors to guide screening frequency, modality, and preventive interventions [4], including the use of germline genetic testing in appropriately selected patients [5].

While the identification of highly penetrant monogenic variants (e.g., in *BRCA1* and *BRCA2*) has revolutionized the management of hereditary breast and ovarian cancer (HBOC) families, a significant clinical dilemma remains: HBOC is a referral category defined by clinical criteria (see Methods), and the majority of patients who meet those criteria do not harbor pathogenic variants in established core genes. For these genetically non-informative patients, actionable guidance is severely limited, highlighting a critical diagnostic gap.

Genome-wide association studies (GWAS) have identified several hundred common single-nucleotide polymorphisms (SNPs) associated with breast cancer risk [6, 7]. However, as individual variants confer only modest effect sizes, they have not proven sufficient for clinical risk stratification. These limitations were targeted by the emerging field of polygenic risk scores (PRS), which aggregate tens to millions of SNPs to stratify disease risk at the population level [8, 9].

The landmark 313-SNP PRS developed by Mavaddat et al. demonstrated an odds ratio of 1.61 per standard deviation and an AUC of 0.630 in independent validation. Based on UK incidence and mortality data, women in the highest centile of this score carried a lifetime breast cancer risk of 32.6%, meeting the UK National Institute for Health and Care Excellence (NICE) definition of high risk and 35% of all breast cancers would be expected to arise in the highest PRS quintile compared with 9% in the lowest [10].

PRS are increasingly integrated into clinical breast cancer risk stratification across multiple healthcare settings. In a German cohort of women counselled for suspected HBOC, incorporating PRS into established risk prediction models altered prevention recommendations in 12.0% of cases [11]. PRS have been integrated at scale into a national randomized personalized screening trial in the United States [12], and the underlying risk models have been validated prospectively in a Dutch cohort [13], reflecting a broader international movement toward PRS-informed, risk-adapted screening strategies.

However, this rapid proliferation has created a new translational bottleneck. These scores differ widely in training populations, variant selection strategies, and SNP counts. This volume has created a "score selection barrier" that complicates the choice of a specific PRS for clinical use. The original 313-SNP PRS and most other scores were developed and validated on SNP array data from varied sources [10, 14], and do not transfer to next-generation sequencing (NGS)-based genotyping without dedicated quality assurance: across five centres of the German Consortium for Hereditary Breast and Ovarian Cancer (GC-HBOC), between 11 and 23 PRS loci showed deviating allele frequencies affecting the final risk prediction [15].

The BRIDGES PRS addressed this limitation by adapting the 313-SNP score for targeted NGS, yielding a 306-SNP panel that retains 295 of the original variants [16]. This adaptation was specifically designed to facilitate clinical implementation alongside monogenic HBOC gene testing within the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) multifactorial risk prediction framework [16]. The BRIDGES PRS thus represents the only publicly available breast cancer PRS explicitly optimized for sequencing-based clinical workflows, making it a natural reference standard.

However, despite advancements in PRS generation, clinical translation has been limited by the lack of systematic, comparable evidence on external predictive accuracy under real-world conditions. Most published PRS have been validated in their derivation cohorts or population-based biobanks, with limited assessment of transferability to independent clinical cohorts [9]. The ClinGen/PGS Catalog Polygenic Risk Score Reporting Standards note that relatively few studies have externally benchmarked the performance of multiple scores, and that no uniformly agreed best practices or widely adopted standards exist that are sufficiently tailored to assess the clinical readiness of a PRS [9].

Recently, Tanha et al. compared the predictive performance of two established PRS, including PGS313, and five newly developed genome-wide PRS across three large population-based cohorts from the UK and Australia, reporting modest improvements of genome-wide approaches over PGS313 in European-ancestry groups [17]. This study was an important advance in comparative PRS evaluation. Yet it was limited to seven scores and relied on SNP array genotyping with imputation rather than sequencing-based variant calling. It was also conducted in unselected population cohorts, without clinical annotation for monogenic variant status or tumor characteristics. A comprehensive head-to-head comparison of all publicly available breast cancer PRS using clinical-grade genome sequencing data has not yet been performed.

Building on the infrastructure of our earlier data [18], we present, to our knowledge, the first systematic benchmarking of all publicly available breast cancer PRS in a clinical-grade genome sequencing cohort in an expanded, retrospective case–control cohort with detailed clinical annotation. Using uniform clinical-grade short genome sequencing (GS) data and standardized analytical pipelines, we evaluated all breast cancer PRS available, at the time of research, in the PGS Catalog as well as the BRIDGES PRS. We compared score performance with respect to discriminative ability, robustness across age strata, and capacity for clinically relevant risk stratification, particularly examining the diagnostic yield among non-informative cases. We further examined the overlap in constituent SNPs across the top-ranked scores. This work contributes to the ongoing validation efforts and provides an evidence base for the selection of a suitable PRS for integration into clinical practice.

Methods

Study design and setting

We present a retrospective case–control study conducted at the Institute of Medical Genetics and Applied Genomics (IMGAG), University Hospital Tübingen, Germany. The study was approved by the local institutional review board (ethics committee of the University of Tübingen, project number 133/2021B01). Patients with additional translational sequencing provided consent to participate in the Ge-Med project study (ClinicalTrials.gov: NCT04760522). All participants provided written informed consent for genetic testing.

Case cohort

The case cohort was recruited from the outpatient clinic for suspected familial cancer predisposition at the University Hospital Tübingen, using the recruitment framework and eligibility criteria laid out by our previous work [18]. In brief, all cases were women diagnosed with breast cancer who concurrently met the clinical criteria for HBOC genetic testing applied in that study, which follow the German interdisciplinary S3 guideline for the early detection, diagnosis, therapy and follow-up of breast cancer and are listed in full in its supplementary tables [18]; counselling and risk-adapted care followed the consensus recommendations of the German Consortium for Hereditary Breast and Ovarian Cancer [19].

Recruitment took place between April 2019 and March 2023 [18]. The final case cohort comprised 775 female breast cancer patients of European genetic ancestry (Fig. 1B). Thirteen established HBOC-associated genes were evaluated as part of routine diagnostics: *ATM*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *PALB2*, *PTEN*, *RAD51C*, *RAD51D*, *STK11*, *TP53*.

Test informativeness

Monogenic variant status for each patient was ascertained through the megSAP diagnostic pipeline [20]. A diagnostic test was considered informative when a (likely) pathogenic variant was identified in an established HBOC gene, and non-informative when no such variant was found (680 non-informative, 95 informative cases); PRS results were stratified accordingly.

Ductal carcinoma in situ (DCIS)

Pure DCIS cases without a documented invasive second tumor were excluded from the analyses to align with the invasive-only validation cohorts of the underlying PRS. Cases with carcinoma in situ (tumor–node–metastasis [TNM] category Tis) as the first tumor but an invasive second tumor were retained. This step removed 48 pure DCIS cases; 2 Tis-first cases with invasive second tumor were retained.

Control cohort

To establish a robust baseline distribution for polygenic risk scores, we curated a control cohort of female individuals sequenced at the same center as the case cohort. To minimize technical batch effects and population stratification, all control samples underwent GS using identical protocols (DNA extraction from ethylenediaminetetraacetic acid (EDTA) blood) and were analyzed with the local megSAP diagnostic pipeline at the IMGAG [20].

Phenotypic orthogonality to breast cancer was enforced by strictly excluding individuals referred for hereditary cancer syndromes (e.g. HBOC) as well as any individual with a documented breast cancer diagnosis. The final control cohort comprised 1,313 female individuals of European ancestry referred for

rare diseases biologically distinct from carcinoma. Many were referred for neurological conditions, including hereditary ataxia (n = 131), hereditary spastic paraplegia (n = 101), and polyneuropathy (n = 52). An additional 2 individuals were referred for cone-rod dystrophy. The remaining controls were referred for singleton genome sequencing (n = 821) or contributed as parental samples in diagnostic trio analyses (n = 206) for a heterogeneous range of rare disease indications. Individual referral diagnoses were not systematically recorded for these subgroups; however, all individuals referred for hereditary cancer syndromes were excluded a priori, and none had a documented diagnosis of breast cancer. This control cohort represents therefore a substantial expansion of the control cohort used by our previous publications [18], enabling more precise z-score normalization and improved statistical power.

Genome sequencing

Clinical short-read genome sequencing was performed at the IMGAG, University Hospital Tübingen, under Deutsche Akkreditierungsstelle (DAkkS) accreditation according to DIN EN ISO 15189:2014 (accreditation annex D-ML-13130-04-00). The sequencing and bioinformatic pipeline has been established for routine clinical diagnostics across multiple disease groups at our institute and has been described in detail previously [20]. Briefly, genomic DNA was extracted from whole blood (FlexiGene DNA kit, Qiagen), processed using the TruSeq PCR-Free Library Prep kit (Illumina), and sequenced predominantly on a NovaSeq 6000 system and in part on a NovaSeq X Plus system (Illumina). Downstream data processing, including variant calling and variant interpretation, was performed using the megSAP pipeline [20].

Genetic ancestry estimation

Genetic ancestry was estimated from GS data using the SampleAncestry tool [21] from the open-source ngs-bits software suite [22], which is integrated within the megSAP diagnostic pipeline [20]. Instead of a genome-wide principal component analysis, this method utilizes a targeted correlation approach based on population-specific variants. For each of the four modelled 1000 Genomes Project super-populations (African [AFR], European [EUR], South Asian [SAS], and East Asian [EAS]), the 1,000 most informative exonic single-nucleotide polymorphisms (SNPs) were evaluated. Participants were assigned to these super-populations by correlating their sample variants against these predefined SNP sets. Population assignment was determined using score thresholds (median and median absolute deviation) established and internally benchmarked on the 1000 Genomes Project dataset, where this approach previously demonstrated a 99.8% assignment accuracy.

Clinical annotation

Molecular subtypes of breast cancer were classified according to the St. Gallen International Expert Consensus [23] using immunohistochemical surrogate definitions based on estrogen receptor, progesterone receptor, HER2, and Ki-67 status. Mammographic breast density was categorized

according to the American College of Radiology (ACR) density classification from routine clinical imaging [24]. Height and weight were ascertained by standardized patient questionnaires at the time of study enrolment.

Polygenic risk score selection

We systematically retrieved all breast cancer PRS deposited in the PGS Catalog [25] that were annotated for the trait "breast carcinoma" (EFO_0000305) or any of its child traits [14, 26]. In addition, we included the BRIDGES PRS, which was adapted from the Mavaddat 313-SNP PRS specifically for next-generation sequencing [16]. A total of 179 distinct PRS were retrieved (178 from the PGS Catalog plus the BRIDGES 306-SNP PRS). Two PGS Catalog scores were excluded because their published definitions lie outside the additive per-allele scoring framework used here and by standard PGS calculators: PGS004299 [27] assigns separate weights to each genotype (dosage_0/1/2_weight) under a non-additive model, and PGS000153 [28] is a population-standardized multiplicative genetic risk score defined as the product of per-allele odds ratios rather than an additive sum; neither is reproducible by additive computation, so both were excluded, leaving 177 PRS for analysis. The individual scores, their PGS identifiers, number of constituent SNPs, development populations, and original publications are listed in Additional file 1: Table S1.

PRS calculation

Polygenic risk score definitions for breast cancer were obtained from the PGS Catalog [14]. We utilized the harmonized dataset versions mapped to the GRCh38 assembly. PRS calculation was performed using an independently developed Python pipeline. Given the variant-only nature of standard GS variant call format (VCF) files [29], we adopted a homozygous reference assumption strategy: any variant present in a PRS definition but absent from a sample's VCF was inferred to be homozygous for the reference allele. The same assumption has been applied when computing PRS from targeted sequencing data, where it was used for the variants not covered by the assay [30]. For each PRS variant present in the VCF, allele alignment between the PGS Catalog definition and the sample's call was tested across all four possible allele and strand orientations (effect allele matching the alternate or the reference base, each on the forward strand or its reverse complement), following the general variant-matching logic used by the PGS Catalog Calculator, which handles multi-allelic variants, strand flips and effect-allele-only scores [26].

Palindromic A/T and C/G SNPs were not excluded a priori; because the reference base was read from the GRCh38 forward strand and each PGS was hmPOS-harmonized to that same strand. A direct same-strand match of the reference base to the effect or other allele is unambiguous even for palindromes, and such variants were retained and scored. A palindromic SNP was excluded only when no direct orientation matched and resolution would have required a reverse-complement interpretation, i.e. the genuinely strand-ambiguous case. Additionally, variants were excluded if they contained non-standard allele codes or mapped to non-canonical chromosomes.

Raw PRS values were converted to z-scores independently for each score, using the mean and standard deviation derived from the control cohort. This per-score normalization enables direct comparison of effect sizes across scores with different numbers of constituent variants and different weight scales, while preserving score-specific distributional properties. For each PRS, quality metrics including total SNPs, directly genotyped SNPs, palindromic exclusions, allele mismatches, and strand flips were recorded (Additional file 1: Table S1)

Validation against the PGS Catalog reference calculator

We cross-validated our Python based workflow against `pgsc_calc` v2.3.0 [26], the open-source workflow of the PGS Catalog, which performs score computation using PLINK2 [31] within a Nextflow [32] workflow built to nf-core community standards [33]. Validation used ten GS samples of diverse genetic ancestry against nine deliberately chosen edge-case scores covering the full method space: size, from PGS000521 with 9 SNPs to PGS004025 with ~ 1.04M; discrimination, from three of the top-ranked additive scores (PGS005346, PGS004025, PGS004040) to the mid-range PGS004170 [34]; indel content, in the Mavaddat landmarks PGS000004 and PGS000007; and study design, in a case-case score built to separate triple-negative from luminal disease (PGS004890 [35]) and the sequencing-adapted BRIDGES 306-SNP score. PGS Catalog scoring files were retrieved automatically, BRIDGES as a custom scoring file. To match the pipeline's retention of palindromic and multiallelic variants, `pgsc_calc` was run with these variants kept (`--keep_ambiguous`, `--keep_multiallelic`); all other parameters were left at their defaults. Because `pgsc_calc` scores from called genotypes alone, the comparison was performed on dosage-based scores; the homozygous reference contribution has no counterpart in `pgsc_calc` and lies outside its scope. The comparison therefore establishes agreement in variant matching, allele and strand orientation, and weighted summation. On this basis the in-house pipeline and `pgsc_calc` agreed on every one of the 90 sample × score combinations ($r = 1.000$; all scores agreed to within < 0.001 SD of the score distribution).

Statistical analysis

Performance. Performance was quantified using the area under the receiver operating characteristic curve (AUC) with 95% confidence intervals obtained by DeLong's method [36]. Effect sizes were additionally expressed as Cohen's d [37] and odds ratios per standard deviation from logistic regression.

Age-stratified analyses. Analyses were conducted for age at diagnosis groups (< 40 years, 40–50 years, ≥ 51 years). Cases were assigned to these groups by age at diagnosis and controls by age at sequencing. Age-matched comparisons restrict both cases and controls to the same age band; unmatched comparisons retain all 1,313 controls. These age groups reflect the established definition of early-onset breast cancer (< 40 years; [38]); the upper boundary approximates the mean age at natural menopause of 49.3 years reported among postmenopausal controls in a 117-study individual-participant meta-analysis [39]. Subgroup analyses were performed by genetic ancestry, contralateral breast cancer status, molecular subtype, mammographic breast density category, BMI category, height, and weight. As a sensitivity analysis, logistic regression adjusting for age at diagnosis (cases) or age at sequencing

(controls) was performed to confirm that PRS stratification was independent of the age difference between cases and controls (median 50 vs. 38 years). The age-adjusted odds ratio per standard deviation and quintile odds ratios were compared to the unadjusted estimates.

Population restriction. The analysis was restricted to female participants of European genetic ancestry [40]; the 152 non-European participants were excluded.

Risk reclassification. PRS-defined risk quintiles were computed based on the control cohort distribution. Odds ratios per quintile were estimated using logistic regression with the lowest quintile as reference. To enable comparison with risk-based screening studies, we additionally applied the risk categorisation of Mars et al. [41], defining high risk above the 90th percentile of the control distribution, low risk below the 10th, and the intermediate 10th–90th percentile as the reference category.

SNP overlap and genomic convergence. For the top 20 PRS by AUC, constituent SNP sets were defined at the level of harmonized GRCh38 chromosomal position, and pairwise similarity was quantified using the Jaccard index [42]. Variants were matched by exact position; variants in high linkage disequilibrium at different positions were not merged, so all overlap measures reported here represent lower bounds. Scores were grouped by complete-linkage hierarchical clustering of the pairwise Jaccard distances. Because the Jaccard index is sensitive to disparities in set size, overlap between the compact distinct scores and the dominant high-overlap group of scores was additionally expressed as the proportion of each score's variants contained within the union of that group's SNP sets. We defined universal variants as those present in all 20 top-ranked PRS (181 variants), representing the most consistently selected loci across PRS development methods, and annotated them with GRCh38 cytogenetic band positions and proximity (± 500 kb) to genome-wide significant breast cancer lead variants ($P < 5 \times 10^{-8}$) for the trait breast carcinoma (EFO_0000305) in the National Human Genome Research Institute–European Bioinformatics Institute (NHGRI-EBI) GWAS Catalog [43]. Position-level overlap is expected to partly reflect the shared GWAS summary statistics, variant catalogues, and linkage-disequilibrium references underlying many published scores. We therefore interpret the overlap structure, a dominant high-overlap cluster alongside compact scores that reach comparable discriminative accuracy with largely non-overlapping variants, descriptively rather than as a test against a random-genome null.

Software. All statistical analyses were performed using R (R Foundation for Statistical Computing, Vienna, Austria) with the packages tidyverse, pROC, gtsummary, pheatmap, ggpubr, cowplot, and scales. The PRS calculation pipeline was implemented in Python [44]. A large language model (Claude, Anthropic) was used to assist with code editing during pipeline development. All outputs were reviewed and validated.

Results

Cohort characteristics

The final study cohort comprised 2,088 female participants of European genetic ancestry, including 775 female breast cancer cases and 1,313 cancer-free female controls (Additional file 1: Table S2). Among cases, 680 (87.7%) had no likely pathogenic or pathogenic variant detected in any of the evaluated HBOC-associated genes (non-informative cases), while 95 (12.3%) harbored a (likely) pathogenic monogenic variant in at least one established HBOC gene (informative cases).

The median age at diagnosis among all cases was 50 years (IQR 42–58; range 24–90). Non-carrier cases had a median age at diagnosis of 50 years (IQR 43–59; range 24–90), while informative cases were diagnosed at a median age of 49 years (IQR 40–56; range 25–82). Controls had a median age at sequencing of 38 years (IQR 20–54; range 0–90). Age group distribution differed substantially between cases and controls: 700 controls (53.3%) were aged < 40 years compared to 153 cases (19.7%), while 364 cases (47%) were aged ≥ 51 years compared to 407 controls (31%).

All 2,088 participants included in the primary analysis were assigned to European genetic ancestry (median European ancestry score 0.330, IQR 0.312–0.348; range 0.221–0.446). This is a correlation-based score against population-specific SNP sets, not an admixture proportion, and the four population scores are not constrained to sum to one; the observed median closely matches the European reference median of 0.325 from the 1000 Genomes Project, confirming correct ancestry assignment. 152 participants were excluded due to non-European ancestry.

Molecular subtypes according to the St. Gallen classification were available for 567 of 775 cases (73.2%); the remaining 208 (26.8%) could not be classified due to incomplete immunohistochemical data. Among the 567 classifiable tumors, the distribution was: Luminal B-like HER2-negative 176 (31%), Luminal A-like 138 (24.3%), triple-negative 115 (20.3%), Luminal B-like HER2-positive 89 (15.7%), HER2-enriched (non-luminal) 38 (6.7%), and HR-discordant (ER–/PR+) 11 (1.9%). Contralateral breast cancer was documented in 55 cases (7.1%), and concurrent breast and ovarian cancer in 16 cases (2.1%).

Mammographic breast density (ACR classification) was available for 517 of 775 cases (66.7%): ACR C 245 (47.4%), B 144 (27.9%), D 108 (20.9%), and A 20 (3.9%). Dense breasts (ACR C/D) were present in 353 cases (68.3%). Among cases with available anthropometric data (n = 607), median BMI was 24.6 kg/m² (IQR 22.0–28.3; range 16.9–51.2), median height was 166.0 cm (IQR 162–170; range 149–186), and median weight was 68 kg (IQR 60–78; range 45–148).

Among the 95 informative cases, the most frequently affected genes were *BRCA1* (n = 32, 33.7%), *BRCA2* (n = 22, 23.2%), *CHEK2* (n = 18, 18.9%), *ATM* (n = 10, 10.5%), *PALB2* (n = 5, 5.3%), *BARD1* (n = 3, 3.2%), *TP53* (n = 3, 3.2%), and *RAD51C* (n = 2, 2.1%). One *BRCA2* carrier additionally harbored a pathogenic *CHEK2* variant; gene frequencies reflect the primary causal variant assigned per carrier.

SNP coverage and technical quality control

A total of 177 PRS were evaluated. The combined input comprised 69,515,852 variants across all scoring files. Following quality control, 69,192,519 variants (99.5%) were retained for analysis. Per-score quality metrics are reported in Additional file 1: Table S1.

Across all 177 PRS, a median of 49.4% of a score's SNPs were present as variant calls in the GS data (IQR 38.9–52.8%; range 16.1–62.1%); the remainder were absent from the VCF and assigned the homozygous reference genotype. Only a small minority of SNPs were lost to quality control, so a median of 100.0% of each score's variants entered the calculation (IQR 99.9–100.0%; range 95.7–100.0%), with 150 of 177 scores (84.7%) above 99% and all 177 above 95%. The BRIDGES PRS comprised 306 SNPs, of which 154 (50.3%) were directly genotyped and 300 of 306 (98.0%) contributed to the final score.

Overall PRS Performance

The primary analysis compared 680 non-informative breast cancer cases to 1,313 cancer-free controls. Informative cases ($n = 95$) were analyzed separately as a subgroup, to test whether polygenic burden is also elevated in carriers of high-penetrance variants. Of 177 evaluated PRS, 173 (97.7%) achieved an AUC significantly exceeding 0.50, 3 showed no significant deviation from chance, and none fell significantly below 0.50. The median AUC across all 177 PRS was 0.601 (IQR 0.574–0.624; range 0.483–0.656); the lowest was PGS004611 (AUC 0.483) (Fig. 2A–C).

The best-performing score was PGS004153 [45] (1,133,268 SNPs; AUC 0.656, 95% CI 0.631–0.681; Cohen's $d = 0.576$; OR per SD 1.77, 95% CI 1.60–1.95), an ensemble score (UK Biobank [UKBB] MultiPRS) that combines multiple individual PRS, closely followed by PGS005346 (AUC 0.655, 95% CI 0.629–0.680; $d = 0.565$) and PGS004025 (AUC 0.654, 95% CI 0.629–0.679; $d = 0.567$), with broadly overlapping confidence intervals, indicating that performance has largely converged across methods (Fig. 2D, Table 1). Three of the highest-ranked scores were ensemble methods (UKBB MultiPRS, PRSmix, PRSmixPlus) that combine multiple individual PRS and are expected to marginally exceed their components; among individually developed scores, the top performers (PGS005346, PGS004025, and PGS004040, all LDpred2-based) spanned a narrow band of AUC 0.654 to 0.655 with overlapping confidence intervals. AUC values among the top 20 PRS ranged from 0.643 to 0.656, with odds ratios per standard deviation spanning 1.66 to 1.77. The top 20 PRS spanned score sizes from 3,820 to 3,666,321 SNPs (Spearman $\rho = 0.529$, $P = 3.76 \times 10^{-14}$ across all 177 PRS).

The BRIDGES 306-SNP PRS ranked 53rd overall (AUC 0.621, 95% CI 0.595–0.647; $d = 0.445$; OR per SD 1.56, 95% CI 1.42–1.71). The original Mavaddat 313-SNP PRS (PGS000004) ranked 45th (AUC 0.624, 95% CI 0.598–0.649; $d = 0.455$).

In a sensitivity analysis adjusting for the case–control age difference, PGS004153 remained independently associated with case status (age-adjusted OR per SD 1.86, 95% CI 1.68–2.07; $P = 8.9 \times 10^{-31}$). All 20 top-ranked PRS remained significant after age adjustment (all $P < 10^{-23}$). Quintile odds ratios were stable (Q5 vs Q1: adjusted OR 7.08, 95% CI 4.82–10.42 vs. unadjusted 5.57, 95% CI 3.93–7.90), confirming that PRS-based risk stratification was independent of the age difference between cases and controls.

In a subgroup analysis restricted to 95 informative cases, the best-performing score was PGS000216 (AUC 0.605). The median AUC was 0.557 (range 0.493–0.605); per-score results for all three case–

control comparisons are given in Additional file 1: Table S3.

Table 1
Top 20 polygenic risk scores ranked by AUC.

Rank	PGS ID	SNPs	AUC (95% CI)	OR per SD (95% CI)	Cohen's d
1	PGS004153	1133268	0.656 (0.631–0.681)	1.77 (1.60–1.95)	0.576
2	PGS005346	961336	0.655 (0.629–0.680)	1.75 (1.59–1.93)	0.565
3	PGS004025	1041298	0.654 (0.629–0.679)	1.75 (1.59–1.94)	0.567
4	PGS004040	1041298	0.654 (0.629–0.679)	1.74 (1.58–1.92)	0.562
5	PGS004737	3666245	0.654 (0.629–0.679)	1.75 (1.59–1.93)	0.563
6	PGS004738	3666321	0.654 (0.629–0.679)	1.75 (1.58–1.93)	0.561
7	PGS004688	1044093	0.652 (0.627–0.677)	1.73 (1.57–1.91)	0.554
8	PGS003999	13086	0.652 (0.627–0.677)	1.75 (1.59–1.94)	0.563
9	PGS000508	1120213	0.651 (0.626–0.676)	1.73 (1.57–1.91)	0.551
10	PGS000335	1079089	0.650 (0.625–0.676)	1.71 (1.55–1.89)	0.543
11	PGS004869	872060	0.650 (0.625–0.675)	1.71 (1.55–1.89)	0.539
12	PGS000507	1118944	0.649 (0.624–0.674)	1.70 (1.55–1.88)	0.538
13	PGS004579	1088163	0.648 (0.623–0.673)	1.70 (1.54–1.88)	0.536
14	PGS004069	869407	0.647 (0.622–0.672)	1.70 (1.54–1.88)	0.538
15	PGS003380	1142448	0.647 (0.621–0.672)	1.69 (1.53–1.86)	0.53
16	PGS004083	1103877	0.647 (0.621–0.672)	1.69 (1.53–1.86)	0.527
17	PGS005347	941918	0.646 (0.620–0.671)	1.70 (1.54–1.87)	0.533
18	PGS004053	869407	0.644 (0.619–0.669)	1.71 (1.55–1.89)	0.54
19	PGS000007	3820	0.644 (0.619–0.670)	1.68 (1.52–1.85)	0.524
20	PGS000499	1118944	0.643 (0.617–0.668)	1.66 (1.51–1.83)	0.509

Non-informative cases (n = 680) versus controls (n = 1,313). Per-score quality control metrics for all 177 evaluated scores are given in Additional file 1: Table S1.

Stratified analyses

All stratified analyses used PGS004153 as the reference score, consistent with its top-ranked performance in the primary analysis. All comparisons were conducted using non-informative cases (n = 680) versus controls (n = 1,313) unless otherwise stated; subgroup counts therefore refer to the non-

informative subset and are smaller than the corresponding totals in Additional file 1: Table S2. AUC values for the top 10 PRS across all clinical subgroups are shown in Additional file 2.

Monogenic variant status. For PGS004153, the AUC was substantially higher among non-informative cases (0.656, 95% CI 0.631–0.681) than among informative cases (0.576, 95% CI 0.516–0.636). The Cohen's d between non-informative cases and controls was 0.576 ($P = 1.8 \times 10^{-32}$), whereas the difference between controls and informative cases was 0.226 ($P = 0.037$). Non-informative cases had significantly higher PRS z-scores than informative cases ($d = 0.345$, $P = 0.002$). Results for BRIDGES were concordant: $d = 0.445$ ($P = 2.7 \times 10^{-20}$) for non-informative versus controls, while the informative versus control comparison did not reach significance ($d = 0.166$, $P = 0.14$) (Fig. 3A, B).

When comparing cases with informative testing directly against non-informative cases, no PRS separated the two groups well enough to be useful at the individual level (best AUC 0.596, 95% CI 0.535–0.657; PGS004153 0.584, 95% CI 0.525–0.643), although the direction was consistent across scores, reflecting the higher polygenic burden of non-informative cases described above.

Age at diagnosis. With controls restricted to the same age band as the cases, PGS004153 accuracy was consistent across age strata in non-informative cases (AUC 0.663 in < 40 years, 0.652 in 40–50 years, and 0.658 in ≥ 51 years; all CIs overlapping), with concordant effect sizes (Cohen's d = 0.578 [< 40 years], 0.585 [40–50 years], and 0.580 [≥ 51 years]; all $P \leq 1.4 \times 10^{-8}$). Figure 3C, D show the same strata against the full control cohort, which is younger than the cases (median 38 vs 50 years), separately for non-informative and informative cases; under that definition mean AUC across the top 20 PRS falls from 0.670 (< 40 years) to 0.664 (40–50 years) and 0.632 (≥ 51 years) in non-informative cases (Fig. 3D, Additional file 1: Table S4). Mean AUC across the top 20 PRS in informative cases was 0.583 (< 40 years), 0.562 (40–50 years) and 0.575 (≥ 51 years) (Fig. 3C, Additional file 1: Table S4). Under age matching, PGS004153 showed no significant discrimination in any informative stratum ($n = 24, 31$ and 40; all $d \leq 0.31$).

Molecular subtypes. For PGS004153, the strongest stratification was registered for hormone receptor-positive subtypes: Luminal B-like HER2-positive (AUC 0.689, 95% CI 0.631–0.747; $n = 81$; $d = 0.690$), Luminal A-like (AUC 0.677, 95% CI 0.628–0.725; $n = 129$; $d = 0.673$), and Luminal B-like HER2-negative (AUC 0.658, 95% CI 0.612–0.703; $n = 149$; $d = 0.582$). Triple-negative breast cancer showed markedly reduced accuracy (AUC 0.568, 95% CI 0.511–0.625; $n = 96$; $d = 0.264$).

Additional clinical subgroups. Stratified analyses by mammographic breast density, BMI, anthropometric measures, contralateral breast cancer, and concurrent ovarian cancer are presented in Additional file 1: Table S5. For PGS004153, AUCs ranged from 0.623 (ACR B density; $n = 120$; $d = 0.458$) to 0.719 (contralateral breast cancer; $n = 44$; $d = 0.848$). PRS discriminative ability was notably high among women with highest breast tissue density values (ACR D: AUC 0.710; $n = 90$; $d = 0.779$). An exploratory PRS x breast density interaction on age at diagnosis was not significant.

Among non-informative cases with available first-degree family history data ($n = 451$ of 680 ; 66.3%), PRS accuracy was maintained regardless of familial breast or ovarian cancer burden. PGS004153 achieved an AUC of 0.644 (95% CI $0.600\text{--}0.688$; $d = 0.552$) in cases without affected first-degree relatives ($n = 171$) and 0.687 (95% CI $0.654\text{--}0.721$; $d = 0.697$) in cases with at least one affected first-degree relative ($n = 280$). A direct comparison between the two case groups showed no significant difference in polygenic burden ($d = 0.12$, 95% CI -0.07 to 0.31 ; $P = 0.24$).

Clinical risk reclassification

For completeness, across the full case series including informative carriers ($n = 775$), 38.7% fell in Q5 compared to 20.0% of controls, and only 8.3% in Q1. A stepwise increase in risk across quintiles was observed, with quintile odds ratios rising from 1.85 (95% CI $1.31\text{--}2.62$) in Q2 to 4.69 (95% CI $3.41\text{--}6.45$; $P = 2.6 \times 10^{-21}$) in Q5 (Fig. 3E, F).

The case-enrichment was more pronounced among non-informative cases ($n = 680$). In this subgroup, 40.1% fell in Q5 and only 7.2% in Q1, yielding a Q5 vs. Q1 odds ratio of 5.57 (95% CI $3.93\text{--}7.90$; $P = 4.9 \times 10^{-22}$), with quintile ORs of 2.13 (Q2), 2.18 (Q3), 3.01 (Q4), and 5.57 (Q5). Under the categorisation of Mars et al. [41], 24.6% of non-informative cases fell above the 90th control percentile against 10% expected (OR 2.70 , 95% CI $2.10\text{--}3.47$; $P = 1.2 \times 10^{-14}$ versus the 10th–90th percentile reference), with a correspondingly reduced proportion below the 10th percentile (OR 0.34 , 95% CI $0.21\text{--}0.54$).

Critically, this PRS-based risk gradient was retained in the clinically most challenging subgroup: non-informative cases with no affected first-degree relative for breast or ovarian cancer ($n = 171$). In this group, 38.0% ($n = 65$) fell in the top PRS quintile compared to 20.0% in controls (OR 2.45 , 95% CI $1.75\text{--}3.43$), and 59.1% in the top two quintiles. Among non-informative cases with ≥ 1 affected first-degree relative ($n = 280$), top-quintile enrichment was higher still at 45.0% (OR 3.27 , 95% CI $2.49\text{--}4.29$ vs controls). PRS therefore identifies a substantial high-risk fraction that are missed by monogenic testing and family-history-based stratification.

Among informative cases ($n = 95$), the quintile distribution was more uniform (Q1: 15.8% , Q2: 14.7% , Q3: 14.7% , Q4: 26.3% , Q5: 28.4%). The Q5 vs. Q1 odds ratio was 1.80 (95% CI $0.94\text{--}3.46$; $P = 0.078$).

PRS redundancy and genomic convergence

Analysis of SNP overlap among the top 20 PRS revealed a structured redundancy pattern. Clustering was dominated by a large cluster of mutually highly similar scores (pairwise Jaccard up to 1.000), comprising 11 scores from six research groups: Monti et al. [45] (4 PRS, LDpred2, PRS-CS and ensemble MultiPRS), Fritsche et al. [46] (3 PRS, PRS-CS), Mars et al. [47] (1 PRS, PRS-CS), Namba et al. [48] (1 PRS, LDpred2), Youssef et al. [49] (1 PRS, PRS-CS), and Hu et al. [50] (1 PRS, SDPR). These scores were developed from overlapping large-scale breast cancer GWAS resources, which explains their near-complete variant overlap despite differing derivation methods.

Three independent smaller clusters emerged outside this dominant cluster: a pair of scores from Tanha et al. [17] (PGS005346, PGS005347; LDpred2 and PRS-CS; Jaccard 0.968), a pair of integrative scores from Truong et al. [51] (PGS004737, PGS004738; PRSmix and PRSmixPlus; Jaccard 1.000), and a trio of MegaPRS scores from Jermy et al. [52] and Monti et al. [45] (PGS004869, PGS004069, PGS004053; Jaccard 0.949–1.000).

Two further scores stood apart from every cluster while still reaching the top performance band: the compact penalized score PGS003999 [45] (13,086 variants; AUC 0.652) and the LASSO-derived Mavaddat score PGS000007 (3,820 variants; AUC 0.644), demonstrating that top-tier discrimination is attainable at a small fraction of the ~ 1-million-variant size. Of these two, PGS000007 shares only a minority of its variants with the dominant cluster: ~81% of its variants lie outside it. The integrative PRSmix scores (PGS004737, PGS004738; AUC 0.654) likewise drew ~ 76% of their much larger variant sets from outside the cluster. This demonstrates that variant sets largely non-overlapping with the dominant cluster can capture breast cancer risk with near-equivalent AUCs. Because overlap was assessed at the level of chromosomal position, non-overlapping variants may nonetheless tag the same underlying signals through linkage disequilibrium; positional distinctness does not imply independent biological signal.

181 variants were universally present in all 20 top-ranked PRS, spanning 22 chromosomes and 142 cytobands (Additional file 1: Table S6). Of these, 81 (45%) mapped within ± 500 kb of a genome-wide significant breast cancer lead variant ($P < 5 \times 10^{-8}$) in the NHGRI-EBI GWAS Catalog [43], including loci near *MAP3K1* (5q11.2) [53], *EBF1* (5q33.3) [43], *PTHLH* (12p11.22) [54], *SLC4A7* (3p24.1) [55], and *ESR1* (6q25.1) [56] (Fig. 4). Chromosome 5 harbored the most universal variants ($n = 19$), followed by chromosome 7 ($n = 18$), chromosome 6 ($n = 17$), and chromosome 8 ($n = 16$) (Fig. 4).

Discussion

We benchmarked all publicly available breast cancer PRS on clinical-grade GS in a large real-world HBOC cohort. The best scores reached an AUC of 0.656 in non-informative cases, and the top 20 fell within a narrow band despite large differences in SNP counts and construction methods. Case-control separation was highest in non-informative cases, the group for whom current HBOC testing provides no genetic risk information.

Overall discriminative ability in context

The best-performing PRS (PGS004153) achieved an AUC of 0.656 (Cohen's $d = 0.576$), exceeding the 0.630 reported by Mavaddat et al. [10]. Relative to our predecessor study, which reported a Cohen's d of 0.476 for BRIDGES at the same center [18], applied like-for-like to all cases BRIDGES yields a somewhat lower effect size here ($d = 0.409$), indicating that cohort expansion alone conferred no advantage. Restricting the analysis to non-informative cases raised the BRIDGES effect size to 0.445 by removing the attenuating effect of high-penetrance monogenic variants, and selecting the best of 177 candidate

scores raised it further to 0.576, though choosing the maximum across many candidates inflates that final increment.

Our analysis evaluates 177 PRS using GS data within the clinical context most relevant for immediate implementation: patients referred for diagnostic HBOC testing who remain genetically non-informative after monogenic analysis. Furthermore, by systematically quantifying SNP overlap across top-performing scores, we provide the first empirical assessment of whether score selection constitutes a genuine barrier to clinical adoption or a perceived one.

The top 20 PRS formed a performance plateau (AUC 0.643–0.656) despite spanning 3,820 to 3.7 million SNPs and diverse construction methods. This suggests that current PRS methodology is approaching a ceiling of discriminative ability from common variants alone in European-ancestry populations. The returns from further methodological refinement are diminishing relative to the gains from clinical implementation.

The OR per SD of 1.77 exceeds the 1.61 reported for PRS313 by Mavaddat et al. [10], but this estimate is the maximum across 177 candidate scores and is therefore inflated by selection. The pre-specified BRIDGES score, which shares 295 of its variants with PRS313, yields 1.56 (95% CI 1.42–1.71) which is statistically indistinguishable from the published value. Mavaddat et al. additionally reported a 3,820-SNP lasso score with OR 1.66 (evaluated here as PGS000007), so the advantage over the best previously published score is marginal. Comparison is further limited by differences in sampling frame (HBOC-enriched referrals vs. population-based studies). Notably, this predictive signal required no input beyond the GS data already generated for monogenic HBOC testing.

Methodological heterogeneity and differences in discriminative ability

The top 20 PRS spanned score sizes from 3,820 to 3,666,321 SNPs yet achieved comparable AUCs. This observation has two implications. First, once the core susceptibility signal is captured, incorporating millions of additional variants yields little incremental gain. Our design cannot distinguish whether these variants carry genuine but negligible effects, represent false-positive associations or simply tag signals already captured by variants in the compact scores through linkage disequilibrium. Second, it demonstrates that PRS constructed using different statistical approaches, including penalized regression, Bayesian shrinkage methods, and pruning-and-thresholding strategies, converge on similar discriminative ability when applied to the same target population, reinforcing the conclusion that the underlying biological signal, rather than the analytical method, is the primary determinant of PRS performance.

Of the 177 evaluated scores, only three scored at or close to an AUC of 0.500, and in each case the score's training objective or development population differs from a European case–control susceptibility comparison. PGS004890 (AUC 0.500) is a case-case GWAS score designed to separate triple-negative from luminal tumors rather than to discriminate cases from controls. PGS004611 (AUC 0.483) is a

polygenic score for tumor aggressiveness, based on the risk-of-recurrence-weighted-on-proliferation (ROR-P) prognostic signature; it was developed in a case-only design to predict breast cancer-specific survival rather than disease occurrence [57]. PGS003758 [58] (AUC 0.493) was trained on a Japanese-ancestry GWAS and is not expected to transfer to a European cohort, whereas its European-GWAS counterpart from the same study, PGS003759, performed as expected (AUC 0.571). No score built for European breast cancer susceptibility performed below chance. These observations illustrate the importance of confirming that a score's training objective and development population match the intended application.

From a clinical infrastructure perspective, this convergence has direct implications for diagnostic strategy. Comprehensive genome sequencing remains resource-intensive, while targeted gene panels remain the more resource-efficient option for routine monogenic HBOC testing. GS nevertheless adds diagnostic value at the monogenic level: in the same cohort it enabled the classification of a *BRCA1* duplication and the detection of a whole-exon *BARD1* inversion and a deep intronic *CHEK2* variant, none of which would have been resolved by conventional panel diagnostics [18].

Compact scores could equally be added to a targeted panel: the 3,820-variant PGS000007 reached an AUC of 0.644, close to the best-performing score, and a variant set of that size needs no separate SNP array. This would, however, fix the score at assay validation, whereas sequencing data can be re-scored as better scores appear. For laboratories already performing clinical GS for monogenic HBOC testing, PRS calculation represents a low-cost additional analysis that exploits data already generated, eliminating the need for a dedicated SNP array. The finding that a compact 306-SNP score (BRIDGES) retains most of the discrimination of scores comprising over 3 million variants further supports this approach: even a minimal set of PRS loci can be reliably extracted from existing GS data without additional laboratory investment.

PRS performance by age

Among non-informative cases, mean AUC across the top 20 scores declined with age at diagnosis: 0.670 before age 40, 0.664 at 40–50 years and 0.632 at ≥ 51 years (Additional file 1: Table S4). This gradient does not survive age matching. The control cohort is younger than the cases (median 38 vs 50 years), so the oldest stratum is compared against the most age-discordant controls; restricting controls to the same age band as the cases gives top-20 mean AUCs of 0.663, 0.639 and 0.658, with the oldest stratum no longer the weakest. We therefore do not interpret the unadjusted gradient as age-dependent attenuation of polygenic effect, in contrast to the attenuation reported for *BRCA1* and *BRCA2* carriers [59].

PRS performance is maintained across all age strata examined, including the pre- and peri-menopausal window. Because this is the period during which screening decisions are most consequential, the finding has direct clinical implications for risk-adapted surveillance. In this window, the difference between high and low polygenic risk may determine whether supplementary MRI screening or modified screening

intervals are warranted, particularly for non-informative patients in whom family history suggests elevated risk but no monogenic cause has been identified. Because discrimination was stable across age strata once controls were age-matched, PRS-based risk modification can be applied at any point, including before the onset of population-based mammography screening.

PRS performance by breast density

PRS accuracy was notably high among women with highest breast tissue density values (ACR D: AUC 0.710). Because high mammographic breast density actively masks tumors and severely limits the sensitivity of standard imaging, a genetic stratification tool that maintains, and even exceeds, its predictive power in this specific subgroup may provide additional clinical value. This is the subgroup for which supplemental screening modalities such as MRI are considered, precisely because mammographic sensitivity is reduced [60]; a genetic stratifier that retains discrimination in this setting could therefore help identify which of these women to prioritise.

A conceptual limitation arises from the partial overlap between PRS-based and density-based risk estimation. Established risk prediction tools such as CanRisk [61] already incorporate mammographic density as an independent risk factor. A subset of the susceptibility SNPs in common PRS are themselves associated with mammographic density [62]. Incorporating both PRS and breast density into a risk model may therefore double-count the same biological signal. This could overestimate absolute risk in high-density, high-PRS individuals. Empirically, the incremental effects of density and PRS in population screening cohorts have been found to be largely complementary rather than redundant [63], and BOADICEA assumes multiplicative joint effects of the PRS and the non-genetic risk factors it incorporates [64]. The residual overlap therefore warrants careful consideration when applying PRS-based triage in clinical settings where density information is routinely available.

PRS in monogenic variant non-carriers

The most urgent unmet need in HBOC diagnostics lies with patients who meet familial testing criteria but test negative for monogenic drivers. The pronounced dose–response for non-informative cases (Q5 vs Q1: OR 5.6, 95% CI 3.93–7.90) identifies a high-risk subgroup.

Current management of these patients leans heavily on family history to triage residual risk. Yet 38.0% of our non-informative cases without an affected first-degree relative fell in the top PRS quintile, nearly twice the 20% expected (OR 2.45, 95% CI 1.75–3.43). This is precisely the subgroup that family-history-based criteria would not flag. PRS thus captures a high-risk fraction orthogonal to both monogenic findings and family history, revealing an unmet diagnostic need for a clinical population that previously fell short of prognostic options.

The top quintile is therefore a candidate decision threshold for supplementary MRI or shortened screening intervals; also women above the 90th control percentile carried 2.70-fold odds relative to the 10th–90th percentile reference (95% CI 2.10–3.47). Its relative-risk scale is comparable to the gradient that triggers MRI in women with extremely dense breasts or a strong family history. Because the quintile

boundaries were derived from the control distribution, this thresholding mirrors how PRS percentiles would be deployed in practice, in line with the risk-based screening framework proposed by Mars et al. [41] and the clinical implementation strategy of the WISDOM trial [12, 65].

Among monogenic carriers, PRS retained a modest association with case status (AUC 0.576, 95% CI 0.516–0.636; $d = 0.226$, $P = 0.037$), while the extreme-quintile contrast was directionally consistent but did not reach significance (Q5 vs Q1 OR 1.80, 95% CI 0.94–3.46; $P = 0.078$), as expected with only 95 carriers. This is consistent with a modest polygenic modifying effect even in the presence of high-penetrance variants, in line with Kuchenbaecker et al. [66] and Barnes et al. [59], and supports the integration of PRS into multifactorial risk models such as BOADICEA [16] for both informative and non-informative populations.

Molecular subtype-specific discriminative ability

PRS performance varied substantially by molecular subtype, with the strongest results for hormone receptor-positive tumors (Luminal B HER2-positive: AUC 0.689; Luminal A-like: 0.677) and markedly reduced power for triple-negative breast cancer (AUC 0.568). This pattern is consistent with the distinct genetic architecture of triple-negative disease, for which only a subset of known breast cancer susceptibility loci are associated [67].

Pathogenic variants in high-penetrance HBOC genes are also enriched among triple-negative cases, the subgroup where PRS discrimination was consistently weakest. This fits the hypothesis that monogenic risk dominates the genetic architecture of triple-negative disease. Current PRS, dominated by loci discovered in overall and ER-positive analyses, capture little of its polygenic component [7].

The strong performance for Luminal B HER2-positive tumors (AUC 0.689, the highest subtype-specific AUC in this study) is noteworthy and may reflect the convergence of ER-positive and HER2-pathway susceptibility loci in these tumors. Conversely, the reduced discriminative ability for triple-negative cases highlights a clinically important gap: the patients for whom PRS-based risk stratification is least effective are those with the most aggressive tumor subtype and the greatest need for early detection. Future PRS development efforts specifically targeting triple-negative susceptibility loci may address this limitation.

Genomic convergence and biological implications

The analysis of SNP overlap revealed that the top-performing PRS were dominated by a large cluster of mutually highly similar PRS, comprising scores from six research groups applying Bayesian shrinkage methods (predominantly PRS-CS and LDpred2) to Breast Cancer Association Consortium (BCAC) GWAS summary statistics, likely reflecting shared development methodology. Despite independent development using different statistical methods, training populations, and variant selection strategies, the top scores share a large set of well-characterized susceptibility variants, consistent with their derivation from overlapping GWAS data.

From a clinical perspective, this redundancy has a reassuring implication: the choice among the best-performing PRS is less consequential than the decision to implement PRS-based risk stratification at all, as most top scores capture essentially the same biological signal.

From a research perspective, the two score families that reach comparable AUCs from largely distinct variant sets are not explained by methodological convergence alone. This suggests that the polygenic architecture of breast cancer includes more than one semi-independent signal cluster. Future functional studies targeting the genomic regions uniquely captured by these distinct scores, and future composite scores integrating both the shared and the distinct signals, may reveal novel biological mechanisms and achieve discriminative accuracy beyond the current plateau of $AUC \approx 0.66$. These regions have remained incompletely characterized despite intensive investigation. This may reflect gene–environment interactions, epistatic effects, or indirect pathways acting through intermediate phenotypes such as mammographic density, all of which are difficult to detect in standard single-variant GWAS.

Of the 181 universal variants present in all 20 top-ranked PRS, 81 (45%) mapped within ± 500 kb of genome-wide significant breast cancer lead variants. The ± 500 kb equals about 1 centimorgan in humans, thereby giving a roughly 1% chance per generation for a crossing over event, which would lead to a decoupling of linkage. The remaining 100 universal variants (55%) did not map to currently catalogued breast cancer risk loci and may represent additional susceptibility signals or tag variants in linkage disequilibrium with loci not captured by current GWAS Catalog lead SNPs. Because variants were matched by exact position, 181 is a lower bound on the true convergence between scores.

Limitations

This study has several limitations. First, the retrospective, single-center design reflects a practical necessity: prospective genome-sequenced cohorts with sufficient sample sizes and follow-up duration remain rare, resource-intensive, and years from completion. This study therefore represents a pragmatic approach to benchmarking PRS performance under real-world clinical sequencing conditions. However, this design limits generalizability and precludes assessment of calibration, the agreement between predicted absolute risk and observed incidence, which requires a population-based longitudinal cohort as reference. While discrimination metrics (AUC, Cohen's d) are valid and interpretable in our case–control setting, calibration remains an essential prerequisite for direct clinical implementation in absolute risk prediction models such as BOADICEA and should be evaluated in future prospective studies. Additionally, the use of a control cohort referred for rare diseases unrelated to breast cancer, rather than a population-representative sample, may introduce subtle selection biases despite ensuring uniform sequencing quality.

Second, clinical annotation was incomplete for several variables, including molecular subtype (unavailable for 26.8% of cases), breast density, anthropometric data, and systematic family history. This reflects the multi-center referral nature of clinical genomic sequencing: a substantial proportion of samples and patients originated from external institutions, which did not consistently provide complete

clinical records alongside sequencing requests. Complete-case exclusion of patients with missing annotation was considered but rejected on both scientific and ethical grounds. GS represents a resource-intensive and clinically irreplaceable investigation, and systematic exclusion of externally referred patients would have substantially reduced cohort size while introducing a selection bias towards cases with more comprehensive clinical documentation. These missing data nonetheless limit the precision of subgroup analyses and preclude full adjustment for potential confounders, and future studies should prospectively capture complete clinical annotation at the point of sequencing referral.

Third, the analysis was restricted to participants of European genetic ancestry ($n = 2,088$), with the 152 non-European participants excluded, consistent with the target population for which the evaluated PRS were originally developed. Transferability to non-European populations was not an aim of this study and cannot be inferred from these data, a deliberate scope limitation rather than a methodological shortcoming. This reflects a fundamental property of GWAS-derived PRS: linkage disequilibrium between genotyped variants and functional causal loci decays with increasing genetic distance from the discovery population, attenuating predictive accuracy in ancestrally distinct cohorts. Systematic benchmarking of PRS performance in diverse populations remains an important and unresolved research priority, but requires dedicated cohorts and purpose-built multi-ancestry PRS, which were beyond the scope of the present study.

Fourth, subgroup analyses were exploratory and not adjusted for multiple comparisons; findings in small subgroups (e.g. contralateral breast cancer, $n = 44$; ACR D, $n = 90$) should be interpreted with caution and require independent confirmation. However, given the resource-intensive nature of GS and the resulting scarcity of sufficiently large, clinically annotated genome-sequenced cohorts, replication in comparably powered datasets remains a substantial practical challenge. These subgroup findings are therefore best understood as hypothesis-generating observations that motivate targeted prospective validation rather than as definitive conclusions.

Future directions

Several avenues for future research emerge from this work. First, prospective validation of PRS-based risk stratification in independent, multi-center cohorts, ideally including both European and non-European populations, is needed to confirm the clinical utility of the risk gradients identified here. Second, integration of PRS with informative testing, clinical risk factors (family history, breast density, reproductive history), and imaging-based risk markers into a unified multifactorial risk model, such as the BOADICEA framework [64], which already accommodates alternative PRS [16], would enable personalized risk estimates that capture both common and rare genetic contributions. Third, functional characterization of the 181 universal variants and their neighboring genomic regions uniquely captured by the distinct PRS may reveal novel biological mechanisms underlying breast cancer susceptibility. Fourth, laboratories implementing PRS from GS data should establish platform-specific reference distributions to ensure comparability across sequencing platforms and with array-derived normative data.

This will matter even more if PRS are ever extended beyond SNPs. Copy number and structural variants are read out from sequencing coverage rather than from individual base calls, which makes them more dependent on how a sample was prepared and on which instrument it was run; a recent genome-wide study of copy number variation in UK Biobank had to discard an entire sequencing release for this reason [68]. Such an extension is nonetheless worth pursuing, because some copy number variants are poorly correlated with the SNPs around them and can carry disease associations that SNP-based studies never detect. This study also guides future evaluations of PRS in other clinical entities.

Conclusions

Multiple breast cancer PRS stratified risk when applied to GS data in a real-world HBOC clinical setting. The best-performing PRS (PGS004153) achieved an AUC of 0.656 and a Cohen's d of 0.576 for discriminating non-informative breast cancer cases from controls. AUC values among the top 20 PRS ranged from 0.643 to 0.656, indicating that multiple independently developed scores achieve comparable discriminative accuracy. The BRIDGES 306-SNP PRS ranked 53rd with an AUC of 0.621 and overlapping confidence intervals; it remains a compact, sequencing-native option well suited to NGS-based clinical workflows.

Risk quintile analysis identified a subgroup of non-informative cases with 5.6-fold elevated odds in the highest PRS quintile compared to the lowest (OR 5.57, 95% CI 3.93–7.90), suggesting that PRS-based risk stratification could meaningfully inform clinical management for non-informative patients. PRS accuracy was strongest for hormone receptor-positive tumor subtypes, while triple-negative breast cancer showed markedly reduced discriminative ability.

Analysis of constituent SNPs revealed that the majority of top-ranked PRS converge on a shared set of established susceptibility loci, while a minority of distinct PRS capture independent genomic signals with comparable discriminative accuracy. These findings support the existence of more than one semi-independent polygenic pathway in breast cancer susceptibility.

These findings provide an evidence-based framework for selecting and integrating PRS into clinical genomic workflows for breast cancer risk assessment. The high redundancy among highest-AUC leading scores implies that the clinical decision of which specific PRS to implement may be less consequential than the decision to implement PRS at all.

Abbreviations

ACR

American College of Radiology

AFR

African

AUC

area under the receiver operating characteristic curve

BCAC
Breast Cancer Association Consortium
BMI
body mass index
BOADICEA
Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm
BRIDGES
Breast Cancer Risk after Diagnostic Gene Sequencing
CI
confidence interval
DAkkS
Deutsche Akkreditierungsstelle (German national accreditation body)
DCIS
ductal carcinoma in situ
EAS
East Asian
EDTA
ethylenediaminetetraacetic acid
ER
estrogen receptor
EUR
European
GC-HBOC
German Consortium for Hereditary Breast and Ovarian Cancer
GS
short-read genome sequencing
GWAS
genome-wide association study
HBOC
hereditary breast and ovarian cancer
HER2
human epidermal growth factor receptor 2
HR
hormone receptor
IMGAG
Institute of Medical Genetics and Applied Genomics
IQR
interquartile range
LLM
large language model

MRI
magnetic resonance imaging
NGS
next-generation sequencing
NHGRI-EBI
National Human Genome Research Institute–European Bioinformatics Institute
NICE
National Institute for Health and Care Excellence
OR
odds ratio
PGS
polygenic score
PR
progesterone receptor
PRS
polygenic risk score
ROR-P
risk of recurrence weighted on proliferation
SAS
South Asian
SD
standard deviation
SNP
single-nucleotide polymorphism
TNM
tumor–node–metastasis classification
UKBB
UK Biobank
VCF
variant call format

Declarations

Ethics approval and consent to participate

This study was approved by the ethics committee of the University of Tübingen (project number 133/2021BO1). Patients with translational sequencing provided consent under the Ge-Med project (ClinicalTrials.gov: NCT04760522 | Study registration dates: First Submitted 16 February 2021; First Submitted that Met QC Criteria 16 February 2021; First Posted 18 February 2021). All participants provided written informed consent.

Consent for publication

Not applicable.

Availability of data and materials

Scoring files for all evaluated polygenic scores are openly available from the PGS Catalog (<https://www.pgscatalog.org>) under the EBI Terms of Use, with PGS identifiers and source publications listed in Additional file 1: Table S1. The datasets generated and analysed during the current study are not publicly available due to patient privacy and the terms of participant consent, but are available from the corresponding author on reasonable request. The PRS calculation pipeline is openly available on GitHub (<https://github.com/ereyol/prs-homref>); the exact version used for this study is archived on Zenodo [44] and released under the GPL-3.0 licence.

Competing interests

GD has received research funding from the European Union's Horizon 2020 research and innovation programme (grant agreement 779257) and consulting fees from Diagenius BV, Amsterdam, the Netherlands, both unrelated to the present work. NC has received research funding from the Deutsche Forschungsgemeinschaft (INST 37/1049-1) and payment for lectures at the University of Reutlingen. SO has received research funding from Illumina Inc. and honoraria from Illumina Inc. for a presentation at the GfH conference 2023, as well as support for attending meetings from Illumina Inc. and ONT in 2023. TBH has received research funding from Illumina Inc. OR has received research funding from Illumina Inc., the German Research Foundation, the European Union, and the German Ministry of Health, and is the representative of the German Ministry of Health in the European +MG project (unpaid). CS has received research funding from Illumina Inc. and research grants from BMS Stiftung Immunonkologie and Westdeutsche Studiengruppe GmbH, outside the submitted work. The remaining authors declare that they have no competing interests.

Funding

This work was partially funded by the European Union via the Can.Heal grant 101080009 to OR, SO, and CS. The funder had no role in the design of the study; in the collection, analysis, and interpretation of data; or in writing the manuscript.

Authors' contributions

EE wrote the original draft of the manuscript. EE and DW developed the PRS calculation pipeline. EE performed the PRS benchmarking, statistical analyses, and data visualization. KB, IG, and AH recruited patients and collected clinical data. OR, SO, and CS acquired funding. DW supervised this manuscript. UF, MS, AS, SG, GD, NC, KM, JW, and TBH provided critical commentary on the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank the technicians of the molecular diagnostics team of the Institute of Medical Genetics and Applied Genomics for their excellent technical support.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71:209–49. <https://doi.org/10.3322/caac.21660>
2. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of Breast Cancer Screening: Systematic Review and Meta-analysis to Update the 2009 U.S. Preventive Services Task Force Recommendation. *Ann Intern Med*. 2016;164:244–55. <https://doi.org/10.7326/M15-0969>
3. Badal K, Maniam A, Otis SU, Esserman LJ. Translating risk-based breast cancer screening to limited-resource settings. *BMC Glob Public Health*. 2025;3:59. <https://doi.org/10.1186/s44263-025-00176-4>
4. Shieh Y, Eklund M, Sawaya GF, Black WC, Kramer BS, Esserman LJ. Population-based screening for cancer: hope and hype. *Nat Rev Clin Oncol*. 2016;13:550–65. <https://doi.org/10.1038/nrclinonc.2016.50>
5. Tung N, Ricker C, Messersmith H, Balmaña J, Domchek S, Stoffel EM, et al. Selection of Germline Genetic Testing Panels in Patients With Cancer: ASCO Guideline. *J Clin Oncol*. 2024;42:2599–615. <https://doi.org/10.1200/JCO.24.00662>
6. Michailidou K, Lindström S, Dennis J, Beesley J, Hui S, Kar S, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. 2017;551:92–4. <https://doi.org/10.1038/nature24284>
7. Zhang H, Ahearn TU, Lecarpentier J, Barnes D, Beesley J, Qi G, et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat Genet*. 2020;52:572–81. <https://doi.org/10.1038/s41588-020-0609-2>
8. Roberts E, Howell S, Evans DG. Polygenic risk scores and breast cancer risk prediction. *Breast*. 2023;67:71–7. <https://doi.org/10.1016/j.breast.2023.01.003>
9. Wand H, Lambert SA, Tamburro C, Iacocca MA, O'Sullivan JW, Sillari C, et al. Improving reporting standards for polygenic scores in risk prediction studies. *Nature*. 2021;591:211–9. <https://doi.org/10.1038/s41586-021-03243-6>
10. Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A, et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet*. 2019;104:21–34. <https://doi.org/10.1016/j.ajhg.2018.11.002>
11. Stiller S, Drukewitz S, Lehmann K, Hentschel J, Strehlow V. Clinical Impact of Polygenic Risk Score for Breast Cancer Risk Prediction in 382 Individuals with Hereditary Breast and Ovarian Cancer Syndrome. *Cancers*. 2023;15:3938. <https://doi.org/10.3390/cancers15153938>
12. Fergus KB, Heise RS, Madlensky L, Fiscalini A, Sabacan L, Theiner S, et al. Integrating breast cancer polygenic risk scores at scale in the WISDOM Study: a national randomized personalized screening

- trial. *Genome Med.* 2025;17:97. <https://doi.org/10.1186/s13073-025-01524-7>
13. Lakeman IMM, Rodríguez-Gironde M, Lee A, Ruiter R, Stricker BH, Wijnant SRA, et al. Validation of the BOADICEA model and a 313-variant polygenic risk score for breast cancer risk prediction in a Dutch prospective cohort. *Genet Med.* 2020;22:1803–11. <https://doi.org/10.1038/s41436-020-0884-4>
 14. Lambert SA, Gil L, Jupp S, Ritchie SC, Xu Y, Buniello A, et al. The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. *Nat Genet.* 2021;53:420–5. <https://doi.org/10.1038/s41588-021-00783-5>
 15. Baumann A, Ruckert C, Meier C, Hutschenreiter T, Remy R, Schnur B, et al. Limitations in next-generation sequencing-based genotyping of breast cancer polygenic risk score loci. *Eur J Hum Genet.* 2024;32:987–97. <https://doi.org/10.1038/s41431-024-01647-2>
 16. Mavaddat N, Ficoirella L, Carver T, Lee A, Cunningham AP, Lush M, et al. Incorporating Alternative Polygenic Risk Scores into the BOADICEA Breast Cancer Risk Prediction Model. *Cancer Epidemiol Biomarkers Prev.* 2023;32:422–7. <https://doi.org/10.1158/1055-9965.EPI-22-0756>
 17. Tanha HM, Law MH, Ingold N, Olsen CM, Pandeya N, Milne RL, et al. Performance of different polygenic risk scores for breast cancer risk prediction: in-depth evaluations across large UK and Australian cohorts. *Eur J Hum Genet.* 2026;34:278–87. <https://doi.org/10.1038/s41431-025-02003-8>
 18. Witt D, Sturm M, Stäbler A, Menden B, Ruisinger L, Bosse K, et al. Clinical genome sequencing in patients with hereditary breast and ovarian cancer: Concept, implementation and benefits. *Breast.* 2025;82:104505. <https://doi.org/10.1016/j.breast.2025.104505>
 19. Rhiem K, Auber B, Briest S, Dikow N, Ditsch N, Dragicevic N, et al. Consensus Recommendations of the German Consortium for Hereditary Breast and Ovarian Cancer. *Breast Care.* 2022;17:199–207. <https://doi.org/10.1159/000516376>
 20. Riess O, Sturm M, Menden B, Liebmann A, Demidov G, Witt D, et al. Genomes in clinical care. *npj Genom Med.* 2024;9:20. <https://doi.org/10.1038/s41525-024-00402-2>
 21. ngs-bits documentation: SampleAncestry. <https://github.com/imgag/ngs-bits/blob/master/doc/tools/SampleAncestry/index.md>. Accessed 11 March 2026.
 22. Schroeder CM, Hilke FJ, Löffler MW, Bitzer M, Lenz F, Sturm M. A comprehensive quality control workflow for paired tumor-normal NGS experiments. Berger B, editor. *Bioinformatics.* 2017;33:1721–2. <https://doi.org/10.1093/bioinformatics/btx032>
 23. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol.* 2013;24:2206–23. <https://doi.org/10.1093/annonc/mdt303>
 24. American College of Radiology. ACR BI-RADS Atlas: Breast Imaging Reporting and Data System. 5th ed. Reston (VA): American College of Radiology; 2013.
 25. PGS Catalog. <https://www.pgscatalog.org>. Accessed 11 March 2026.

26. Lambert SA, Wingfield B, Gibson JT, Gil L, Ramachandran S, Yvon F, et al. Enhancing the Polygenic Score Catalog with tools for score calculation and ancestry normalization. *Nat Genet.* 2024;56:1989–94. <https://doi.org/10.1038/s41588-024-01937-x>
27. Ohta R, Tanigawa Y, Suzuki Y, Kellis M, Morishita S. A polygenic score method boosted by non-additive models. *Nat Commun.* 2024;15:4433. <https://doi.org/10.1038/s41467-024-48654-x>
28. Shi Z, Yu H, Wu Y, Lin X, Bao Q, Jia H, et al. Systematic evaluation of cancer-specific genetic risk score for 11 types of cancer in The Cancer Genome Atlas and Electronic Medical Records and Genomics cohorts. *Cancer Med.* 2019;8:3196–205. <https://doi.org/10.1002/cam4.2143>
29. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics.* 2011;27:2156–8. <https://doi.org/10.1093/bioinformatics/btr330>
30. Ponelle-Chachuat F, Lepage M, Viala S, Kelleci M, Floch EL, Dandine-Roulland C, et al. Development and validation of a next-generation sequencing-based method for calculating the breast cancer polygenic risk score PRS313. *Breast.* 2025;84:104580. <https://doi.org/10.1016/j.breast.2025.104580>
31. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4:s13742-015-0047–8. <https://doi.org/10.1186/s13742-015-0047-8>
32. Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. *Nat Biotechnol.* 2017;35:316–9. <https://doi.org/10.1038/nbt.3820>
33. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol.* 2020;38:276–8. <https://doi.org/10.1038/s41587-020-0439-x>
34. Raben TG, Lello L, Widen E, Hsu SDH. Biobank-scale methods and projections for sparse polygenic prediction from machine learning. *Sci Rep.* 2023;13:11662. <https://doi.org/10.1038/s41598-023-37580-5>
35. Sun X, Verma SP, Jia G, Wang X, Ping J, Guo X, et al. Case-Case Genome-Wide Analyses Identify Subtype-Informative Variants That Confer Risk for Breast Cancer. *Cancer Res.* 2024;84:2533–48. <https://doi.org/10.1158/0008-5472.CAN-23-3854>
36. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44:837-45.
37. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.
38. Anders CK, Johnson R, Litton J, Phillips M, Bleyer A. Breast Cancer Before Age 40 Years. *Seminars in Oncology.* 2009;36:237–49. <https://doi.org/10.1053/j.seminoncol.2009.03.001>
39. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol.* 2012;13:1141–51. [https://doi.org/10.1016/S1470-2045\(12\)70425-4](https://doi.org/10.1016/S1470-2045(12)70425-4)

40. Liu C, Zeinomar N, Chung WK, Kiryluk K, Gharavi AG, Hripcsak G, et al. Generalizability of Polygenic Risk Scores for Breast Cancer Among Women With European, African, and Latinx Ancestry. *JAMA Netw Open*. 2021;4:e2119084. <https://doi.org/10.1001/jamanetworkopen.2021.19084>
41. Mars N, Kerminen S, Tamlander M, Pirinen M, Jakkula E, Aaltonen K, et al. Comprehensive Inherited Risk Estimation for Risk-Based Breast Cancer Screening in Women. *J Clin Oncol*. 2024;42:1477–87. <https://doi.org/10.1200/JCO.23.00295>
42. Jaccard P. Étude comparative de la distribution florale dans une portion des Alpes et du Jura. Imprimerie Corbaz & Comp.; 1901 [cited 2026 Aug 9]; <https://doi.org/10.5169/SEALS-266450>
43. Cerezo M, Sollis E, Ji Y, Lewis E, Abid A, Bircan KO, et al. The NHGRI-EBI GWAS Catalog: standards for reusability, sustainability and diversity. *Nucleic Acids Research*. 2025;53:D998–1005. <https://doi.org/10.1093/nar/gkae1070>
44. Eyol E, Witt D. prs-homref: a pipeline for polygenic risk score retrieval, quality control and scoring under a homozygous-reference assumption [Internet]. Zenodo; 2026 [cited 2026 Aug 9]. <https://doi.org/10.5281/ZENODO.21861191>
45. Monti R, Eick L, Hudjashov G, Läll K, Kanoni S, Wölford BN, et al. Evaluation of polygenic scoring methods in five biobanks shows larger variation between biobanks than methods and finds benefits of ensemble learning. *Am J Hum Genet*. 2024;111:1431–47. <https://doi.org/10.1016/j.ajhg.2024.06.003>
46. Fritsche LG, Patil S, Beesley LJ, VandeHaar P, Salvatore M, Ma Y, et al. Cancer PRSweb: An Online Repository with Polygenic Risk Scores for Major Cancer Traits and Their Evaluation in Two Independent Biobanks. *Am J Hum Genet*. 2020;107:815–36. <https://doi.org/10.1016/j.ajhg.2020.08.025>
47. Mars N, Widén E, Kerminen S, Meretoja T, Pirinen M, Della Briotta Parolo P, et al. The role of polygenic risk and susceptibility genes in breast cancer over the course of life. *Nat Commun*. 2020;11:6383. <https://doi.org/10.1038/s41467-020-19966-5>
48. Namba S, Saito Y, Kogure Y, Masuda T, Bondy ML, Gharahkhani P, et al. Common Germline Risk Variants Impact Somatic Alterations and Clinical Features across Cancers. *Cancer Research*. 2023;83:20–7. <https://doi.org/10.1158/0008-5472.CAN-22-1492>
49. Youssef O, Loukola A, Zidi-Mouaffak YHS, Tamlander M, Ruotsalainen S, Kilpeläinen E, et al. High-Resolution Genotyping of Formalin-Fixed Tissue Accurately Estimates Polygenic Risk Scores in Human Diseases. *Lab Invest*. 2024;104:100325. <https://doi.org/10.1016/j.labinv.2024.100325>
50. Hu J, Ye Y, Zhou G, Zhao H. Using clinical and genetic risk factors for risk prediction of 8 cancers in the UK Biobank. *JNCI Cancer Spectrum*. 2024;8:pkae008. <https://doi.org/10.1093/jncics/pkae008>
51. Truong B, Hull LE, Ruan Y, Huang QQ, Hornsby W, Martin H, et al. Integrative polygenic risk score improves the prediction accuracy of complex traits and diseases. *Cell Genom*. 2024;4:100523. <https://doi.org/10.1016/j.xgen.2024.100523>
52. Jermy B, Läll K, Wölford BN, Wang Y, Zguro K, Cheng Y, et al. A unified framework for estimating country-specific cumulative incidence for 18 diseases stratified by polygenic risk. *Nat Commun*.

2024;15:5007. <https://doi.org/10.1038/s41467-024-48938-2>

53. Glubb DM, Maranian MJ, Michailidou K, Pooley KA, Meyer KB, Kar S, et al. Fine-Scale Mapping of the 5q11.2 Breast Cancer Locus Reveals at Least Three Independent Risk Variants Regulating MAP3K1. *Am J Hum Genet.* 2015;96:5–20. <https://doi.org/10.1016/j.ajhg.2014.11.009>
54. Netherlands Collaborative Group on Hereditary Breast and Ovarian Cancer (HEBON), Familial Breast Cancer Study (FBCS), The Gene Environment Interaction of Breast Cancer in Germany (GENICA) Network, kConFab Investigators, Australian Ovarian Cancer Study Group, Ghoussaini M, et al. Genome-wide association analysis identifies three new breast cancer susceptibility loci. *Nat Genet.* 2012;44:312–8. <https://doi.org/10.1038/ng.1049>
55. SEARCH, The GENICA Consortium, kConFab, Australian Ovarian Cancer Study Group, Ahmed S, Thomas G, et al. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat Genet.* 2009;41:585–90. <https://doi.org/10.1038/ng.354>
56. EMBRACE, GEMO Study Collaborators, HEBON, kConFab Investigators, Dunning AM, Michailidou K, et al. Breast cancer risk variants at 6q25 display different phenotype associations and regulate ESR1, RMND1 and CCDC170. *Nat Genet.* 2016;48:374–86. <https://doi.org/10.1038/ng.3521>
57. Shieh Y, Roger J, Yau C, Wolf DM, Hirst GL, Swigart LB, et al. Development and testing of a polygenic risk score for breast cancer aggressiveness. *npj Precis Onc.* 2023;7:42. <https://doi.org/10.1038/s41698-023-00382-z>
58. Ohbe H, Hachiya T, Yamaji T, Nakano S, Miyamoto Y, Sutoh Y, et al. Development and validation of genome-wide polygenic risk scores for predicting breast cancer incidence in Japanese females: a population-based case-cohort study. *Breast Cancer Res Treat.* 2023;197:661–71. <https://doi.org/10.1007/s10549-022-06843-6>
59. Barnes DR, Rookus MA, McGuffog L, Leslie G, Mooij TM, Dennis J, et al. Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of BRCA1 and BRCA2 pathogenic variants. *Genet Med.* 2020;22:1653–66. <https://doi.org/10.1038/s41436-020-0862-x>
60. Melnikow J, Fenton JJ, Whitlock EP, Miglioretti DL, Weyrich MS, Thompson JH, et al. Supplemental Screening for Breast Cancer in Women With Dense Breasts: A Systematic Review for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2016;164:268–78. <https://doi.org/10.7326/M15-1789>
61. CanRisk. <https://www.canrisk.org>. Accessed 11 March 2026.
62. Lindström S, Thompson DJ, Paterson AD, Li J, Gierach GL, Scott C, et al. Genome-wide association study identifies multiple loci associated with both mammographic density and breast cancer risk. *Nat Commun.* 2014;5:5303. <https://doi.org/10.1038/ncomms6303>
63. Evans DGR, Van Veen EM, Harkness EF, Brentnall AR, Astley SM, Byers H, et al. Breast cancer risk stratification in women of screening age: Incremental effects of adding mammographic density, polygenic risk, and a gene panel. *Genet Med.* 2022;24:1485–94. <https://doi.org/10.1016/j.gim.2022.03.009>

64. Lee A, Mavaddat N, Wilcox AN, Cunningham AP, Carver T, Hartley S, et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med*. 2019;21:1708–18. <https://doi.org/10.1038/s41436-018-0406-9>
65. Esserman LJ, Ficalini AS, Naeim A, Van'T Veer LJ, Kaster A, Scheuner MT, et al. Risk-Based vs Annual Breast Cancer Screening: The WISDOM Randomized Clinical Trial. *JAMA*. 2026;335:763. <https://doi.org/10.1001/jama.2025.24784>
66. Kuchenbaecker KB, McGuffog L, Barrowdale D, Lee A, Soucy P, Dennis J, et al. Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. *JNCI: Journal of the National Cancer Institute* [Internet]. 2017 [cited 2026 Aug 9];109. <https://doi.org/10.1093/jnci/djw302>
67. Purrington KS, Slager S, Eccles D, Yannoukakos D, Fasching PA, Miron P, et al. Genome-wide association study identifies 25 known breast cancer susceptibility loci as risk factors for triple-negative breast cancer. *Carcinogenesis*. 2014;35:1012–9. <https://doi.org/10.1093/carcin/bgt404>
68. Eisenberg M, Packer R, Shrine N, Demidov G, Pack H, Hollox EJ, et al. Phenome-wide association of multiallelic copy number variation in 422,170 UK Biobank individuals reveals novel genetic loci associated with disease [Internet]. *Genetic and Genomic Medicine*; 2026 [cited 2026 Aug 9]. <https://doi.org/10.64898/2026.06.03.26354825>

Figures

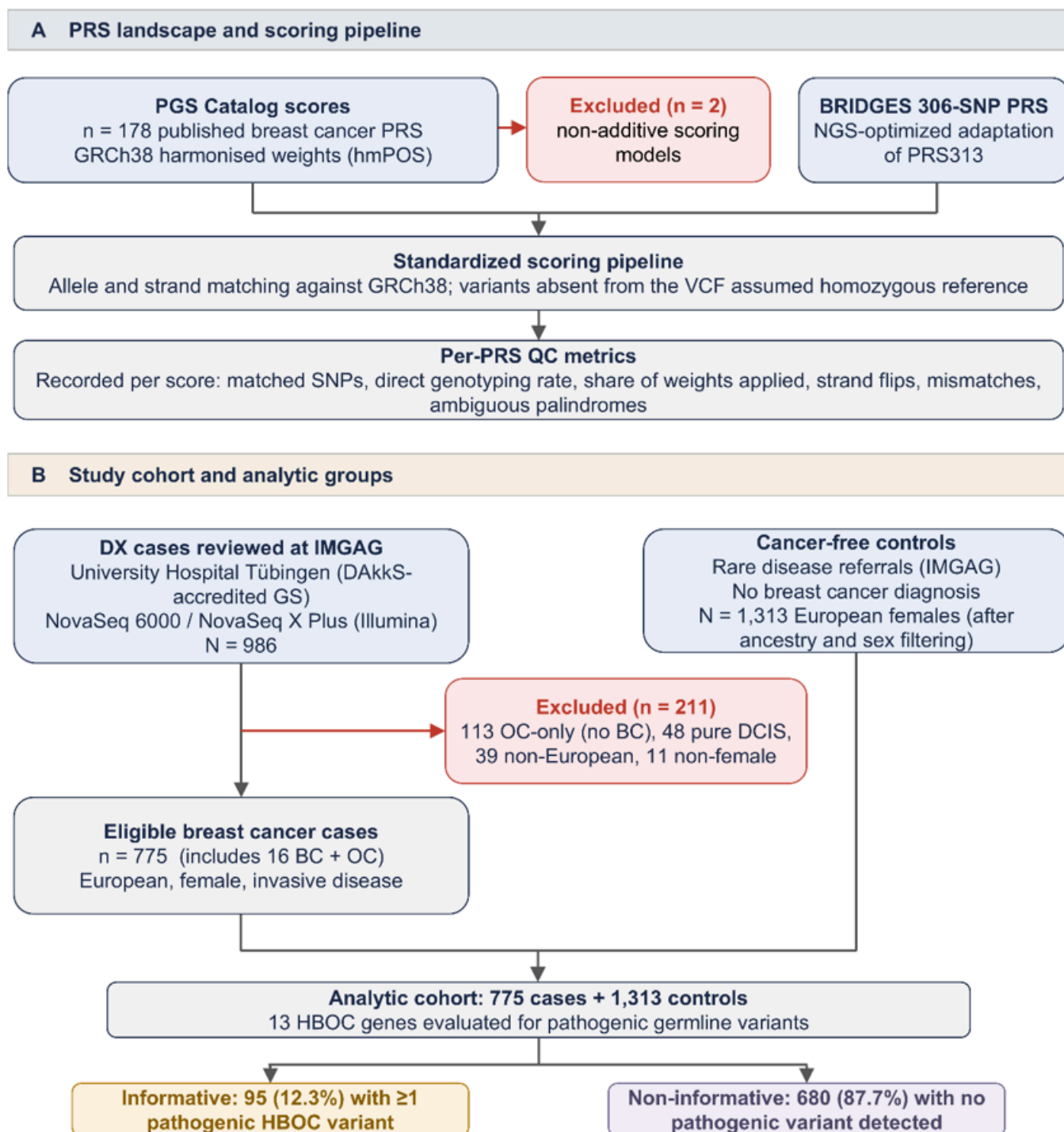


Figure 1

Study design and cohort. (A) PRS landscape and scoring pipeline. (B) Cohort assembly into 775 breast cancer cases and 1,313 controls, with stratification by test informativeness.

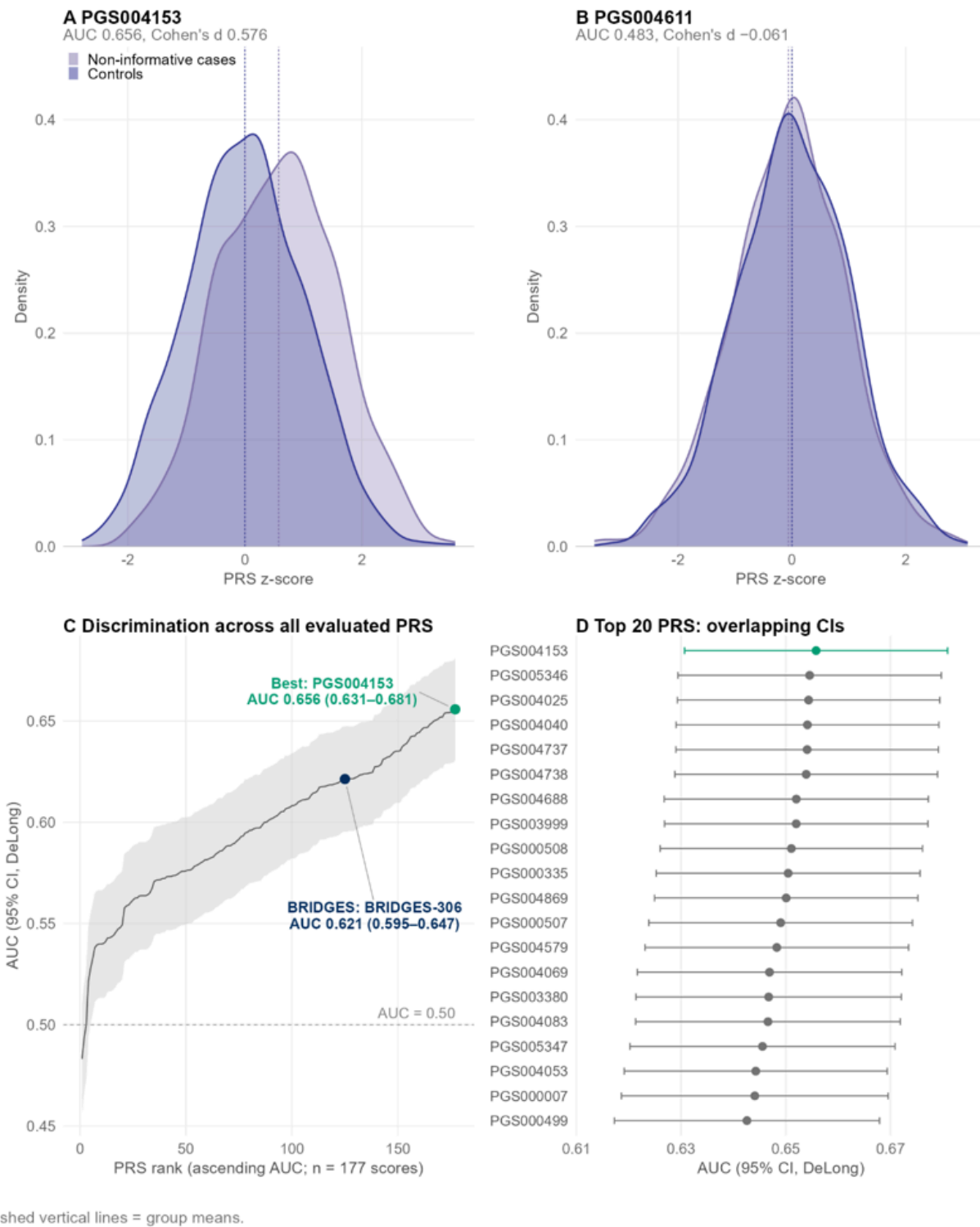


Figure 2

PRS discrimination across 177 evaluated scores. PRS z-score distributions for non-informative cases ($n = 680$) versus controls ($n = 1,313$) for (A) the best-performing score PGS004153 (AUC 0.656, Cohen's $d = 0.576$) and (B) PGS004611 (AUC 0.483, $d = -0.061$), a score developed to predict tumor aggressiveness rather than disease risk. (C) AUC with 95% DeLong confidence intervals for all 177 scores ranked by

AUC, with the 0.50 null line and the best and BRIDGES 306-SNP scores highlighted. (D) The top 20 scores, showing that their confidence intervals overlap throughout.

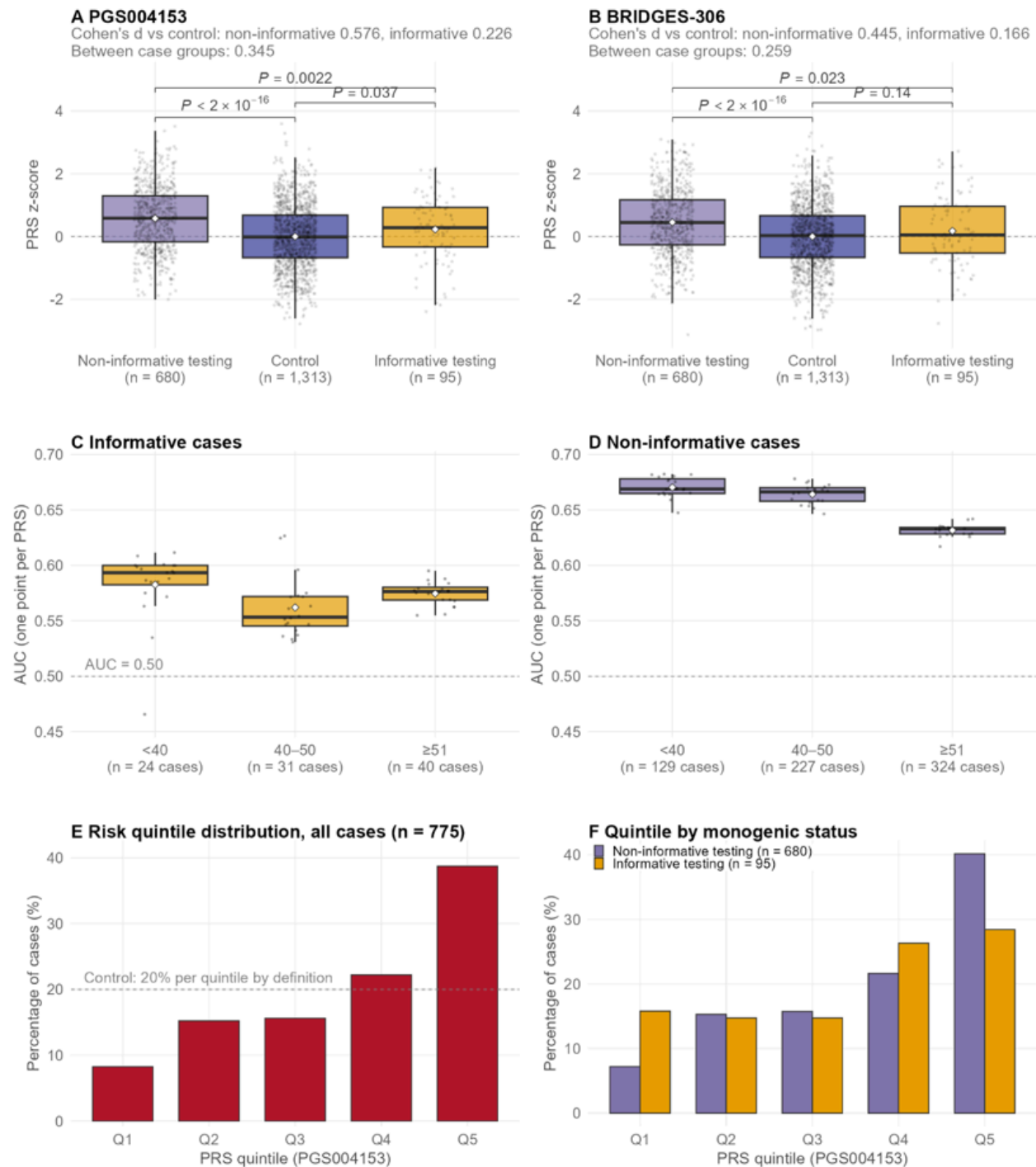


Figure 3

PRS performance across clinical strata. All panels use PGS004153 unless stated otherwise, and all comparisons use the same 1,313 controls. (A, B) PRS z-score distributions by test informativeness for

(A) PGS004153 and (B) BRIDGES-306; axis labels give full group sizes. Boxplots show median and IQR, with pairwise Welch t-tests and Cohen's d. (C, D) Discriminative ability (AUC) by age at diagnosis in (C) informative and (D) non-informative cases. Axis labels give case counts per age band, summing to 95 and 680 respectively; each band is compared against all 1,313 controls. Each dot is one of the top 20 PRS. (E, F) Risk quintile distribution, quintiles defined on the control distribution: (E) all cases against the control reference of 20% per quintile, (F) cases only, split by test informativeness.



Figure 4

Ideogram of universal SNPs. Cytogenetic band positions of the 181 variants present in all 20 top-ranked PRS, across 22 chromosomes and 142 cytobands.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.xlsx](#)
- [Additionalfile2.pdf](#)