HHS Public Access

Author manuscript

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2024 May 01.

Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2023 November 01; 32(11): 1558–1563.

doi:10.1158/1055-9965.EPI-23-0174.

Assessing a Polygenic Risk Score for Lung Cancer Susceptibility in Non-Hispanic White and Black Populations

Matthew R. Trendowski[1,2,*], **Christine M. Lusk**[1,2,*], **Angela S. Wenzlaff**[1,2], **Christine Neslund-Dudas**[3,4], **Shirish M. Gadgeel**[4], **Ayman O. Soubani**[5,2], **Ann G. Schwartz**[1,2]

[1]Department of Oncology, Wayne State University School of Medicine, Detroit, MI, USA.

[2]Karmanos Cancer Institute, Detroit, MI, USA.

[3]Department of Public Health Sciences, Henry Ford Health, Detroit, MI, USA

[4]Henry Ford Cancer Institute, Henry Ford Health, Detroit, MI, USA

[5]Division of Pulmonary, Critical Care and Sleep Medicine, Wayne State University School of Medicine, Detroit, MI, USA

Abstract

Background: Polygenic risk scores (PRS) have become an increasingly popular approach to evaluate cancer susceptibility, but have not adequately represented Black populations in model development.

Materials and Methods: We used a previously published lung cancer PRS based on 80 SNPs associated with lung cancer risk in the OncoArray cohort and validated in UK Biobank. The PRS was evaluated for association with lung cancer risk adjusting for age, sex, total pack-years, family history of lung cancer, history of COPD, and the top five principal components for genetic ancestry.

Results: Among the 80 PRS SNPs included in the score, 14 were significantly associated with lung cancer risk ($p<0.05$) in INHALE White participants, while there were no significant SNPs among INHALE Black participants. After adjusting for covariates, the PRS was significantly associated with risk in Whites (continuous score $p=0.007$), but not in Blacks (continuous score $p=0.88$). The PRS remained a statistically significant predictor of lung cancer risk in Whites ineligible for lung cancer screening under current USPSTF guidelines (p = 0.02).

Conclusions: Using a previously validated PRS, we did find some predictive ability for lung cancer in INHALE White participants beyond traditional risk factors. However, this effect was not observed in Black participants, indicating the need to develop and validate ancestry-specific lung cancer risk models.

Corresponding Author: Matthew R. Trendowski, Ph.D., Karmanos Cancer Institute, 4100 John R, Mailcode: MM04EP, Detroit, MI 48201, hf3877@wayne.edu, Phone: 315-807-8991.

*Both authors contributed equally to this manuscript.

Conflict of Interest: The authors declare no potential conflicts of interest.

Impact: While a previously published lung cancer PRS was able to stratify White participants into different levels of risk, the model was not predictive in Blacks. Our findings highlight the need to develop and validate ancestry-specific lung cancer risk models.

Introduction

Advances in genomics have enabled a more detailed understanding of how genetic variation influences cancer predisposition. Rather than simply analyzing the association of mutations in coding regions of oncogenes and tumor suppressor genes in established signaling pathways, the analysis of single nucleotide polymorphisms (SNPs) through genome-wide association studies (GWAS) or other analytical approaches has elucidated the importance of common genetic variants in carcinogenesis (1). However, due to their typically modest effect sizes, there has been a shift towards using the cumulative effect of multiple SNPs in the creation of polygenic risk scores (PRS) that have the potential to more accurately determine overall cancer risk. The benefit of using PRS over individual SNPs resides in the fact that on a population basis, the risk conferred by multiple susceptibility SNPs is likely greater than the risk from either a single common low-penetrance SNP or a rare (population-specific) high-penetrance mutation, especially for populations without any family history of cancer. In a recent study, participants who had a high PRS for breast, prostate or colorectal cancer often did not have a family history of the disease, indicating that PRS may identify a new subset of the population at high risk for these cancers who would not have been identified as high-risk based on current risk-assessment criteria (2). In addition, a panel of 18 SNPs in *BRCA1/BRCA2* was used to stratify breast cancer risk beyond classic risk factors and mammographic density, demonstrating the additional value of incorporating PRS to evaluate risk in women participating in a nonselective national screening program (3). Similarly, an 86-SNP PRS was able to predict breast cancer status in women of European ancestry who were negative for pathogenic variants in traditional hereditary-cancer genetic testing of 28 cancer-predisposition genes (4). Therefore, genotyping multiple candidate SNPs enables an overall risk estimate for an individual to be derived that can provide predictive power for cancer susceptibility beyond family history, classic risk factors, and established oncogenes.

Lung cancer is a particularly intriguing target for PRS because the underlying genetics of susceptibility are not as well-established as for other common malignancies, such as breast or colorectal cancer, with attention mostly focused on the inherent environmental risk factors of smoking, air pollution, and occupational history (5,6). Consequently, the role of genetic factors in lung cancer development remains poorly understood, largely masked by the influence of environmental factors. Nevertheless, the risk of lung cancer in individuals with a first-degree family history is increased by approximately 50% compared to those without a family history irrespective of gender, ethnicity, histological types, and other known lung cancer risk factors (7). Further, GWAS have reported more than 45 lung cancer susceptibility loci specific to histological subtypes or ethnicity (8). These studies have enabled several PRS models to be developed for lung cancer (9–12), including a recent study that demonstrated PRS can influence differential trajectories of 5-year and cumulative absolute risk, suggesting the potential utility of using a patient's genetic background to more optimally predict the appropriate time to begin low-dose computed tomography (LDCT) screening (11).

Nevertheless, these studies have yet to establish a reliable PRS for Blacks who are uniquely susceptible to lung cancer. The age-adjusted lung cancer incidence rate is approximately 32% higher in Blacks compared to non-Hispanic Whites (13), with Blacks being more frequently diagnosed with late-stage disease and less likely to receive the recommended course of treatment based on disease stage (14). In addition, it has been demonstrated that Black men between 40 and 54 years of age are 2- to 4-times more likely to develop lung cancer compared with men of European ancestry, even after adjusting for smoking (15). Genetics may contribute to this disparity, as Coté et al. (16) demonstrated that first-degree relatives of Blacks with early-onset lung cancer have a greater risk of lung cancer compared with non-Hispanic Whites. Understanding the influence of genetics in lung cancer susceptibility in this vulnerable population may therefore enable those at greatest risk to be identified for LDCT, facilitating earlier diagnoses.

In this study, we use a previously developed PRS model (11) for lung cancer to determine its predictive power of evaluating risk in non-Hispanic Whites and Blacks from the INHALE dataset (17) regardless of family history and other risk factors.

Materials and Methods

Study Population

The INHALE study was initiated in 2012 and concluded in 2018 (refer to Schwartz et al. (17) for a detailed description). Briefly, lung cancer cases were enrolled at the Karmanos Cancer Institute, Henry Ford Health System (HFHS), or their respective network sites, within 12 months of diagnosis. Volunteer controls were enrolled from the same geographic area as the cases were drawn (Metropolitan Detroit). Cases and controls were 21–89 years of age, had never taken amiodarone or been diagnosed with bronchiectasis or cystic fibrosis. Additionally, controls carried health insurance (in the event of a finding on CT that required follow-up), never had surgical removal of any portion of either lung, and had never been diagnosed with lung cancer. Participants were asked to complete a written questionnaire, a low-dose chest CT scan, a pulmonary function test (PFT) and provide blood, saliva, and tumor tissue (for cases). The Wayne State University, HFHS and McLaren Health Care Institutional Review Boards approved the procedures used in collecting and processing participant information, and written informed consent was obtained from all subjects prior to participation.

Study Measures

Age at diagnosis, sex, family history of lung cancer, history of COPD and smoking history were collected in interviews. Race was also collected from interview data; the respondents included in this analysis identified as either ‘White or Caucasian’ or ‘Black or African American’. Pack-years were calculated by multiplying the number of years smoked by the average number of cigarettes smoked per day divided by 20.

Selection of Single Nucleotide Polymorphisms

We considered 128 SNPs used in the lung cancer polygenic risk score developed by Hung et al (11). Of the 128 SNPs, 35 SNPs were genotyped directly in INHALE samples, as they

were included in the Multi-Ethnic Genotype Array (MEGA). Imputation was also performed on INHALE samples using the Trans-Omic for Precision Medicine (TOPMed) reference panel on the Michigan Imputation Server (18,19). The imputed data captured an additional 81 PRS SNPs; seven of these SNPs failed imputation QC (r^2<0.4). There were 12 PRS SNPs that were neither genotyped directly nor imputed. We additionally excluded SNPs missing > 10% genotypes, resulting in a total of 80 SNPs for risk modeling. (Supplementary Table S1).

We evaluated associations with lung cancer risk for each PRS SNP in INHALE samples using logistic regression. We also used linear regression to test associations between PRS SNPs and smoking intensity defined by pack years in INHALE controls. SNPs were evaluated separately in non-Hispanic Whites and Blacks using self-reported race to define each population. Statistical significance was defined as $p < 0.05$. Covariates in the logistic regression model were selected *a priori* based on previous associations with lung cancer risk, and included age (20), sex (21), total pack-years (22), family history of lung cancer (5), and history of COPD (23). Population stratification was evaluated using the function eigenstrat in the R package AssocTests (24). The top five principal components (PCs) from this analysis were included as covariates to account for population substructure, and were significantly associated with lung cancer status in individual SNP tests. We also estimated the proportion of African ancestry using a panel of 122 ancestry informative markers (AIMs) and expected genotype frequencies from 1000 Genomes European and African populations. The proportion of African ancestry was estimated as $f(AFR_j)/(f(AFR_j+f(EUR_j))$, where $f(AFR_j)$ is the expected frequency for genotype j in Africans and $f(EUR_j)$ is the expected genotype frequency for genotype j in Europeans. Samples were assigned a probability of African ancestry for each AIM corresponding to their observed genotype, and these probabilities were then summed and scaled for each individual. All SNPs were included in the PRS estimation regardless of statistical significance in INHALE.

PRS Modeling

Consistent with Hung et al. (11), PRS was estimated based on additive dosages of the individual effect alleles weighted by the effect estimate. Effect estimates were derived from an analysis of more than 23,000 lung cancer cases and controls from the International Lung Cancer Consortium (described in Hung et al (11)). In the original analysis, ORs were adjusted for age, sex and the first five principal components. Although the PRS was calculated based on previously published effect estimates, the risk score was categorized by quintiles according to race-specific distributions in INHALE White and Black samples. The resulting PRS was standardized and evaluated in a continuous manner as well as by quintile, stratified by race.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Results

Cohort Characteristics

Sociodemographic characteristics for non-Hispanic Whites and Blacks included in the study are provided in Table 1. In total, there were 1,915 non-Hispanic Whites (cases: 1,103; controls: 812), and 1,123 Blacks (cases: 558; controls: 565). As expected, median age at enrollment and total pack-years were greater in lung cancer cases versus controls. The majority of both non-Hispanic White (52.8%) and Black (56.5%) participants were female. Cases were more likely to report a family history of lung cancer and history of COPD.

The median proportion of African ancestry was 0.25 (IQR=0.04) for self-reported Whites and 0.69 (IQR=0.09) for self-reported Blacks. African ancestry was nearly perfectly correlated with the first principal component (r = 0.99), modestly correlated with the second PC (r = 0.22), and very weakly correlated with the next three PCs (r = −0.01, 0.01 and −0.04, respectively). Since the top five PCs captured variability explained by African ancestry in addition to variability from other ancestral populations, we used PCs to account for ancestry in risk modeling.

Association of SNPs with Lung Cancer Status

Individual PRS SNP modeling results by race are presented in Supplementary Table S2. Of the 80 PRS SNPs tested in non-Hispanic White cases and controls, 14 SNPs were significantly associated with lung cancer risk ($p<0.05$) after adjusting for age, sex, pack years, family history of lung cancer, history of COPD and the top five PCs. The most significant SNP was a common variant (q=0.41 in White controls) located in the *TERT* gene region (chr5:1287079, rs2853677), with an OR=1.31 (95%CI: 1.14, 1.51, $p = 0.0001$). Other significant gene regions included *AK5* (chromosome 1p31.1), *CLPTM1L* (chr. 5p15.33), *REXO4* (chr. 9q34.2), *ADAMTS7* (chr. 15q25.1), *MORF4L1* (chr. 15q25.1) and *OTOP3* (chr. 17q25.1). There were also significant SNPs in several intergenic regions: rs112401627 on 5p15.33; rs116822326 on 6p21.33; rs9602270 on 13q31.1; rs77468143 on 15q21.1 and rs11855650 on 15q23. Five PRS SNPs were invariant or had a very rare minor allele relative frequency (i.e., no minor allele carriers in cases) in Black participants. Of the 75 valid SNP tests, none were significantly associated with lung cancer risk at an alpha level of 0.05 after adjusting for covariates.

Association of SNPs with Smoking Intensity

Since smoking intensity differs between cases and controls and some of the PRS SNPs have also been previously associated with nicotine addiction, we evaluated associations between each of the PRS SNPs and pack-years in control participants, stratified by race (Supplementary Table S3). Only two of the 80 SNPs were significantly associated with pack-years in White INHALE controls after adjusting for covariates. Both SNPs were located in intergenic regions: rs78334599 (chr11:116128039) was positively associated with pack-years (β=18.4, 95% CI: 4.3, 32.4, p=0.01) and rs9602270 (chr13:83706928) was negatively associated with pack-years (β=−6.9, 95% CI: −12.4, −1.4, p=0.02). Among Black INHALE controls, only rs78062588 (chr1:154593749) was significantly positively associated with pack-years (β=16.9, 95% CI: 1.7, 32.1, p=0.03).

Evaluation of Polygenic Risk Scores

SNPs were weighted by published effect sizes (11) and summed for each individual, and we evaluated whether the resulting score was validated in INHALE Whites and whether it predicted risk in INHALE Blacks. Risk score effect estimates were adjusted for age, sex, pack years, family history of lung cancer, history of COPD and the top five PCs. PRS results are presented in Table 2. When scaled and treated as a continuous score (change in odds per 1 SD increase), the PRS was significantly associated with lung cancer risk in non-Hispanic Whites (OR = 1.16, 95% CI: 1.04–1.29, p = 0.007). The PRS was not associated with lung cancer risk in Black participants after adjusting for covariates (OR = 0.99, 95% CI: 0.85–1.15, p = 0.88). The risk score was also categorized into quintiles, using race-specific distributions. In INHALE Whites, the trend was not strictly increasing, with the largest differences observed between the first (lowest) and second quintiles (OR=1.67, 95%CI=1.19, 2.35, p=0.003) and the lowest and highest (fifth) quintiles (OR=1.65, 95%CI=1.17, 2.33, p=0.004). The test for trend was also statistically significant (p=0.029). In INHALE Black participants, none of the quintile categories were significantly different from the reference group (lowest quintile), and consequently the test for trend was not significant (p=0.715).

In addition, we examined whether the PRS would predict risk specifically among those who did not satisfy the criteria for lung cancer screening under the current USPSTF guidelines: adults aged 50 to 80 years who have a 20 pack-year smoking history and currently smoke or have quit within the past 15 years (Table 1). The PRS remained significant for screening-ineligible Whites (OR = 1.21, 95% CI: 1.03–1.43, p = 0.02) after adjusting for covariates. The PRS was not significantly associated with lung cancer risk among screening-ineligible Blacks (OR = 1.06, 95% CI: 0.84–1.35, p = 0.62).

Discussion

This study makes an important contribution to the field of cancer genomics, providing additional evidence of the potential utility of using PRS to predict lung cancer susceptibility. Although the PRS generated from the methods of Hung et al. (11) was not significant in Black participants, we were able to demonstrate 14 of the selected SNPs were significantly associated with lung cancer status in non-Hispanic Whites. The PRS also was significantly associated with lung cancer status in Whites, demonstrating the potential importance of the selected SNPs in evaluating genetic susceptibility. Therefore, the current study provides additional evidence in support of using the aggregate effect of these SNPs in PRS to assess overall lung cancer risk.

The significant association of our PRS in White participants only is likely attributed to the fact that the OncoArray project data and histologically confirmed lung cancer cases and controls used to create the PRS generated by Hung et al. (11) were predominantly of European ancestry. Further, its validation using UK Biobank data in the same study again relies on a cohort who are almost exclusively of European descent. Since the PRS in our study is only significantly associated with risk in Whites, it is likely that SNPs influencing predisposition to lung cancer in Whites and Blacks have some level of variability, underscoring the importance of creating race-specific models to accurately

predict genetic risk. The observation that 5 of the 80 PRS SNPs were too low frequency to generate valid tests in Black INHALE samples supports this assertion. The importance of generating race-specific PRS for Blacks has previously been demonstrated in prostate cancer in which models derived from largely European ancestry-based GWAS were more effective for Europeans than for Africans (25). In addition, the study indicated that existing PRS were largely unable to predict whether Africans develop aggressive forms of prostate cancer, as specified by higher tumor stages or Gleason scores. Even where generalizable models have been sufficient for demonstrating a statistical association between PRS and breast cancer predisposition (26), women of Hispanic or African ancestries did not have a significant association with incidence at the extremes of the PRS distribution where the association with breast cancer risk was the strongest for women of European ancestry. Further, the effect sizes for women with African ancestry were smaller because of differences in risk allele frequencies and linkage disequilibrium patterns, leading the authors to conclude that representation of diverse populations in genomic research cohorts needs substantial improvement.

Currently, lung cancer remains the leading cause of cancer-related mortality in the United States. The 5-year relative survival rate for non-small cell lung cancer (NSCLC) is 26%, while for small cell lung cancer (SCLC) it is only 7% (Surveillance Epidemiology and End Results (RRID:SCR_006902)). However, the survival rate dramatically increases when these malignancies are found in earlier stages (NSCLC localized disease: 64%; SCLC localized disease: 29%) (Surveillance Epidemiology and End Results (RRID:SCR_006902)). This observation has provided the rationale for performing LDCT on at risk populations according to USPSTF guidelines (between the ages of 50–80 with a 20 pack-year history, and currently smoke or have quit within the last 15 years). The efficacy of this screening measure is quite profound with the initial National Lung Screening Trial (NLST) demonstrating a 20% relative reduction in lung cancer mortality, with multiple studies also confirming a survival benefit (27). However, screening has also resulted in radiation-induced cancer, false-positive results leading to unnecessary tests and invasive procedures, overdiagnosis, incidental findings, and an increase in distress (28). Notably, incidental findings (4.4%–40.7% of individuals screened) and overdiagnosis (0%–67% chance that a lung cancer was overdiagnosed) have been a common complication (29). The PRS generated in our study for non-Hispanic Whites remained significant for those who currently do not satisfy USPSTF guidelines. Consequently, the incorporation of PRS into lung cancer screening has the potential to identify individuals at risk for developing lung cancer who do not meet current screening criteria. In addition, our PRS provides further insight into the utility of incorporating genetic susceptibility into LDCT screening, as the original model from Hung et al. (11) demonstrated that ever smokers reached the LDCT screening threshold at different ages depending on their family history of lung cancer and their risk predicted by the PRS.

The potential utility of incorporating PRS into clinical care has been gaining support, with the most widely examined setting being breast cancer. Recently, Liu et al. (26) demonstrated that multiple previously developed PRS were associated with breast cancer risk with the strength of the association greatest in the extremes of the score distribution and the effect size larger for estrogen receptor (ER)-positive than for ER-negative breast

cancer. Importantly, the authors found that the European-based PRS performed well in individuals of European and Hispanic ancestry, and less well in individuals of African ancestry. As such, the development of separate PRS for populations of African ancestry are warranted, especially due to the known differences in allele frequencies between individuals of European and African ancestry (30, 31). Recently, Myriad Genetics has developed RiskScore®, a PRS-based assessment that enables women of all ancestries to receive personalized polygenic breast cancer risk assessment. It has been validated in several cohorts defined by self-reported ancestry, and was superior to a previously described 86-SNP PRS for women of European ancestry (31). These promising results indicate that PRS can be applied to common cancers, and warrants further investigation of using genetic variants to predict lung cancer susceptibility.

Our study demonstrated the potential utility of a previously validated PRS in evaluating genetic susceptibility to lung cancer. However, its efficacy may be race-specific because the PRS was only statistically significant in White participants. Development of a PRS in individuals of African ancestry is essential because risk prediction models developed from European ancestry populations may not accurately reflect allele variation that affects carcinogenesis in African ancestry populations. Black patients already experience poorer treatment outcomes for lung cancer and other malignancies. Development of a race-specific PRS for Blacks would ultimately help identify those at highest risk and subsequently decrease the impact of health disparities this population currently experiences. Nevertheless, the PRS remained significantly associated with lung cancer status in Whites who did not meet the criteria for screening under current USPSTF guidelines, indicating the model may be able to identify individuals at increased risk despite not having traditional risk factors. Further validation of the PRS model in independent cohorts would support the incorporation of a cumulative genetic risk assessment in clinical evaluation of lung cancer susceptibility and screening. Through the development of a reliable genetic risk factor prediction model, clinicians will have another method by which to evaluate lung cancer susceptibility, potentially leading to earlier diagnoses that portend more favorable treatment outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

This study was supported by the National Institutes of Health grants R01CA141769 (A.G. Schwartz, C. Neslund-Dudas, S. Gadgeel, and A.O. Soubani) and P30CA022453 (A.G. Schwartz), as well as the Herrick Foundation (A.G. Schwartz).

Abbreviations:

GWAS	genome-wide association studies
PRS	polygenic risk scores
SNP	single nucleotide polymorphism

References

1). Sud A, Kinnersley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives. Nat Rev Cancer. 2017 Nov;17(11):692–704. [PubMed: 29026206]

2). Conran CA, Shi Z, Resurreccion WK, Na R, Helfand BT, Genova E, et al. Assessing the clinical utility of genetic risk scores for targeted cancer screening. J Transl Med. 2021 Jan 22;19(1):41. [PubMed: 33482857]

3). van Veen EM, Brentnall AR, Byers H, Harkness EF, Astley SM, Sampson S,et al. Use of Single-Nucleotide Polymorphisms and Mammographic Density Plus Classic Risk Factors for Breast Cancer Risk Prediction. JAMA Oncol. 2018 Apr 1;4(4):476–482. [PubMed: 29346471]

4). Hughes E, Tshiaba P, Gallagher S, Wagner S, Judkins T, Roa B,et al. Development and Validation of a Clinical Polygenic Risk Score to Predict Breast Cancer Risk. JCO Precis Oncol. 2020 Jun 8;4:PO.19.00360. [PubMed: 32923876]

5). Kanwal M, Ding XJ, Cao Y. Familial risk for lung cancer. Oncol Lett. 2017 Feb;13(2):535–542. [PubMed: 28356926]

6). de Alencar VTL, Formiga MN, de Lima VCC. Inherited lung cancer: a review. Ecancermedicalscience. 2020 Jan 29;14:1008. [PubMed: 32104210]

7). Bailey-Wilson JE, Amos CI, Pinney SM, Petersen GM, de Andrade M, Wiest JS,et al. A major lung cancer susceptibility locus maps to chromosome 6q23–25. Am J Hum Genet. 2004 Sep;75(3):460–74. [PubMed: 15272417]

8). Bossé Y, Amos CI. A Decade of GWAS Results in Lung Cancer. Cancer Epidemiol Biomarkers Prev. 2018 Apr;27(4):363–379. [PubMed: 28615365]

9). Dai J, Lv J, Zhu M, Wang Y, Qin N, Ma H,et al. Identification of risk loci and a polygenic risk score for lung cancer: a large-scale prospective cohort study in Chinese populations. Lancet Respir Med. 2019 Oct;7(10):881–891. [PubMed: 31326317]

10). Kachuri L, Graff RE, Smith-Byrne K, Meyers TJ, Rashkin SR, Ziv E,et al. Pan-cancer analysis demonstrates that integrating polygenic risk scores with modifiable risk factors improves risk prediction. Nat Commun. 2020 Nov 27;11(1):6084. [PubMed: 33247094]

11). Hung RJ, Warkentin MT, Brhane Y, Chatterjee N, Christiani DC, Landi MT,et al. Assessing Lung Cancer Absolute Risk Trajectory Based on a Polygenic Risk Model. Cancer Res. 2021 Mar 15;81(6):1607–1615. [PubMed: 33472890]

12). Graff RE, Cavazos TB, Thai KK, Kachuri L, Rashkin SR, Hoffman JD,et al. Cross-cancer evaluation of polygenic risk scores for 16 cancer types in two large cohorts. Nat Commun. 2021 Feb 12;12(1):970. [PubMed: 33579919]

13). Ryan BM. Lung cancer health disparities. Carcinogenesis. 2018 May 28;39(6):741–751. [PubMed: 29547922]

14). Jones CC, Mercaldo SF, Blume JD, Wenzlaff AS, Schwartz AG, Chen H,et al. Racial Disparities in Lung Cancer Survival: The Contribution of Stage, Treatment, and Ancestry. J Thorac Oncol. 2018 Oct;13(10):1464–1473. [PubMed: 29885480]

15). Schwartz AG, Swanson GM. Lung carcinoma in African Americans and whites. A population-based study in metropolitan Detroit, Michigan. Cancer. 1997 Jan 1;79(1):45–52. [PubMed: 8988725]

16). Coté ML, Liu M, Bonassi S, Neri M, Schwartz AG, Christiani DC,et al. Increased risk of lung cancer in individuals with a family history of the disease: a pooled analysis from the International Lung Cancer Consortium. Eur J Cancer. 2012 Sep;48(13):1957–68. [PubMed: 22436981]

17). Schwartz AG, Lusk CM, Wenzlaff AS, Watza D, Pandolfi S, Mantha L,et al. Risk of Lung Cancer Associated with COPD Phenotype Based on Quantitative Image Analysis. Cancer Epidemiol Biomarkers Prev. 2016 Sep;25(9):1341–7. [PubMed: 27383774]

18). Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A,et al. Next-generation genotype imputation service and methods. Nat Genet. 2016 Oct;48(10):1284–1287. [PubMed: 27571263]

19). Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R,et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature. 2021 Feb;590(7845):290–299. [PubMed: 33568819]

20). de Groot PM, Wu CC, Carter BW, Munden RF. The epidemiology of lung cancer. Transl Lung Cancer Res. 2018 Jun;7(3):220–233. [PubMed: 30050761]

21). Sagerup CM, Småstuen M, Johannesen TB, Helland Å, Brustugun OT. Sex-specific trends in lung cancer incidence and survival: a population study of 40,118 cases. Thorax. 2011 Apr;66(4):301–7. [PubMed: 21199818]

22). Remen T, Pintos J, Abrahamowicz M, Siemiatycki J. Risk of lung cancer in relation to various metrics of smoking history: a case-control study in Montreal. BMC Cancer. 2018 Dec 19;18(1):1275. [PubMed: 30567516]

23). Young RP, Hopkins RJ, Christmas T, Black PN, Metcalf P, Gamble GD. COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. Eur Respir J. 2009 Aug;34(2):380–6. [PubMed: 19196816]

24). Wang L, Zhang W and Li Q AssocTests: An R package for Genetic Association Studies, J Statistical Software 2020.

25). Kim MS, Naidoo D, Hazra U, Quiver MH, Chen WC, Simonti CN, et al. . Testing the generalizability of ancestry-specific polygenic risk scores to predict prostate cancer in sub-Saharan Africa. Genome Biol. 2022 Sep 13;23(1):194. [PubMed: 36100952]

26). 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 Aug 2;4(8):e2119084. [PubMed: 34347061]

27). Copeland A, Criswell A, Ciupek A, King JC. Effectiveness of Lung Cancer Screening Implementation in the Community Setting in the United States. J Oncol Pract. 2019 Jul;15(7):e607–e615. [PubMed: 31150312]

28). Jonas DE, Reuland DS, Reddy SM, Nagle M, Clark SD, Weber RP,et al. Screening for Lung Cancer With Low-Dose Computed Tomography: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2021 Mar 9;325(10):971–987. [PubMed: 33687468]

29). Barnholtz-Sloan JS, McEvoy B, Shriver MD, Rebbeck TR. Ancestry estimation and correction for population stratification in molecular epidemiologic association studies. Cancer Epidemiol Biomarkers Prev. 2008 Mar;17(3):471–7. [PubMed: 18349264]

30). 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. [PubMed: 18593304]

31). Hughes E, Tiemeny P, Gallagher S, Meek S, Eng C, Gary M,et al. Ancestrally unbiased polygenic breast cancer (BC) risk assessment. J Clin Oncol 2021. 39:15_suppl, 10502–10502.

Table 1.

Clinical and Sociodemographic Characteristics of Cases and Controls Enrolled in the INHALE Study.

	Non-Hispanic White		Black	
Characteristic	**Controls**	**Lung Cancer Cases**	**Controls**	**Lung Cancer Cases**
Total	812	1,103	565	558
Age at Enrollment				
Median (Range)	61 (28–89)	64 (31–89)	60 (34–87)	62 (26–90)
< 50	101 (12.4%)	99 (9.0%)	69 (12.2%)	46 (8.2%)
50–59	270 (33.3%)	274 (24.8%)	212 (37.5%)	176 (31.5%)
60–69	290 (35.7%)	406 (36.8%)	202 (35.8%)	192 (34.4%)
70–79	134 (16.5%)	264 (23.9%)	70 (12.4%)	104 (18.6%)
80+	17 (2.1%)	60 (5.4%)	12 (2.1%)	40 (7.2%)
Sex				
Male	383 (47.2%)	510 (46.2%)	246 (43.5%)	243 (43.5%)
Female	429 (52.8%)	593 (53.7%)	319 (56.5%)	315 (56.5%)
Total Pack-Years	29 (0–240)	41 (0–216)	21 (0–205)	32 (0–162)
0	97 (11.9%)	118 (10.8%)	50 (8.9%)	38 (7.0%)
1–24	251 (30.9%)	176 (16.1%)	281 (49.8%)	168 (30.8%)
25–49	286 (35.2%)	376 (34.5%)	180 (31.9%)	199 (36.5%)
50–99	159 (19.6%)	346 (31.7%)	48 (8.5%)	122 (22.4%)
100+	19 (2.3%)	74 (6.8%)	5 (0.9%)	18 (3.3%)
Not Reported	0	13	1	13
First Degree Family History of Lung Cancer				
No	668 (82.3%)	829 (75.3%)	480 (85.1%)	445 (79.7%)
Yes	144 (17.7%)	273 (24.7%)	84 (14.9%)	113 (20.3%)
Not Reported	0	1	1	0
History of COPD				
No	634 (78.2%)	707 (66.8%)	459 (81.5)	351 (66.7%)
Yes	177 (21.8%)	351 (33.2%)	104 (18.5%)	175 (33.3%)

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2024 May 01.

	Non-Hispanic White		Black	
Characteristic	**Controls**	**Lung Cancer Cases**	**Controls**	**Lung Cancer Cases**
Not Reported	1	45	2	32
Qualify for Lung Cancer Screening*				
No	417 (51.4%)	461 (41.8%)	295 (52.2%)	210 (37.6%)
Yes	395 (48.6%)	642 (58.2%)	270 (47.8%)	348 (62.4%)

*Lung cancer screening criteria were based on USPSTF guidelines: adults aged 50 to 80 years who have a 20 pack-year smoking history and currently smoke or have quit within the past 15 years.

Table 2.

Association between Hung et al. (2021) Polygenic Risk Score and Lung Cancer Risk in INHALE Lung Cancer Cases and Control Stratified by Race.

	Whites		Blacks	
PRS	**OR (95% CI)**	***p*-value**	**OR (95% CI)**	***p*-value**
Continuous (per SD)	1.16 (1.04, 1.29)	0.007	0.99 (0.85, 1.15)	0.880
Quintiles				
0–20%	1.00 (Reference)		1.00 (Reference)	
20–40%	1.67 (1.19, 2.35)	0.003	0.82 (0.52, 1.31)	0.415
40–60%	1.28 (0.90, 1.80)	0.165	0.95 (0.60, 1.50)	0.823
60–80%	1.46 (1.04, 2.06)	0.029	1.07 (0.67, 1.69)	0.782
80–100%	1.65 (1.17, 2.33)	0.004	0.78 (0.49, 1.25)	0.308
P_{trend}		0.029		0.715

Effect estimates are adjusted for age, sex, pack years, history of COPD, family history of lung cancer and the top five ancestry-related PCs.