Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

[Student's name typed] Date

Approval Sheet

The approval sheet should be designed according to the plan below. See next page for several comments.

Polygenic Risk Scores for High-Grade Serous Ovarian Cancer in African American Women within the African American Cancer Epidemiology Study (AACES)

By

Alexa Kifer

Degree to be awarded: **MPH**

Department of Epidemiology

Joellen M. Schildkraut, Ph.D.

Committee Chair

_____ [Member's signature]

[Member's name, typed]

Committee Member

_____ [Member's signature]

[Member's name, typed]

Committee Member

_____ [Member's signature]

[Member's name, typed]

Committee Member

Polygenic Risk Scores for High-Grade Serous Ovarian Cancer in African American Women within the African American Cancer Epidemiology Study (AACES)

By

Alexa Kifer

B.S. Biological Health Sciences

University of South Florida

2022

Thesis Committee Chair: Dr. Joellen Schildkraut, PhD

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Epidemiology

2025

Abstract

Polygenic Risk Scores for High-Grade Serous Ovarian Cancer in African American Women within the African American Cancer Epidemiology Study (AACES)

By: Alexa Kifer

Background:
African American women experience disproportionately poor outcomes from epithelial ovarian cancer (EOC), yet they remain underrepresented in genetic studies that inform risk prediction. Polygenic risk scores (PRS), derived primarily from European ancestry populations, may not accurately capture EOC susceptibility in diverse populations.

Methods:
Using data from the African American Cancer Epidemiology Study (AACES), the largest population-based cohort of African American women with EOC in the U.S., we evaluated a previously developed PRS constructed from 24 genome-wide significant loci. Logistic and multinomial logistic regression models were used to assess associations between standardized PRS and EOC risk, adjusting for age, ancestry, family history, and geographic region. Subtype-specific analyses were conducted to examine histologic variation in genetic risk.

Results:
Among 592 Black women with EOC, a one-unit increase in PRS was associated with significantly increased odds of high-grade serous ovarian cancer (HGSOC) (adjusted OR = 1.37; 95% CI: 1.16–1.62; p = 0.0004). No significant association was observed between PRS and non-HGSOC histotypes. Participants in the top 5% of the PRS distribution had more than twice the odds of HGSOC compared to those in the bottom 80%, indicating a dose-response relationship. These associations persisted after adjusting for potential confounders. Binary logistic models confirmed the robustness and specificity of the PRS to HGSOC.

Conclusions:
This study demonstrates that PRS constructed from European ancestry GWAS data can capture meaningful genetic risk for HGSOC among African American women, despite reduced transferability. These findings underscore the need for ancestry-informed PRS models and inclusive genomic research to advance equitable precision medicine. Incorporating PRS into risk stratification frameworks may improve early detection and prevention efforts for high-risk Black women disproportionately affected by EOC disparities.

Keywords: polygenic risk score, high-grade serous ovarian cancer, African American women, genetic epidemiology, AACES, health disparities, ovarian cancer, GWAS, ancestry bias

Contents

Polygenic Risk Scores for High-Grade Serous Ovarian Cancer in African American Women within the African American Cancer Epidemiology Study (AACES)

By

Alexa Kifer

B.S. Biological Health Sciences

University of South Florida

2022

Thesis Committee Chair: Dr. Joellen Schildkraut, PhD

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Epidemiology

2025

Introduction

Epithelial ovarian cancer (EOC) is a significant public health concern, specifically within the United States. Ovarian cancer is the second most common gynecological cancer, with the majority of those cases being EOC. EOC accounts for approximately 95% of all ovarian cancers, while about 5% are non-epithelial cancers such as germ cell, sex-cord stromal, and small cell ovarian cancers (1). Ovarian cancer is the leading cause of death among women diagnosed with gynecologic cancers. Its high mortality is largely attributed to the disease's nonspecific clinical symptoms and the lack of preventative screening, which likely leads to delayed diagnosis. Consequently, most individuals are diagnosed with advanced-stage disease (2).

Demographics and Mortality/Survival Rates

In the United States, ovarian cancer remains a relatively rare but highly lethal disease. According to the American Cancer Society, approximately 20,890 new cases of ovarian cancer will be diagnosed in 2025 with over 92% classified as epithelial ovarian cancer. Annually, 12,730 women are expected to die from the disease, making it the sixth leading cause of cancer-related death among women (3). Lifetime risk estimates from 2002 indicated that about 1.1% of women would develop ovarian cancer (3). Despite advancements in treatment, the overall five-year relative survival rate remains low at 49%, primarily due to the fact that the disease is most often diagnosed at advanced stages (2).

Stage at diagnosis plays a critical role in survival. When ovarian cancer is diagnosed at a localized stage (Stage I), the five-year survival rate is approximately 92%. However, only about 15% of women are diagnosed at this early stage. For cases diagnosed at regional spread (Stage II/III), survival rates drop to 73–36%, and for distant metastatic disease (Stage IV),

which accounts for the majority of cases, the survival rate falls to 31% (3). This disparity underscores the urgent need for better early detection strategies.

Demographic characteristics such as age, race, and ethnicity significantly influence both the incidence and outcomes of epithelial ovarian cancer. Ovarian cancer primarily affects older women, with a median age at diagnosis of 63 years (4). Incidence rates increase with age, peaking between ages 55 and 64, and declining thereafter. According to recent national SEER data, the age-adjusted incidence rate is approximately 10.3 per 100,000 women per year in the U.S., with variation by race and ethnicity (5).

White women have the highest reported incidence rates of ovarian cancer, followed by Hispanic, Asian/Pacific Islander, and Black women. However, Black women consistently experience the worst survival outcomes. The five-year relative survival rate for white women is approximately 50%, while it is only 41% for Black women, even after adjusting for stage at diagnosis (5). Hispanic and Asian/Pacific Islander women have somewhat better survival outcomes, with five-year survival rates closer to 55–57%, although the reasons for this remain unclear (5).

Several studies have demonstrated that Black women are more likely to be diagnosed at later stages, receive suboptimal surgical debulking, and are less likely to receive chemotherapy, all of which contribute to poorer survival (6). For example, Bristow et al. found that Black women were 25% less likely to receive guideline-concordant treatment, and this lack of optimal care was strongly associated with reduced survival (6). Even after controlling for stage, tumor characteristics, and socioeconomic status, survival disparities persist, suggesting a multifactorial interplay between structural inequities and biological differences.

Social determinants of health likely play a central role in driving these disparities. Limited access to specialized care, lack of insurance, geographic barriers, and provider bias all contribute to delayed diagnosis, under-treatment, and ultimately, higher mortality among Black women (7). Underdiagnosis and progression of disease due to these systemic factors exacerbate the survival gap and reduce opportunities for early intervention (7).

In addition to social factors, emerging research suggests that biological differences may also contribute. Analyses from the Ovarian Cancer in Women of African Ancestry Consortium (OCWAA) have highlighted the potential influence of reproductive history, comorbid conditions, and hormone use on survival differences between Black and white women (8). Furthermore, some studies suggest that specific genetic variants may be more prevalent in African American populations, potentially increasing the risk for more aggressive histologic subtypes of ovarian cancer (8). These findings underscore the importance of developing ancestry-informed genomic research and risk models tailored to underrepresented populations.

Understanding how demographics, social determinants, and tumor biology intersect is essential for improving equity in ovarian cancer care. Targeted efforts to address both systemic barriers and population-specific biological risks will be crucial to advancing personalized prevention and treatment strategies and reducing mortality among women most affected by ovarian cancer disparities.

Risk Factors

One positive thing to note is that while this is still a prevalent cancer within the United States, the overall incidence of ovarian cancer has been declining in recent decades (9). This decline is hypothesized to likely be due to better treatments and more use of oral contraceptives,

which seem to both have a protective factor (9). While the etiological of ovarian cancer is not clearly understood, there seem to be several factors that increase one's risk for the development of epithelial ovarian cancer. Some of the risk factors may include advanced age, early onset of menarche, late onset of menopause, family history, nulliparity, obesity, perineal talc use, smoking, endometriosis, and hormone replacement therapy (1). Additionally, inflammatory conditions are also thought to potentially lead to the development of ovarian cancer due to oxidative stress and deoxyribonucleic acid damage (10)(11). A positive personal or family history of breast or ovarian cancer is a well-established risk factor for ovarian cancer. A personal or first-degree family history of breast or ovarian cancer significantly elevates ovarian cancer risk. Specifically, women with a first-degree relative diagnosed with ovarian cancer have a three- to four-fold higher lifetime risk compared to women with no family history (12). Germline mutations in BRCA1 or BRCA2 genes are prevalent cause of underlying malignancy risk in individuals. Compared with a lifetime risk of 2% for the general population, the average cumulative risks by age 70 for ovarian cancer among BRCA1 or BRCA2 mutations is 59% (95% CI: 43–76) and 16.5% (95% CI: 7.5–34) respectively (12). According the American Cancer Society, there are no recommendations for screening tests for Ovarian cancer for women who are not at high risk or do not have symptoms (13). The only current recommendation is that only women who are at high risk for ovarian cancer should get screened. The 2 tests used most often (in addition to a complete pelvic exam) to screen for ovarian cancer are transvaginal ultrasound (TVUS) and the CA-125 blood test (13).This is primarily due to the limited evidence screening reduces the risk of dying from ovarian cancer, and often results in high specificity test rates (13).

Histologic Variation in Risk Factors

Epithelial ovarian cancer (EOC) encompasses several histologic subtypes, each with distinct molecular, clinical, and epidemiologic features. These subtypes include high-grade serous carcinoma, low-grade serous carcinoma, endometrioid carcinoma, clear cell carcinoma, and mucinous carcinoma vary in tumor origin, risk factors, genetic drivers, treatment response, and prognosis (14). Recognizing this heterogeneity is essential for understanding disease etiology and tailoring risk prediction models.

High-Grade Serous Ovarian Carcinoma (HGSOC) is the most common and lethal subtype, representing approximately 70% of all EOC cases (15). It is characterized by frequent TP53 mutations (observed in over 95% of cases) and genomic instability (16). HGSOC typically originates from the distal fallopian tube epithelium, not the ovarian surface, as previously thought (16). Key risk factors for HGSOC include increasing age, nulliparity, early menarche, late menopause, and particularly, germline mutations in BRCA1 and BRCA2 genes (17). HGSOC is typically diagnosed at advanced stages (III/IV) and responds initially to platinum-based chemotherapy, although recurrence is common. Low Grade Serous Ovarian Carcinoma (LGSOC) is a rare subtype, comprising less than 5% of EOC cases. Unlike HGSOC, it is indolent in behavior, often affecting younger women and progressing more slowly (18). It is genetically distinct, typically involving mutations in KRAS, BRAF, or NRAS, but lacking TP53 alterations (18). LGSOC is frequently resistant to standard chemotherapy and may arise from serous borderline tumors. There is limited evidence regarding specific environmental or reproductive risk factors for LGSOC, but oral contraceptive use may be protective, as it is with other subtypes. Endometrioid carcinoma accounts for approximately 10% of EOC cases and is often diagnosed at earlier stages with a favorable prognosis (19). It shares clinical and molecular

features with endometrial cancer and often arises in the context of endometriosis or estrogen excess. Mutations in CTNNB1, PTEN, PIK3CA, and ARID1A are common (19). Risk factors include obesity, unopposed estrogen exposure, infertility, and endometriosis. Endometrioid tumors are more common among white and Hispanic women, though population-based comparisons are limited (19). Clear cell carcinoma represents 10–12% of EOC and is also associated with endometriosis. It is more prevalent in East Asian populations, accounting for up to 25% of ovarian cancers in Japan, compared to ~12% in the U.S. (20). These tumors are typically chemo resistant and associated with poor prognosis, particularly in advanced-stage disease. Molecular alterations include ARID1A, PIK3CA, and HNF-1β mutations (20). Risk factors include endometriosis, nulliparity, and potential environmental exposures, though research is ongoing. Mucinous carcinoma is the rarest of the five major subtypes, making up 3–4% of EOC cases (21). These tumors often present at early stages and resemble gastrointestinal epithelium, which can make diagnosis challenging. Molecularly, mucinous tumors are typically driven by KRAS mutations. Risk factors are less well-defined but may include smoking. Mucinous carcinoma is more commonly diagnosed in younger women and white women, although precise racial/ethnic patterns are not well-characterized (21).

Some evidence demonstrates that the distribution of epithelial ovarian cancer (EOC) histologic subtypes varies significantly by race and ethnicity, and these differences may help explain disparities in incidence, treatment response, and survival outcomes (22).

Among white women, high-grade serous carcinoma (HGSOC) remains the predominant subtype, accounting for approximately 70–75% of EOC diagnoses. In contrast, studies suggest that Black women are proportionally more likely to be diagnosed with HGSOC and more likely to present with rare, high-grade, or poorly differentiated tumors, which are associated with worse

prognosis (23). Some data also suggest that Black women may have a slightly higher prevalence of clear cell and mucinous carcinomas, though these findings vary by cohort and may reflect underlying misclassification or differences in tumor biology.

In Asian women, particularly those from East Asia, the distribution of Histotypes is notably different. Clear cell carcinoma represents a substantial proportion of EOC cases, ranging from 20–25% in countries like Japan, compared to about 10–12% in Western populations (15). This elevated prevalence may reflect genetic predisposition, such as increased frequency of ARID1A mutations, or environmental factors, including higher rates of endometriosis, which is a known precursor lesion for clear cell tumors.

Hispanic and Latina women also exhibit unique patterns in histotype distribution. Some studies have found higher rates of endometrioid carcinoma and slightly lower frequencies of high-grade serous tumors compared to white women (23). The reasons behind these patterns are still unclear but may involve differences in reproductive factors, hormone use, genetic ancestry, and metabolic conditions such as obesity.

Prognostic Factors in Risk Factors

Beyond influencing cancer risk, histologic subtype plays a critical role in determining treatment response, survival outcomes, and long-term prognosis. Subtypes vary in chemosensitivity, recurrence rates, and survival trajectories (24). These biological differences intersect with patient-level and systemic factors to shape prognosis.

Histologic subtype significantly affects survival. While HGSOC, though aggressive, is initially chemosensitive, particularly in the presence of BRCA mutations or homologous recombination deficiency (HRD), other subtypes such as clear cell and mucinous carcinomas

tend to be chemoresistant, resulting in poorer prognosis. Low-grade serous ovarian carcinoma (LGSOC) progresses more slowly but is often resistant to standard chemotherapy. In contrast, endometrioid carcinomas are typically detected earlier and have better outcomes.

These survival disparities are further complicated by racial and ethnic variation in histotype prevalence. Among white women, HGSOC is the predominant subtype, accounting for approximately 70–75% of EOC diagnoses. In contrast, Black women in the United States are proportionally less likely to be diagnosed with HGSOC and more likely to present with rare, high-grade, or poorly differentiated tumors, which are associated with a worse prognosis (23). Some data also suggest that Black women may have a higher prevalence of clear cell and mucinous carcinomas, although findings vary across cohorts and may reflect tumor biology, environmental exposures, or diagnostic misclassification. In Asian women, particularly those from East Asia, clear cell carcinoma represents a notably higher proportion of EOC cases, up to 20–25% in countries like Japan, compared to 10–12% in Western populations (25). This higher prevalence may be influenced by genetic predisposition (e.g., increased ARID1A mutations) and a higher rate of endometriosis, a known precursor to clear cell carcinoma. Hispanic and Latina women tend to have higher rates of endometrioid carcinoma and lower frequencies of HGSOC compared to white women (26). These patterns may reflect differences in reproductive history, hormone use, genetic ancestry, and comorbidities such as obesity and diabetes.

These histotype variations by race and ethnicity carry substantial clinical implications. Subtypes that are more common among Black women are associated with greater resistance to chemotherapy and worse survival outcomes (8). Even after accounting for the stage at diagnosis, studies have shown that Black women experience worse survival outcomes than other racial groups in both early and late stages of the disease (8). Social determinants of health likely

contribute to these disparities. Limited access to healthcare among Black women often delays diagnosis and treatment, leading to underdiagnosis or progression of the disease (8). Meanwhile, HGSOC, more prevalent among white women, tends to be more chemosensitive, especially when associated with BRCA mutations, which are more commonly observed in individuals of European ancestry (27).

Importantly, histotype variation does not exist in isolation. It intersects with social determinants of health, such as access to care, quality of treatment, insurance coverage, and geographic proximity to gynecologic oncology centers. For instance, Black women are disproportionately more likely to receive care at low-volume centers, are less likely to undergo optimal cytoreductive surgery, and are less frequently offered guideline-concordant treatment, regardless of tumor histology (28). These systemic inequities further compound the biological challenges posed by aggressive histotypes.

The underrepresentation of non-European populations in genetic research, particularly genome-wide association studies (GWAS) and polygenic risk score (PRS) development, has created major gaps in our understanding of how genetic susceptibility interacts with histologic subtype by race. Without equitable representation, risk prediction tools may fail to capture the aggressive subtypes more commonly seen in women of color, thus limiting opportunities for early detection and personalized prevention.

Efforts to improve ovarian cancer outcomes must therefore address both molecular tumor features and sociostructural barriers to care. Recognizing and addressing racial and ethnic variation in histotype prevalence, prognosis, and access is essential for achieving equitable precision medicine in ovarian cancer prevention and treatment.

Genetic Risk Prediction

Genome-wide studies have been conducted to understand epithelial ovarian cancer susceptibility, In the more recent years, genome-wide association studies GWAS have increased in frequency in predicting various women's specific cancers. Ovarian cancer GWAS have identified about 33 common, low-penetrant EOC susceptibility alleles. (29). Polygenic risk scores (PRSs) offer a promising approach for predicting individual susceptibility to EOC by improving risk stratification and identifying high-risk individuals. PRSs are derived by summing the risk alleles across multiple genetic variants, with each allele weighted based on its estimated contribution to disease risk (30). Despite advancements in PRS models, the optimal selection of genetic variants and their weights to maximize predictive accuracy remains an ongoing challenge. Recent developments in PRS modeling have demonstrated their potential in assessing cancer risk, including epithelial ovarian cancer, and their utility in clinical settings (30). For example, PRSs can help identify individuals at elevated risk for adverse drug events or predict treatment responses, offering a personalized approach to cancer management and prevention (31). While rare variants in high- and moderate-penetrance susceptibility genes explain approximately 40% of the inherited risk of EOC, common genetic variants contribute significantly to the remaining heritability. Studies, such as those by Dareng et al. estimate that common susceptibility variants account for roughly 6% of EOC heritability (30).

A critical issue in PRS development is the lack of representation of non-European populations, which undermines the generalizability and accuracy of these scores. This problem, commonly referred to as ancestry bias, arises because the vast majority of genome-wide association studies (GWAS), which form the basis for PRS construction, have been conducted in individuals of European descent. This bias poses and lack of generalizability to the African

American population causes significant challenges for public health, particularly in accurately assessing disease risk among underrepresented populations, such as African American women (31). The lack of development and validation among only European ancestries leads to various implications such as reduced predictive accuracy for non-European Populations, loss of generalizability, and implications for clinical translation. (31)

From a genetic perspective, ancestry bias in polygenic risk score (PRS) development stems from fundamental population-level differences in genomic architecture. African populations, which represent the most genetically diverse human populations due to their deep evolutionary history, have distinct allele frequency distributions, shorter linkage disequilibrium (LD) blocks, and higher levels of haplotype diversity compared to European populations (32). These characteristics pose specific challenges for GWAS-based discovery and PRS construction.

For example, the shorter LD blocks found in African ancestry genomes mean that single-nucleotide polymorphisms (SNPs) identified in European populations as being associated with disease may not adequately tag the same causal variants in African populations (32). This leads to a reduced ability of PRSs constructed from European GWAS data to capture true underlying disease risk in non-European populations, especially those of African descent. Furthermore, certain causal variants may be population-specific, present only or primarily in African populations, and thus entirely missed in discovery efforts focused exclusively on European cohorts.

Additionally, African genomes exhibit greater allelic heterogeneity, meaning that multiple different genetic variants may influence the same trait within this population (32). This diversity dilutes the contribution of any single SNP to disease risk, reducing the predictive power of PRSs that rely on fixed variant weights derived from European samples (32). Moreover, rare variants,

which may play important roles in disease etiology, are more difficult to detect in underrepresented populations due to smaller sample sizes and lack of statistical power in non-European datasets (32).

The underlying genetic architecture of complex diseases such as epithelial ovarian cancer (EOC) may also differ subtly between populations due to interactions between ancestry-specific variants and environmental exposures, which further complicates risk prediction (33). As a result, PRSs constructed from European datasets may systematically underestimate or misclassify risk among African ancestry individuals, exacerbating disparities in early detection and preventive care (33).

From a statistical perspective, ancestry bias in polygenic risk score (PRS) development arises from several methodological assumptions and limitations that reduce performance when models are transferred across populations. One major issue is the assumption of effect size homogeneity, where the effect size of a single nucleotide polymorphism (SNP) to disease risk is presumed to be constant across all ancestral groups (34). Due to genetic drift, population bottlenecks, local adaptation, and varying environmental exposures, SNP effect sizes often differ between populations, which can compromise the accuracy of PRSs when applied outside their discovery cohort (35).

Additionally, linkage disequilibrium (LD) patterns, which refer to non-random associations between alleles at different loci, vary considerably by ancestry. PRS construction relies on these LD structures to tag causal variants, but if the LD relationships differ between the discovery and target populations, the PRS will tag different genomic regions, further reducing predictive power. For example, SNPs in high LD in European populations may not be in LD in African populations, leading to misestimation of risk in the latter group (36).

Imputation quality also plays a critical role in PRS performance and is generally poorer in non-European populations due to underrepresentation in genomic reference panels. Low imputation accuracy can introduce measurement error into genotype estimates, attenuating effect sizes and degrading PRS performance (36).

Another key issue is population stratification bias, a confounding problem in GWAS where allele frequency differences between ancestral groups are correlated with differences in disease prevalence due to non-genetic (e.g., social or environmental) factors. If not properly controlled for, stratification can produce spurious associations or misestimate SNP effects, especially in admixed populations like African Americans. Although methods like principal component analysis (PCA) or mixed models are used to account for ancestry differences, residual confounding often remains and can skew PRS estimates (37).

These combined issues contribute to what is known as the polygenic score portability problem, where scores developed in European-ancestry datasets show significantly reduced performance in other populations. Empirical studies have demonstrated that PRS accuracy in African ancestry populations is often lower than in European populations for a range of complex traits and diseases (38). This drop in performance has profound implications for health equity, particularly when PRSs are used for clinical decision-making or early intervention.

Lastly, statistical power is also a concern. Smaller sample sizes in non-European populations limit the discovery of ancestry-specific variants and reduce the precision of effect size estimates. This results in PRSs with higher standard errors and lower reliability, compounding their limited generalizability and further widening the predictive gap between populations.

While it is important to note that there have been suggestive genetic loci associated with ovarian cancer, it is likely that this burden we see among African American women with epithelial ovarian cancer is likely due to a multi-faceted risks.

This paper seeks to address this critical gap by leveraging data from the African American Cancer Epidemiology Study (AACES)**,** the largest cohort of African American women with epithelial ovarian cancer in the United States. The AACES dataset, which includes robust geographical diversity and comprehensive epidemiological and genetic information, represents a unique resource for understanding the genetic drivers of EOC disparities in Black women. (8) Due to the gap in understanding the applicability and accuracy of polygenic risk scores among this diverse population, this paper will provide a unique opportunity to address the effectiveness of polygenic risk scores. We hope to enhance risk prediction for epithelial ovarian cancer among African American women, potentially improving the ability to identify high-risk individuals early. Given the multi-faceted nature of EOC risk in African American women, encompassing genetic predisposition and social determinants, this study provides an opportunity to validate PRS models and assess their clinical relevance. By validating and further analyzing an existing polygenic risk score in a cohort of African American women, we aim to assess its predictive utility, refine its application for this population, and contribute to reducing disparities in ovarian cancer outcomes. This work helps bridge critical gaps in the performance of PRS across diverse ancestries and supports more equitable precision medicine efforts.

Materials and Methods

Study Population

In this study, we used the African American Cancer Epidemiology Study (AACES), which is a multicenter population-based case-control study designed to investigate EOC among self-reported African American/Black women, a population historically underrepresented in epithelial ovarian cancer research. This cohort represents the largest study of its kind and leverages a unique design to address disparities and explore genetic, environmental, and lifestyle factors associated with EOC risk and outcomes in Black women.

For this specific analysis, the study population included participants diagnosed with Epithelial Ovarian Cancer between 2010 and 2015. Eligibility was restricted to women residing in one of the 11 geographic regions across the United States (Alabama, Georgia, New Jersey, Louisiana, South Carolina, North Carolina, Tennessee, Ohio, Illinois, Detroit Metropolitan area, and Texas) at the time of diagnosis. These regions were selected to ensure a representative sample of Black women from diverse populations, including those from both urban and rural areas, encompassing a broad range of social and environmental contexts. Eligible participants within the cohort included women who self-identified as Black or African American, newly diagnosed with EOC, and aged between 20 and 79 years at diagnosis (8).

All study procedures were approved by the Institutional Review Boards (IRBs) at each participating site. Written informed consent was obtained from all participants prior to data collection, including collection of biospecimens and genetic information. The current analysis was conducted using de-identified data from the African American Cancer Epidemiology Study (AACES), in accordance with approved data use agreements and ethical guidelines.

Data Collection

Cases were identified through rapid case ascertainment to identify and enroll eligible participants within months from the date of diagnosis, aiming to minimize survival. Controls were recruited using random digit dialing, and were frequency-matched to cases based on age and residence. This population-based case-control design enabled a robust comparison between groups with respect to genetic and epidemiological risk factors for epithelial ovarian cancer in African American women.

Statistical Analysis

This study aimed to evaluate and further analyze an existing polygenic risk score (PRS) for EOC in African American women, originally developed using genome-wide significant and suggestive loci identified by Manichaikul et al. (2020) (8,39). The PRS was derived from 24 single nucleotide polymorphisms (SNPs) previously associated with EOC in European ancestry populations and evaluated for transferability and predictive performance in women of African ancestry.

We utilized genotype data from the African American Cancer Epidemiology Study (AACES) to assess the performance of this PRS in a cohort of African American women. Association testing between the PRS and EOC risk, overall and by histologic subtype was conducted using logistic and multinomial logistic regression models, adjusting for age, ancestry principal components, and study site. Principal components was conducted using genome-wide genotype data.

Data Preparation and Quality Control

The genotype data were initially converted to PLINK binary format using PLINK v1.90. Standard quality control measures were applied to ensure data integrity. The analysis began with data preparation and quality control measures. Participants were filtered based on predefined inclusion criteria. This reduced the data from 2,257 participants to 1,008 individuals eligible for analysis.

Quality control measures were applied to ensure the validity of the dataset. Quality control procedures were applied at both the sample and variant levels. At the sample level, exclusions were based on low call rates, sex discrepancies, and high heterozygosity (39). A low call rate refers to the percentage of missing genotype calls for an individual or SNP. Individuals or SNPs with high proportions of missing data can compromise the analysis and introduce bias. Individuals with a call rate below a predefined threshold were excluded to ensure that the dataset contained reliable and complete genotype information. Sex discrepancies occur when the recorded sex of an individual does not match their inferred sex based on genetic data. Individuals with mismatched sex information were excluded due to eligibility criteria. Heterozygosity refers to the proportion of heterozygous alleles observed in an individual's genome. An unusually high level of heterozygosity can indicate potential issues such as DNA contamination, relatedness, or underlying population structure differences. Individuals with heterozygosity levels above the predefined threshold were removed. At the variant level, several criteria were applied to ensure the quality and reliability of the genetic data used in the analysis. SNPs with a minor allele frequency below 0.01 were excluded. SNPs with very low MAFs have insufficient representation in the data set, reducing their statistical power to detect meaningful associations (39). Rare SNPs can also be prone to genotyping errors, which can disproportionately impact analyses. Low

MAFs SNPs may also not be as relevant to the general population, as they might be less likely to provide information signals in association studies, such as this one. SNPs with a missingness rate above a predefined threshold of greater than 5% were excluded. Missingness refers to the proportion of samples for which a particular SNP does not have genotyped values. High rates of missing data can lead to biased estimates in statistical analyses (39). Including SNPs with many missing values can reduce the overall quality of the dataset and potentially distort results. Lastly, SNPs that deviated significantly from the Hardy-Weinberg Equilibrium were excluded. This was assessed using a chi-squared test with significance threshold ($p < 1 \times 10^{-6}$). Deviation from HWE may indicate genotyping errors, such as allele miscalls. Large deviations might suggest population stratification or non-random mating, which can confound genetic analyses. The final dataset contained 652,512 SNPs with a 100% genotyping rate.

18,007 SNPs were identified from prior genome-wide association studies (GWAS) and were used for PRS construction. To address potential errors due to duplicate SNPs a multi-step resolution process was employed. Duplicate SNPs were defined as variants with identical chromosomal positions but differing alleles. Duplicates were identified by grouping SNPs based on their chromosome and base-pair positions. The duplicate SNPs were matched against a dataset, which included validated alleles for each SNP. SNPs with allele matches to the reference were retained based on a logical comparison of allele SNPs with allele matches to the reference were retained, resulting in 225 high-quality variants. This ensured accuracy and consistency in PRS calculation.

Polygenic Risk Score Construction

The polygenic risk score (PRS) for each individual was calculated as the weighted sum of risk alleles across all selected SNPs. SNP weights were based on effect sizes reported in prior

genome-wide association studies (GWAS), primarily conducted in European ancestry populations (39). Although these effect estimates may not fully capture the genetic architecture in African American populations, they were used to enable assessment of PRS transferability and predictive performance. The PRS was calculated using the following formula:

$$PRS_i = \sum_{j=1}^{m} \beta_j * G_{ij}$$

Where: PRS_i is the polygenic risk score for individual i, m is the total number of SNPs, β_j is the effect size of SNP j, G_{ij} is the genotype of individual i for SNP j, coded as 0, 1, 2. The PRS was standardized to have a mean of zero and a standard deviation of one to facilitate interpretation and comparison across individuals.

Logistic Regression Model Construction

The Logistic regression model was employed to evaluate the association between PRS and EOC risk. The model adjusted for confounders, including age and geographic region. The logistic regression model was expressed as:

$$logit(P) = \beta_0 + \beta_1 x\ PRS + \ \beta_2\ x\ Age + \beta_3\ x\ Geographic\ Region$$

Where: P is the probability of developing EOC, β0 is the intercept, β1,β2,…,βk are the coefficients for PRS, age, and geographic region.

Odds ratios (ORs) with 95% confidence intervals were reported for the PRS and other covariates to quantify their associations with EOC risk. The model's performance was assessed using the area under the receiver operating characteristic curve (AUC), which measured its ability to distinguish between cases and controls. Calibration plots compared predicted probabilities with

observed outcomes, ensuring alignment between the model's predictions and actual data. Sensitivity analyses were conducted to ensure the robustness of the PRS. These included removing individuals with extreme PRS values, re-evaluating the model after excluding SNPs with borderline imputation quality scores, and subgroup analyses by demographic characteristics (e.g., age groups, geographic regions). These steps assessed the stability and generalizability of the PRS across diverse scenarios.

All analyses were conducted using established bioinformatics and statistical tools. Genotype processing and QC were performed in PLINK v1.90, while statistical modeling and visualization were conducted in R (version 4.0 or later). Imputation of missing genotypes was performed using reference panels, such as the 1000 Genomes Project or TOPMed, ensuring comprehensive coverage of variants relevant to African ancestry.

The results were interpreted to evaluate the PRS's effectiveness in predicting EOC risk among self-reported African American women. Key performance metrics, such as AUC and calibration plots, were used to assess the model's accuracy. The study highlighted the potential clinical utility of PRS in identifying high-risk individuals and improving personalized prevention strategies. This comprehensive analysis addressed critical gaps in genetic epidemiology and contributed to advancing equitable healthcare outcomes for underrepresented populations.

Results

Study Population

A total of 592 Black women diagnosed with epithelial ovarian cancer (EOC) were enrolled in the African American Cancer Epidemiology Study (AACES) between 2010 and 2015. The median age at diagnosis was 58 years, with 52.2% of participants diagnosed between ages

41–60. The majority of participants (67.7%) were diagnosed with high-grade serous carcinoma, and 63% were diagnosed at advanced FIGO stages (III or IV), underscoring the aggressive nature and late detection of EOC in this population. Over two-thirds of women had optimal debulking (66.6%), though 30.3% had suboptimal cytoreduction. However, data on residual disease status were missing for a notable proportion of participants. Although survival is not the outcome of interest in this analysis, residual disease status remains a key clinical variable in ovarian cancer research due to its strong association with tumor biology, treatment response, and disease burden. Therefore, the missing data may introduce bias or limit the ability to fully characterize surgical outcomes across the cohort.. Nearly 60% of participants had a BMI of 30 kg/m^2 or higher, and more than half (62.7%) reported talc use. A majority of participants (81.6%) reported having at least one full-term pregnancy. From a socioeconomic perspective, 51.0% of women had a high school education or less, 45.0% reported annual household income under $25,000, and 32.0% were uninsured or covered by Medicaid at diagnosis. About 6.6% reported a prior breast cancer diagnosis, and 23.8% had a Charlson Comorbidity Index score of 3 or more, indicating a substantial burden of chronic illness. Approximately 39% of participants were alive as of October 2024 with a median overall survival time of 4.8 years. Full baseline and clinical characteristics, including inflammatory, reproductive, and treatment-related variables, are summarized in Table 1.

Table 1. Baseline Characteristics of AACES Participants

Characteristic		**N(%)**
Age at Diagnosis	20-40	29 (4.9%)
	41-60	309 (52.2%)

	61-75	254 (42.9%)
Education	High school or less	302 (51.0%)
	Some college	106 (17.9%)
	College graduate	112 (18.9%)
	Graduate/professional school	72 (12.2%)
Annual Family Income	< $10,000	115 (21.7%)
	$10,000 - $24,999	125 (23.5%)
	$25,000 - $49,999	132 (24.9%)
	$50,000 - $74,999	79 (14.9%)
	$75,000 - $100,000	48 (9.0%)
	> $100,000	32 (6.0%)
Insurance Status	Uninsured	50 (9.3%)
	Medicaid	122 (22.6%)
	Medicare only	123 (22.8%)
	Private + Medicare	25 (4.6%)
	Private	189 (35.1%)
	Other	30 (5.6%)
FIGO Stage	Stage I	130 (25.6%)
	Stage II	57 (11.2%)
	Stage III	227 (44.7%)
	Stage IV	94 (18.5%)
Histotype	High-grade serous	397 (67.7%)

	Low-grade serous	17 (2.9%)
	Endometrioid	57 (9.7%)
	Clear cell	23 (3.9%)
	Mucinous	29 (4.9%)
	Carcinosarcoma	18 (3.1%)
	Other epithelial	45 (7.7%)
Family History	Breast Cancer	
	Yes	114 (19.5%)
	No	472 (80.5%)
	Ovarian Cancer	
	Yes	95 (14.4%)
	No	563 (85.6%)
Debulking Status	Optimal	255 (66.6%)
	Suboptimal	116 (30.3%)
	No surgery	12 (3.1%)
BMI ≥ 30 kg/m²	Yes	347 (59.0%)
Talc Use (ever)	Yes	371 (62.7%)
Parity (≥1 Pregnancy)	Yes	483 (81.6%)
Charlson Comorbidity Index ≥3	Yes	141 (23.8%)
Prior Breast Cancer	Yes	39 (6.6%)
Alive	Yes	228 (39.5%)

Median Overall Survival (years)		4.8 years

PRS Estimates

Multinomial logistic regression analyses were conducted to examine the association between standardized polygenic risk scores (PRS) and epithelial ovarian cancer (EOC) histotypes. The outcomes were categorized into high-grade serous ovarian cancer (HGSOC), other EOC histotypes, and controls. Key variables included standardized PRS, age at reference, percent African ancestry, family history of breast or ovarian cancer, and U.S. region of residence.

Using binary logistic regression, all EOC cases (regardless of histotype) were combined and compared to controls. A one-unit increase in standardized PRS was associated with a 28% increased odds of overall EOC (OR = 1.28; 95% CI: 1.11–1.47; $p = 0.0005$). In the unadjusted model, a one-unit increase in standardized PRS was associated with a 30% increased odds of HGSOC compared to controls (OR = 1.30; 95% CI: 1.12–1.51; p = 0.00035) (Table 2), indicating a strong and statistically significant association between genetic risk and disease outcome. The adjusted model, which controlled for age at reference, percent African ancestry, family history of breast and ovarian cancer, and geographic region, yielded an even stronger association (OR = 1.37; 95% CI: 1.16–1.62; p = 0.0004) (Table 2). Although age was used as a matching variable in the case–control study design, it was retained in the model due to potential residual differences introduced by the subset of participants with available genetic data. Therefore, the observed effect of age (OR = 1.04; 95% CI: 1.03–1.60; $p < 0.001$) likely reflects selection-related variability rather than confounding per se. This reinforces that while age is a

known risk factor for HGSOC, its adjustment in this model primarily accounts for analytic imbalance rather than confounding bias introduced by the study design.

Both family history variables were positively and significantly associated with HGSOC. Specifically, family history of breast cancer was associated with a 26% increase in odds of HGSOC (OR = 1.26; 95% CI: 1.02–1.98; p = 0.004), while family history of ovarian cancer conferred more than double the odds of disease (OR = 2.23; 95% CI: 1.32–3.43; p < 0.001).

In contrast, no statistically significant associations were found between PRS and non-HGSOC histotypes in either the unadjusted or adjusted models. For example, in the adjusted model, the OR for PRS predicting other histotypes was 1.16 (95% CI: 0.93–1.44; p = 0.125) (Table 2), suggesting PRS may not be a robust predictor of these subtypes.

Table 2. Adjusted Odds Ratios from Multinomial Logistic Regression Models Evaluating PRS in Relation to EOC Histotypes

Outcome Group	Variable	Odds Ration	95% CI	P-Value
HGSOC vs. Controls	PRS	1.37	1.16-1.62	0.0004
	Age at Reference	1.04	1.03-1.06	<0.001
	Family History (Breast)	1.26	1.02-1.98	0.004

	Family History (Ovarian)	2.23	1.32-3.43	<0.001
Other Histotypes vs. Control	PRS	1.16	0.93-1.44	0.125
	Age at Reference	1.00	0.98-1.02	0.974
	Family History (Breast)	1.11	0.86-1.58	0.32
	Family History (Ovarian)	1.61	1.02-3.85	0.01

To further evaluate the relationship between polygenic risk score (PRS) and risk of high-grade serous ovarian cancer (HGSOC), participants were categorized into four percentile groups: 0–80% (reference), 80–90%, 90–95%, and 95–100%. Logistic regression models were used to estimate the odds ratios for each percentile category compared to the reference group, adjusting for age at reference, percent African ancestry, family history of breast and ovarian cancer, and geographic region.

Table 3. Adjusted Odds Ratios for HGSOC by PRS Percentile

PRS Percentile Group	Odds Ratio	95% CI	P-Value
0-80% (reference)	1.00	-	-
80-90%	1.60	0.97-2.63	0.066
90-95%	2.12	1.03-4.36	0.043
95-100%	2.08	1.09-3.98	0.027

In the adjusted models, the 80–90% group showed an odds ratio of 1.60 (95% CI: 0.97–2.63; p = 0.066) (Table 3), suggesting a potential increased risk that approached statistical significance. Participants in the 90–95% group demonstrated a significantly elevated risk with an odds ratio of 2.12 (95% CI: 1.03–4.36; p = 0.043) (Table 3), while those in the highest PRS bracket (95–100%) also had significantly greater odds of HGSOC (OR = 2.08; 95% CI: 1.09–3.98; p = 0.027) (Table 3). These findings indicate a consistent trend of increasing HGSOC risk across ascending PRS categories, which is consistent with a dose-response relationship. Notably, while the 80–90% group did not reach conventional levels of statistical significance, the point estimate suggests a potentially meaningful increase in risk. The widening confidence intervals in higher PRS groups also reflect variability that may stem from sample size limitations, but the fact that both the 90–95% and 95–100% groups demonstrated statistically significant elevations in risk reinforces the pattern. This consistent elevation suggests that individuals in higher genetic risk brackets are systematically more likely to be diagnosed with high-grade serous ovarian cancer, even after adjustment potential confounders like, ancestry, and family history. The pattern

held even after controlling for multiple covariates, suggesting that PRS may serve as a meaningful stratification tool in identifying high-risk individuals.

To assess the robustness of our multinomial regression findings and align with conventional modeling approaches, we conducted two binary logistic regression models: HGSOC vs. controls and Other EOC histotypes vs. controls.

The binary logistic regression for HGSOC confirmed the association observed in the multinomial model. A unit increase in PRS was associated with 37% higher odds of HGSOC (OR = 1.37; 95% CI: 1.16–1.62; p = 0.0004), while age (OR = 1.04; 95% CI: 1.03–1.06; p < 0.001. The association between PRS and other histotypes remained nonsignificant (OR = 1.16; 95% CI: 0.93–1.44; p = 0.13).

These binary models provide consistent results compared to the multinomial approach and reinforce the specificity of the PRS effect to the HGSOC subtype. By separating the comparisons into two focused binary models, we are able to affirm that the association observed in the multinomial framework was not an artifact of outcome categorization or model specification. The replication of results in a more conventional logistic regression framework suggests that the observed effect is robust and not driven by modeling assumptions. This analytical choice also allowed for more intuitive interpretation of subtype-specific risk factors. For instance, while PRS showed a robust and statistically significant effect in predicting HGSOC, its lack of association with other histotypes strengthens the argument for biologically distinct genetic mechanisms across subtypes. Furthermore, these binary models align with conventional epidemiologic methods and offer an added layer of interpretability for clinicians and researchers evaluating polygenic risk within a more familiar framework.

Discussion

This study provides compelling evidence that higher polygenic risk scores (PRS) are associated with significantly increased odds of developing high-grade serous ovarian cancer (HGSOC) among self-reported African American/Black women. The observed association persisted even after adjusting for potential confounders such as age, ancestry, family history, and geographic region. Our findings are in line with earlier studies conducted in predominantly European ancestry populations, which identified common genetic variants as contributors to EOC heritability.

In our study, a one standard deviation increase in PRS was associated with a 38% higher odds of HGSOC (adjusted OR = 1.38; 95% CI: 1.17–1.62), underscoring the potential of PRS to capture germline susceptibility to this aggressive subtype among African American/Black women. These findings are broadly consistent with previous reports from European ancestry populations. For instance, Dareng et al. (2022) developed subtype-specific PRS using over 22,000 EOC cases and 40,000 controls, and found that a one standard deviation increase in HGSOC-specific PRS was associated with an OR of approximately 1.38 (95% CI range: ~1.3–1.5 across different cohorts), nearly identical to our estimate (30). Similarly, Flaum et al. (2020) reported that women in the top 20% of a HGSOC-specific PRS distribution had a 2.5-fold increased risk of disease compared to women in the middle quintile (OR ≈ 2.5), with continuous models showing ORs in the 1.3–1.4 range per standard deviation (33). These studies collectively demonstrate that common genetic variants confer a measurable and consistent risk for HGSOC, supporting the reproducibility of PRS effects across populations and analytic frameworks.

More importantly, our findings extend this body of work to African ancestry populations, which have been historically underrepresented in genome-wide association studies (GWAS) and

PRS development. Dareng et al. (2022) addressed this gap by evaluating both cross-ancestry and ancestry-specific PRS models for ovarian cancer in the Africa, Asia, and Europe (AANE) consortium (30). Among women of African ancestry, Dareng et al. reported that HGSOC-specific PRS were associated with increased disease risk, with effect estimates per standard deviation increase ranging from OR = 1.20 to 1.40 depending on the weighting scheme and model used (30). These effect sizes are largely consistent across ancestries and closely align with our findings in a U.S.-based cohort of African American women, where we observed an adjusted OR of 1.38 (95% CI: 1.17–1.62) per standard deviation increase in PRS. Notably, although PRS constructed from European-ancestry GWAS data had reduced predictive accuracy in non-European populations, the direction and magnitude of associations remained largely consistent across ancestries.

What distinguishes our work from prior reports is its focus on a U.S.-based, population-specific cohort of African American women from the AACES study, which allows for analysis within a relatively homogeneous sociocultural and healthcare context. Unlike many prior efforts that pooled diverse populations across continents, our study controls for environmental and structural factors more specific to Black women in the U.S. In addition, we examined PRS associations not only with overall EOC but also stratified by histologic subtype, including HGSOC and non-serous tumors, while adjusting for relevant covariates such as percent African ancestry, family history, and geographic region. This enhances the clinical interpretability of PRS in admixed African ancestry populations and informs future strategies for equitable genetic risk prediction.

Our findings align closely with those of Dareng et al., reinforcing the notion that while predictive performance may vary by ancestry due to differences in linkage disequilibrium

structure and allele frequencies, the underlying biological mechanisms captured by PRS may be shared (30). In this context, the similarity in effect size between our study (adjusted OR = 1.38) and both Dareng et al. and European studies supports the external validity of the association between common germline variants and HGSOC (30). However, our results also emphasize the importance of continuing to refine and validate PRS models specifically within African ancestry populations. Numerous studies have shown that PRS developed in European ancestry populations often perform poorly when applied to individuals of African descent due to differences in linkage disequilibrium patterns, allele frequencies, and genetic architecture (41,42). As such, increasing the representation of African ancestry populations in GWAS is essential to improve model calibration, reduce prediction bias, and enhance the clinical utility of PRS in diverse populations. Recent efforts have demonstrated that ancestry-informed models and multi-ancestry GWAS can meaningfully improve risk prediction in non-European groups (41,42), underscoring the need for continued investment in inclusive genomic research to ensure equitable translation of PRS into clinical and public health contexts.

These results suggest that polygenic risk scores (PRS) can be a valuable tool for risk stratification among African American women, a population historically underrepresented in genetic risk modeling. Importantly, the ability of PRS to stratify individuals by relative genetic susceptibility offers a mechanism to identify those at substantially higher risk of developing high-grade serous ovarian cancer (HGSOC). As an example, women in the top 5% of the PRS distribution had more than twice the odds of HGSOC diagnosis compared to those in the bottom 80%, even after adjusting for known epidemiologic risk factors. This level of stratification highlights the clinical relevance of PRS and supports their use in both individual risk counseling and population-level risk prediction frameworks.

The potential to identify women at highest risk before disease onset presents a critical opportunity to improve outcomes through more proactive management. Specifically, PRS could be used to prioritize high-risk individuals for enhanced clinical surveillance, such as more frequent pelvic imaging, blood-based biomarkers like CA125, or enrollment in early detection trials. Additionally, knowledge of elevated genetic risk may facilitate shared decision-making regarding preventive strategies, such as risk-reducing salpingo-oophorectomy, particularly in cases with intermediate family history but elevated PRS (43).

Incorporating PRS into clinical risk assessments may also help in selecting individuals for intensive screening or chemoprevention trials, especially since ovarian cancer often presents with vague, nonspecific symptoms and lacks effective screening modalities for the general population (12). The ability to improve early detection, when ovarian cancer is most treatable, could have a significant impact on survival outcomes, particularly for African American/Black women, who have historically experienced worse prognoses and limited access to early intervention strategies.

However, one important limitation of this analysis is the inability to account for BRCA1 and BRCA2 mutation status, which are high-penetrance, Mendelian-inherited risk factors for HGSOC. These mutations confer substantially higher lifetime ovarian cancer risks and are especially relevant for risk stratification and clinical decision-making (12). While we adjusted for self-reported family history of breast and ovarian cancer, family history is an imperfect proxy for underlying pathogenic variants. In the context of regression modeling, adjusting for family history means we are conditioning on it, that is, we are estimating the association between the polygenic risk score (PRS) and cancer risk independent of any effect attributable to family history. This does not incorporate family history into an integrated risk prediction model, but

rather statistically accounts for its potential confounding or mediating role. Consequently, the PRS estimates reflect associations over and above the contribution of familial risk, allowing us to isolate the effect of common, low-penetrance variants captured by the PRS It may fail to capture mutation carriers from small or male-dominated families, individuals unaware of their relatives' diagnoses, or those who do not meet traditional clinical testing guidelines. Consequently, the observed associations between PRS and HGSOC risk may be partially confounded or diluted by unmeasured BRCA mutation status. Future analyses should aim to jointly model both PRS and monogenic risk variants to better disentangle their independent and interactive contributions to disease risk.

Nonetheless, our findings support the continued integration of polygenic tools into precision public health frameworks, particularly for addressing persistent disparities in ovarian cancer outcomes. As PRS models improve and become more tailored to diverse populations, they could complement existing risk models to identify high-risk individuals who might benefit from enhanced surveillance or preventive measures. Importantly, implementation must be guided by attention to clinical validity, utility, and ethical concerns, especially for African American/Black populations, who face systemic barriers such as limited access to genetic counseling, historical mistrust of medical institutions, and inconsistent insurance coverage for genetic testing (12). Without addressing these structural inequities, PRS-based initiatives risk reinforcing rather than reducing disparities in cancer prevention and care.

Our findings also highlight the histotype-specific nature of genetic risk for epithelial ovarian cancer (EOC). The absence of a statistically significant association between polygenic risk scores (PRS) and non-HGSOC histotypes reinforces the importance of stratifying by histologic subtype in genetic risk prediction studies. This specificity likely reflects the distinct

molecular, cellular, and etiologic characteristics of each EOC subtype, which are increasingly understood to represent biologically separate diseases rather than variants along a single continuum.

These molecular differences also align with divergent risk factor profiles. For instance, endometriosis is a strong risk factor for clear cell and endometrioid tumors but is not associated with HGSOC. Conversely, BRCA1/2 pathogenic variants strongly increase risk for HGSOC, but not for mucinous or low-grade serous cancers. Hormonal, reproductive, and lifestyle risk factors such as parity, oral contraceptive use, and obesity also exhibit subtype-specific associations (9).

Therefore, it is not surprising that a PRS developed primarily from GWAS of HGSOC, where most of the heritability signal has been concentrated, would not perform well in predicting risk for other histotypes with different genetic underpinnings. The lack of association in our study supports the hypothesis that the genetic architecture of ovarian cancer is subtype-specific and that PRS models must be histotype-targeted to be both biologically meaningful and clinically useful.

Strengths of this study include the use of a large, geographically diverse cohort of African American women and a focus on a specific histologic subtype with high clinical relevance. The African American Cancer Epidemiology Study (AACES) remains one of the few datasets that captures the genetic and epidemiologic diversity of Black women with epithelial ovarian cancer. Leveraging this unique resource enabled an in-depth analysis of polygenic risk scores (PRS) within a historically understudied population, enhancing the relevance and potential impact of our findings.

The use of multinomial logistic regression allowed for simultaneous estimation of associations across multiple histologic subtypes of EOC, offering refined and specific insights into PRS associations for HGSOC versus other subtypes. This is particularly valuable in ovarian cancer research, where etiologic heterogeneity across histotypes may obscure findings in traditional binary analyses. Furthermore, the addition of binary logistic regression as a sensitivity analysis strengthens the rigor of our findings and enhances interpretability audiences who are more familiar with dichotomous outcome models. The consistency of effect estimates across modeling approaches increases confidence in the robustness of the observed associations

Another methodological strength lies in our incorporation of multiple covariates including age, percent African ancestry, and family history of breast and ovarian cancer into the adjusted models. These variables reflect known confounders in ovarian cancer epidemiology, and their inclusion helps isolate the independent contribution of PRS to disease risk. Additionally, our stratified and percentile-based analyses offer a nuanced understanding of the distribution and impact of genetic risk across population subgroups.

Nonetheless, limitations exist. Wide confidence intervals, particularly in regional and ancestry-adjusted models, suggest that some of the estimates may be imprecise due to limited sample sizes in specific strata. While this does not invalidate the findings, it does warrant caution in over-interpreting marginal effects. Moreover, the PRS model used in this study was derived from multi-ancestry genome-wide association studies (GWAS), which may still limit predictive performance in African ancestry populations despite our use of ancestry-adjusted models. Future efforts to recalibrate PRS using African ancestry-specific summary statistics are needed to enhance accuracy. While recalibrating polygenic risk scores (PRS) using African ancestry-specific summary statistics is a promising direction, it currently remains challenging but

essential, and its success will depend on several key developments. To address this gap, several strategies can be pursued to improve the accuracy and applicability of PRS for individuals of African ancestry. First, expanding the representation of African ancestry individuals in large-scale genome-wide association studies (GWAS) and meta-analyses is essential. Although initiatives such as H3Africa and TOPMed have begun to address this imbalance, further expansion, particularly in cancer-focused studies, is critically needed (44). Second, trans-ethnic or multi-ancestry PRS development approaches can be employed. These models incorporate differences in linkage disequilibrium (LD) structure across populations and use methods such as PRS-CSx or multi-ancestry LDpred2 to improve score transferability and predictive performance (45). Third, fine-mapping and functional annotation efforts can leverage the greater genetic diversity of African populations to more precisely identify causal variants and prioritize biologically relevant loci, even with smaller sample sizes (46). Fourth, advanced statistical frameworks, including machine learning and Bayesian shrinkage-based methods, offer the flexibility to integrate cross-ancestry data while accommodating LD structure specific to African ancestry populations. Lastly, equitable investment in research infrastructure is vital. This includes ensuring access to genomic data, sequencing technologies, and computational resources, particularly for institutions and investigators working with African ancestry cohorts in both the U.S. and global contexts. Together, these efforts are crucial for building accurate and inclusive PRS tools that support equitable implementation of genomic medicine. In addition, limited imputation accuracy among individuals of African ancestry, stemming from lower linkage disequilibrium and underrepresentation in reference panels, remains a key technical barrier that can compromise SNP-level estimates and, by extension, PRS performance.

As with many case-control designs, recall and selection bias may also be concerns. Certain exposures may modify the penetrance of polygenic risk and could be crucial in understanding the full landscape of ovarian cancer etiology, particularly among African American/Black women. However, detecting such gene–environment interactions typically requires very large, well-powered epidemiologic studies with detailed exposure assessment, resources that are not always available.

To address this gap, experimental and molecular studies can serve as valuable complements to population-based research. For example, in vitro functional assays and cell line models can help identify how specific SNPs or genetic pathways respond to environmental stimuli such as endocrine-disrupting chemicals, oxidative stress, or hormonal exposures (47). Organoid models derived from fallopian tube or ovarian epithelium may offer insight into how cells with different genetic risk profiles react to environmental triggers at a molecular level (47). Animal models, such as genetically engineered mice carrying risk variants in genes like BRCA1/2 or TP53, could be used to explore how diet, inflammation, or carcinogen exposure modulates cancer risk in a controlled setting (48). Additionally, epigenomic and transcriptomic profiling of ovarian tumors stratified by polygenic risk or environmental exposures could reveal gene expression signatures or methylation patterns indicative of GxE mechanisms (48).

Together, these approaches can help narrow the search space for population-level GxE studies and inform the development of integrative, biologically grounded risk prediction models. Future efforts should prioritize multidisciplinary collaborations that combine molecular epidemiology, genomics, and experimental systems to better understand the interplay between inherited genetic risk and modifiable exposures in ovarian cancer development.

Additionally, while our study adjusted for ancestry proportion, population stratification remains a potential source of residual confounding in genetic association studies. The potential for misclassification of histologic subtypes, particularly in cases diagnosed at community hospitals with limited pathology review, may also introduce classification bias. This limitation is very unlikely as the study design accounts for centralized pathology review by an expert pathologist. Another challenge is the static nature of PRS, which does not account for temporal changes in risk factors or exposures, nor does it incorporate dynamic health behaviors over the life course. Finally, the clinical implementation of PRS remains limited by challenges in translation to actionable interventions, particularly in the absence of guidelines specific to PRS use in ovarian cancer prevention and screening among underrepresented groups.

Future work should focus on validating this PRS model in independent African ancestry cohorts and exploring the integration of genetic, environmental, and social risk factors into a unified risk prediction framework. External validation in diverse datasets is crucial for assessing the model's generalizability, refining predictive accuracy, and identifying population-specific thresholds for clinical use. In addition to replication, efforts should explore calibration and re-weighting of PRS using African ancestry-specific effect sizes derived from larger GWAS consortia (40). However, this remains a highly challenging goal. As highlighted by Manichaikul et al., the development and validation of PRS in African ancestry populations are hindered by limited representation in GWAS, differences in linkage disequilibrium structure, allele frequency divergence, and lower trans-ancestry portability of risk estimates. These factors can lead to reduced discrimination, poor calibration, and ultimately limited clinical utility of PRS if not addressed through ancestry-aware modeling approaches. Although, achieving this goal will require substantial investment in the recruitment of African ancestry participants, increased

collaboration across consortia, and the use of statistical methods that explicitly account for ancestry-specific genomic architecture.

Investigating gene–environment interactions could uncover additional modifiers of genetic risk and further personalize prevention strategies. For instance, the impact of modifiable factors such as obesity, hormone therapy, and reproductive history may differ by PRS strata, providing insight into risk mitigation strategies tailored to genetic susceptibility (31). Studies that incorporate exposome data, neighborhood-level socioeconomic factors, and life-course exposures could lead to more holistic and actionable risk models (31).

Finally, community-engaged research approaches should be prioritized to ensure that PRS implementation aligns with the values and preferences of African American women. This includes qualitative work on patient acceptability, preferences for return of results, and strategies for culturally sensitive communication of genetic risk. Ultimately, the goal is to move toward equitable precision medicine. Developing robust, ancestry-informed PRS models represents an essential step toward that objective. In tandem, addressing structural barriers to care, such as access, insurance, and provider bias, will be necessary to translate genetic risk insights into improved health outcomes for African American women.

References

1. Arora T, Mullangi S, Vadakekut ES, et al. Epithelial Ovarian Cancer. [Updated 2024 May 6]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK567760/

2. American Cancer Society. *Ovarian Cancer Early Detection, Diagnosis, and Staging.* Available at: https://www.cancer.org/cancer/ovarian-cancer/detection-diagnosis-staging.html

3. American Cancer Society. (2025). Ovarian cancer statistics: How common is ovarian cancer. Ovarian Cancer Statistics | How Common is Ovarian Cancer | American Cancer Society. https://www.cancer.org/cancer/types/ovarian-cancer/about/key-statistics.html

4. American Cancer Society. (2025a). Ovarian cancer risk factors: Risk factors for ovarian cancer. Risk Factors for Ovarian Cancer | American Cancer Society. https://www.cancer.org/cancer/types/ovarian-cancer/causes-risks-prevention/risk-factors.html

5. NIH. (2025). Cancer of the ovary - cancer stat facts. SEER. https://seer.cancer.gov/statfacts/html/ovary.html

6. Bristow RE, Powell MA, Al-Hammadi N, Chen L, Miller JP, Roland PY, Mutch DG, Cliby WA. Disparities in ovarian cancer care quality and survival according to race and socioeconomic status. J Natl Cancer Inst. 2013 Jun 5;105(11):823-32. doi: 10.1093/jnci/djt065. Epub 2013 Mar 28. PMID: 23539755; PMCID: PMC3691950.

7. Mei S, Chelmow D, Gecsi K, Barkley J, Barrows E, Brooks R, Huber-Keener K, Jeudy M, O'Hara JS, Burke W. Health Disparities in Ovarian Cancer: Report From the Ovarian Cancer Evidence Review Conference. Obstet Gynecol. 2023 Jul 1;142(1):196-210. doi:

10.1097/AOG.0000000000005210. Epub 2023 Jun 7. PMID: 37348095; PMCID: PMC10278570.

8. Schildkraut JM, Johnson C, Dempsey LF, Qin B, Terry P, Akonde M, Peters ES, Mandle H, Cote ML, Peres L, Moorman P, Schwartz AG, Epstein M, Marks J, Bondy M, Lawson AB, Alberg AJ, Bandera EV. Survival of epithelial ovarian cancer in Black women: a society to cell approach in the African American cancer epidemiology study (AACES). Cancer Causes Control. 2023 Mar;34(3):251-265. doi: 10.1007/s10552-022-01660-0. Epub 2022 Dec 15. PMID: 36520244; PMCID: PMC9753020.

9. Gnagy S, Ming EE, Devesa SS, Hartge P, Whittemore AS. Declining ovarian cancer rates in U.S. women in relation to parity and oral contraceptive use. Epidemiology. 2000 Mar;11(2):102-5. doi: 10.1097/00001648-200003000-00004. PMID: 11021604.

10. Varier L, Sundaram SM, Gamit N, Warrier S. An Overview of Ovarian Cancer: The Role of Cancer Stem Cells in Chemoresistance and a Precision Medicine Approach Targeting the Wnt Pathway with the Antagonist sFRP4. Cancers (Basel). 2023 Feb 17;15(4):1275. doi: 10.3390/cancers15041275. PMID: 36831617; PMCID: PMC9954718.

11. Friedenreich CM, Ryder-Burbidge C, McNeil J. Physical activity, obesity and sedentary behavior in cancer etiology: epidemiologic evidence and biologic mechanisms. Mol Oncol. 2021 Mar;15(3):790-800. doi: 10.1002/1878-0261.12772. Epub 2020 Aug 18. PMID: 32741068; PMCID: PMC7931121.

12. Ali AT, Al-Ani O, Al-Ani F. Epidemiology and risk factors for ovarian cancer. Prz Menopauzalny. 2023 Jun;22(2):93-104. doi: 10.5114/pm.2023.128661. Epub 2023 Jun 14. PMID: 37674925; PMCID: PMC10477765.

13. American Cancer Society. (2025a). How to check for ovarian cancer: Ovarian cancer screening. How to Check for Ovarian Cancer | Ovarian Cancer Screening | American Cancer Society. https://www.cancer.org/cancer/types/ovarian-cancer/detection-diagnosis-staging/detection.html

14. Köbel M, Kang EY. The Evolution of Ovarian Carcinoma Subclassification. Cancers (Basel). 2022 Jan 14;14(2):416. doi: 10.3390/cancers14020416. PMID: 35053578; PMCID: PMC8774015.

15. Zhu M, Sun S, Huang L, Chen M, Cai J, Wang Z, Cai L. Case report: diagnosis and treatment of advanced high-grade serous ovarian carcinoma aided by ^{68}Ga-FAPI PET/MR scan. Am J Nucl Med Mol Imaging. 2024 Feb 20;14(1):72-77. doi: 10.62347/LHUD5322. PMID: 38500744; PMCID: PMC10944375.

16. Kim J, Park EY, Kim O, Schilder JM, Coffey DM, Cho CH, Bast RC Jr. Cell Origins of High-Grade Serous Ovarian Cancer. Cancers (Basel). 2018 Nov 12;10(11):433. doi: 10.3390/cancers10110433. PMID: 30424539; PMCID: PMC6267333.

17. Wilczyński J, Paradowska E, Wilczyński M. High-Grade Serous Ovarian Cancer-A Risk Factor Puzzle and Screening Fugitive. Biomedicines. 2024 Jan 19;12(1):229. doi: 10.3390/biomedicines12010229. PMID: 38275400; PMCID: PMC10813374.

18. Ilenkovan, N., Ryan, N. A. J., Roxburgh, P., Bell, S., & Gourley, C. (2025). Controversies in the management of serous borderline tumors and low-grade serous ovarian cancer. International Journal of Gynecological Cancer, 35(3), Article 101673. https://doi.org/10.1016/j.ijgc.2025.101673

19. Chen S, Li Y, Qian L, Deng S, Liu L, Xiao W, Zhou Y. A Review of the Clinical Characteristics and Novel Molecular Subtypes of Endometrioid Ovarian Cancer. Front Oncol. 2021 Jun 3;11:668151. doi: 10.3389/fonc.2021.668151. PMID: 34150634; PMCID: PMC8210668.

20. Zhu C, Xu Z, Zhang T, Qian L, Xiao W, Wei H, Jin T, Zhou Y. Updates of Pathogenesis, Diagnostic and Therapeutic Perspectives for Ovarian Clear Cell Carcinoma. J Cancer. 2021 Feb 22;12(8):2295-2316. doi: 10.7150/jca.53395. PMID: 33758607; PMCID: PMC7974897.

21. Babaier, A., & Ghatage, P. (2020). Mucinous Cancer of the Ovary: Overview and Current Status. *Diagnostics*, *10*(1), 52. https://doi.org/10.3390/diagnostics10010052

22. Arter, Z.L., Desmond, D., Berenberg, J.L. *et al.* Epithelial ovarian cancer survival by race and ethnicity in an equal-access healthcare population. *Br J Cancer* **130**, 108–113 (2024). https://doi.org/10.1038/s41416-023-02471-z

23. Peres LC, Risch H, Terry KL, Webb PM, Goodman MT, Wu AH, Alberg AJ, Bandera EV, Barnholtz-Sloan J, Bondy ML, Cote ML, Funkhouser E, Moorman PG, Peters ES, Schwartz AG, Terry PD, Manichaikul A, Abbott SE, Camacho F, Jordan SJ, Nagle CM; Australian Ovarian Cancer Study Group; Rossing MA, Doherty JA, Modugno F, Moysich K, Ness R, Berchuck A, Cook L, Le N, Brooks-Wilson A, Sieh W, Whittemore A, McGuire V, Rothstein J, Anton-Culver H, Ziogas A, Pearce CL, Tseng C, Pike M, Schildkraut JM; African American Cancer Epidemiology Study and the Ovarian Cancer Association Consortium. Racial/ethnic differences in the epidemiology of ovarian cancer: a pooled analysis of 12 case-control studies. Int J Epidemiol. 2018 Apr 1;47(2):460-472. doi: 10.1093/ije/dyx252. Erratum in: Int J Epidemiol. 2018 Jun 1;47(3):1011. doi: 10.1093/ije/dyy054. PMID: 29211900; PMCID: PMC5913601.

24. Passaro A, Al Bakir M, Hamilton EG, Diehn M, André F, Roy-Chowdhuri S, Mountzios G, Wistuba II, Swanton C, Peters S. Cancer biomarkers: Emerging trends and clinical implications for personalized treatment. Cell. 2024 Mar 28;187(7):1617-1635. doi: 10.1016/j.cell.2024.02.041. PMID: 38552610; PMCID: PMC7616034.

25. Kim SI, Ha HI, Eoh KJ, Lim J, Won YJ, Lim MC. Trends in the Incidence and Survival Rates of Primary Ovarian Clear Cell Carcinoma Compared to Ovarian Serous Carcinoma in Korea. Front Oncol. 2022 Apr 7;12:874037. doi: 10.3389/fonc.2022.874037. PMID: 35463304; PMCID: PMC9021727.

26. Rodriguez VE, Tanjasiri SP, Ro A, Hoyt MA, Bristow RE, LeBrón AMW. Trends in endometrial cancer incidence in the United States by race/ethnicity and age of onset from 2000 to 2019. Am J Epidemiol. 2025 Jan 8;194(1):103-113. doi: 10.1093/aje/kwae178. PMID: 38960701; PMCID: PMC12034838.

27. Punzón-Jiménez, P., Lago, V., Domingo, S., Simón, C., & Mas, A. (2022). Molecular Management of High-Grade Serous Ovarian Carcinoma. *International Journal of Molecular Sciences*, *23*(22), 13777. https://doi.org/10.3390/ijms232213777

28. Towner M, Kim JJ, Simon MA, Matei D, Roque D. Disparities in gynecologic cancer incidence, treatment, and survival: a narrative review of outcomes among black and white women in the United States. Int J Gynecol Cancer. 2022 Jul 4;32(7):931-938. doi: 10.1136/ijgc-2022-003476. PMID: 35523443; PMCID: PMC9509411.

29. Lai, PH., Tyrer, J.P., Pharoah, P. *et al.* Characterizing somatic mutations in ovarian cancer germline risk regions. *Commun Biol* **8**, 676 (2025). https://doi.org/10.1038/s42003-025-08072-1

30. Dareng EO, Tyrer JP, Barnes DR, Jones MR, Yang X, Aben KKH, Adank MA, Agata S, Andrulis IL, Anton-Culver H, Antonenkova NN, Aravantinos G, Arun BK, Augustinsson A, Balmaña J, Bandera EV, Barkardottir RB, Barrowdale D, Beckmann MW, Beeghly-Fadiel A, Benitez J, Bermisheva M, Bernardini MQ, Bjorge L, Black A, Bogdanova NV, Bonanni B, Borg A, Brenton JD, Budzilowska A, Butzow R, Buys SS, Cai H, Caligo MA, Campbell I, Cannioto R, Cassingham H, Chang-Claude J, Chanock SJ, Chen K, Chiew YE, Chung WK, Claes KBM, Colonna S; GEMO Study Collaborators; GC-HBOC Study Collaborators; EMBRACE Collaborators; Cook LS, Couch FJ, Daly MB, Dao F, Davies E, de la Hoya M, de Putter R, Dennis J, DePersia A, Devilee P, Diez O, Ding YC, Doherty JA, Domchek SM, Dörk T, du Bois A, Dürst M, Eccles DM, Eliassen HA, Engel C, Evans GD, Fasching PA, Flanagan JM, Fortner RT, Machackova E, Friedman E, Ganz PA, Garber J, Gensini F, Giles GG, Glendon G, Godwin AK, Goodman MT, Greene MH, Gronwald J; OPAL Study Group; AOCS Group; Hahnen E, Haiman CA, Håkansson N, Hamann U, Hansen TVO, Harris HR, Hartman M, Heitz F, Hildebrandt MAT, Høgdall E, Høgdall CK, Hopper JL, Huang RY, Huff C, Hulick PJ, Huntsman DG, Imyanitov EN; KConFab Investigators; HEBON Investigators; Isaacs C, Jakubowska A, James PA, Janavicius R, Jensen A, Johannsson OT, John EM, Jones ME, Kang D, Karlan BY, Karnezis A, Kelemen LE, Khusnutdinova E, Kiemeney LA, Kim BG, Kjaer SK, Komenaka I, Kupryjanczyk J, Kurian AW, Kwong A, Lambrechts D, Larson MC, Lazaro C, Le ND, Leslie G, Lester J, Lesueur F, Levine DA, Li L, Li J, Loud JT, Lu KH, Lubiński J, Mai PL, Manoukian S,

Marks JR, Matsuno RK, Matsuo K, May T, McGuffog L, McLaughlin JR, McNeish IA, Mebirouk N, Menon U, Miller A, Milne RL, Minlikeeva A, Modugno F, Montagna M, Moysich KB, Munro E, Nathanson KL, Neuhausen SL, Nevanlinna H, Yie JNY, Nielsen HR, Nielsen FC, Nikitina-Zake L, Odunsi K, Offit K, Olah E, Olbrecht S, Olopade OI, Olson SH, Olsson H, Osorio A, Papi L, Park SK, Parsons MT, Pathak H, Pedersen IS, Peixoto A, Pejovic T, Perez-Segura P, Permuth JB, Peshkin B, Peterlongo P, Piskorz A, Prokofyeva D, Radice P, Rantala J, Riggan MJ, Risch HA, Rodriguez-Antona C, Ross E, Rossing MA, Runnebaum I, Sandler DP, Santamariña M, Soucy P, Schmutzler RK, Setiawan VW, Shan K, Sieh W, Simard J, Singer CF, Sokolenko AP, Song H, Southey MC, Steed H, Stoppa-Lyonnet D, Sutphen R, Swerdlow AJ, Tan YY, Teixeira MR, Teo SH, Terry KL, Terry MB; OCAC Consortium; CIMBA Consortium; Thomassen M, Thompson PJ, Thomsen LCV, Thull DL, Tischkowitz M, Titus L, Toland AE, Torres D, Trabert B, Travis R, Tung N, Tworoger SS, Valen E, van Altena AM, van der Hout AH, Van Nieuwenhuysen E, van Rensburg EJ, Vega A, Edwards DV, Vierkant RA, Wang F, Wappenschmidt B, Webb PM, Weinberg CR, Weitzel JN, Wentzensen N, White E, Whittemore AS, Winham SJ, Wolk A, Woo YL, Wu AH, Yan L, Yannoukakos D, Zavaglia KM, Zheng W, Ziogas A, Zorn KK, Kleibl Z, Easton D, Lawrenson K, DeFazio A, Sellers TA, Ramus SJ, Pearce CL, Monteiro AN, Cunningham J, Goode EL, Schildkraut JM, Berchuck A, Chenevix-Trench G, Gayther SA, Antoniou AC, Pharoah PDP. Polygenic risk modeling for prediction of epithelial ovarian cancer risk. Eur J Hum Genet. 2022 Mar;30(3):349-362. doi: 10.1038/s41431-021-00987-7. Epub 2022 Jan 14. Erratum in: Eur J Hum Genet. 2022 May;30(5):630-631. doi: 10.1038/s41431-022-01085-y. PMID: 35027648; PMCID: PMC8904525.

31. Wang Y, Zhu M, Ma H, Shen H. Polygenic risk scores: the future of cancer risk prediction, screening, and precision prevention. Med Rev (2021). 2022 Feb 14;1(2):129-149. doi: 10.1515/mr-2021-0025. PMID: 37724297; PMCID: PMC10471106.

32. Campbell MC, Tishkoff SA. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. Annu Rev Genomics Hum Genet. 2008;9:403-33. doi: 10.1146/annurev.genom.9.081307.164258. PMID: 18593304; PMCID: PMC2953791.

33. Flaum N, Crosbie EJ, Edmondson RJ, Smith MJ, Evans DG. Epithelial ovarian cancer risk: A review of the current genetic landscape. Clin Genet. 2020 Jan;97(1):54-63. doi: 10.1111/cge.13566. Epub 2019 May 29. PMID: 31099061; PMCID: PMC7017781.

34. Patel RA, Musharoff SA, Spence JP, Pimentel H, Tcheandjieu C, Mostafavi H, Sinnott-Armstrong N, Clarke SL, Smith CJ; V.A. Million Veteran Program; Durda PP, Taylor KD, Tracy R, Liu Y, Johnson WC, Aguet F, Ardlie KG, Gabriel S, Smith J, Nickerson DA, Rich SS, Rotter JI, Tsao PS, Assimes TL, Pritchard JK. Genetic interactions drive heterogeneity in causal variant effect sizes for gene expression and complex traits. Am J Hum Genet. 2022 Jul 7;109(7):1286-1297. doi: 10.1016/j.ajhg.2022.05.014. Epub 2022 Jun 17. PMID: 35716666; PMCID: PMC9300878.

35. Slunecka JL, van der Zee MD, Beck JJ, Johnson BN, Finnicum CT, Pool R, Hottenga JJ, de Geus EJC, Ehli EA. Implementation and implications for polygenic risk scores in healthcare. Hum Genomics. 2021 Jul 20;15(1):46. doi: 10.1186/s40246-021-00339-y. PMID: 34284826; PMCID: PMC8290135.

36. Fatumo S, Sathan D, Samtal C, Isewon I, Tamuhla T, Soremekun C, Jafali J, Panji S, Tiffin N, Fakim YJ. Polygenic risk scores for disease risk prediction in Africa: current challenges and future directions. Genome Med. 2023 Oct 30;15(1):87. doi: 10.1186/s13073-023-01245-9. PMID: 37904243; PMCID: PMC10614359.

37. Peterson RE, Kuchenbaecker K, Walters RK, Chen CY, Popejoy AB, Periyasamy S, Lam M, Iyegbe C, Strawbridge RJ, Brick L, Carey CE, Martin AR, Meyers JL, Su J, Chen J, Edwards AC, Kalungi A, Koen N, Majara L, Schwarz E, Smoller JW, Stahl EA, Sullivan PF, Vassos E, Mowry B, Prieto ML, Cuellar-Barboza A, Bigdeli TB, Edenberg HJ, Huang H, Duncan LE. Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. Cell. 2019 Oct 17;179(3):589-603. doi: 10.1016/j.cell.2019.08.051. Epub 2019 Oct 10. PMID: 31607513; PMCID: PMC6939869.

38. Majara L, Kalungi A, Koen N, Tsuo K, Wang Y, Gupta R, Nkambule LL, Zar H, Stein DJ, Kinyanda E, Atkinson EG, Martin AR. Low and differential polygenic score generalizability among African populations due largely to genetic diversity. HGG Adv. 2023 Feb 13;4(2):100184. doi: 10.1016/j.xhgg.2023.100184. PMID: 36873096; PMCID: PMC9982687.

39. Manichaikul A, Peres LC, Wang XQ, Barnard ME, Chyn D, Sheng X, Du Z, Tyrer J, Dennis J, Schwartz AG, Cote ML, Peters E, Moorman PG, Bondy M, Barnholtz-Sloan JS, Terry P, Alberg AJ, Bandera EV, Funkhouser E, Wu AH, Pearce CL, Pike M, Setiawan VW, Haiman CA; African American Breast Cancer Consortium (AABC); African Ancestry Prostate Cancer Consortium (AAPC); Palmer JR, LeMarchand L, Wilkens LR, Berchuck A, Doherty JA, Modugno F, Ness R, Moysich K, Karlan BY, Whittemore AS, McGuire V, Sieh W, Lawrenson K,

Gayther S, Sellers TA, Pharoah P, Schildkraut JM; African American Cancer Epidemiology Study (AACES) and the Ovarian Cancer Association Consortium (OCAC). Identification of novel epithelial ovarian cancer loci in women of African ancestry. Int J Cancer. 2020 Jun 1;146(11):2987-2998. doi: 10.1002/ijc.32653. Epub 2019 Oct 8. PMID: 31469419; PMCID: PMC7523187.

40. Flaum N, Crosbie EJ, Edmondson RJ, Smith MJ, Evans DG. Epithelial ovarian cancer risk: A review of the current genetic landscape. Clin Genet. 2020 Jan;97(1):54-63. doi: 10.1111/cge.13566. Epub 2019 May 29. PMID: 31099061; PMCID: PMC7017781.

41. Martin, A.R., Kanai, M., Kamatani, Y. *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* **51**, 584–591 (2019). https://doi.org/10.1038/s41588-019-0379-x

42. Duncan, L., Shen, H., Gelaye, B. *et al.* Analysis of polygenic risk score usage and performance in diverse human populations. *Nat Commun* **10**, 3328 (2019). https://doi.org/10.1038/s41467-019-11112-0

43. Cherry C, Ropka M, Lyle J, Napolitano L, Daly MB. Understanding the needs of women considering risk-reducing salpingo-oophorectomy. Cancer Nurs. 2013 May-Jun;36(3):E33-8. doi: 10.1097/NCC.0b013e3182642cb5. PMID: 22964868; PMCID: PMC3637391.

44. Amaeshi LC, Okunade KS, Anorlu RI. Current landscape of cancer genomics research in sub-Saharan Africa - a review of literature. Front Oncol. 2025 Apr 17;15:1512005. doi: 10.3389/fonc.2025.1512005. PMID: 40313245; PMCID: PMC12043467.

45. Ruan Y, Lin YF, Feng YA, Chen CY, Lam M, Guo Z; Stanley Global Asia Initiatives; He L, Sawa A, Martin AR, Qin S, Huang H, Ge T. Improving polygenic prediction in ancestrally diverse populations. Nat Genet. 2022 May;54(5):573-580. doi: 10.1038/s41588-022-01054-7. Epub 2022 May 5. Erratum in: Nat Genet. 2022 Aug;54(8):1259. doi: 10.1038/s41588-022-01144-6. PMID: 35513724; PMCID: PMC9117455.

46. Yuan K, Longchamps RJ, Pardiñas AF, Yu M, Chen TT, Lin SC, Chen Y, Lam M, Liu R, Xia Y, Guo Z, Shi W, Shen C; Schizophrenia Workgroup of Psychiatric Genomics Consortium; Daly MJ, Neale BM, Feng YA, Lin YF, Chen CY, O'Donovan MC, Ge T, Huang H. Fine-mapping across diverse ancestries drives the discovery of putative causal variants underlying human complex traits and diseases. Nat Genet. 2024 Sep;56(9):1841-1850. doi: 10.1038/s41588-024-01870-z. Epub 2024 Aug 26. PMID: 39187616; PMCID: PMC11888783.

47. Rotroff DM, Dix DJ, Houck KA, Knudsen TB, Martin MT, McLaurin KW, Reif DM, Crofton KM, Singh AV, Xia M, Huang R, Judson RS. Using in vitro high throughput screening assays to identify potential endocrine-disrupting chemicals. Environ Health Perspect. 2013 Jan;121(1):7-14. doi: 10.1289/ehp.1205065. Epub 2012 Sep 28. PMID: 23052129; PMCID: PMC3546348.

48. Walrath JC, Hawes JJ, Van Dyke T, Reilly KM. Genetically engineered mouse models in cancer research. Adv Cancer Res. 2010;106:113-64. doi: 10.1016/S0065-230X(10)06004-5. PMID: 20399958; PMCID: PMC3533445.