Received: 12 June 2018 | Revised: 16 May 2019 | Accepted: 31 May 2019

DOI: 10.1002/ajmg.b.32748

RESEARCH ARTICLE

American Journal of Medical Genetics Part B Neuropsychiatric Genetics WILEY

Examining sex differences in pleiotropic effects for depression and smoking using polygenic and gene-region aggregation techniques

Lauren L. Schmitz[1] | Arianna M. Gard[2] | Erin B. Ware[1]

[1]Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, Michigan

[2]Department of Psychology, University of Michigan, Ann Arbor, Michigan

Correspondence
Lauren L. Schmitz, Survey Research Center, Institute for Social Research, University of Michigan, 426 Thompson Street, Ann Arbor, MI 48109.
Email: lschmitz@umich.edu

Funding information
Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: T32 HD007109-36; National Institute on Aging, Grant/Award Numbers: K99 AG056599, P30 AG012846, R01 AG055406, R25 AG053227, T32 AG000221

Abstract

Sex differences in rates of depression are thought to contribute to sex differences in smoking initiation (SI) and number of cigarettes smoked per day (CPD). One hypothesis is that women smoke as a strategy to cope with anxiety and depression, and have difficulty quitting because of concomitant changes in hypothalamic–pituitary–adrenocortical (HPA) axis function during nicotine withdrawal states. Despite evidence of biological ties, research has not examined whether genetic factors that contribute to depression-smoking comorbidity differ by sex. We utilized two statistical aggregation techniques—polygenic scores (PGSs) and sequence kernel association testing—to assess the degree of pleiotropy between these behaviors and moderation by sex in the Health and Retirement Study (N = 8,086). At the genome-wide level, we observed associations between PGSs for depressive symptoms and SI, and measured SI and depressive symptoms (all $p < .01$). At the gene level, we found evidence of pleiotropy in *FKBP5* for SI ($p = .028$), and sex-specific pleiotropy in females in *NR3C2* ($p = .030$) and *CHRNA5* ($p = .025$) for SI and CPD, respectively. Results suggest bidirectional associations between depression and smoking may be partially accounted for by shared genetic factors, and genetic variation in genes related to HPA-axis functioning and nicotine dependence may contribute to sex differences in SI and CPD.

KEYWORDS
HPA-axis, pleiotropy, polygenic score (PGS), sequence kernel association testing (SKAT), smoking behavior

1 | INTRODUCTION

Nearly 36.5 million (15.1%) U.S. adults are current cigarette smokers (Jamal et al., 2016). Tobacco use is the leading cause of preventable morbidity and mortality; the health effects of smoking include many types of cancer, respiratory diseases, cardiovascular disease, and adverse reproductive outcomes (Centers for Disease Control and Prevention [CDC], 2014).

Though 20th century sex-related demographic trends in smoking prevalence have persisted, with men smoking more than women do, this gap is narrowing. Shrinking rates in smoking prevalence by sex have been consistent (CDC, 2002; Cheng & Kenkel, 2010; Hammond, 2009; Jamal et al., 2016; Peters, Huxley, & Woodward, 2014). Convergence in smoking rates by sex appear to be especially pronounced in teens (CDC, 2002), with some reports showing that girls are smoking more than boys (Substance Abuse Mental Health Services Administration [SAMHSA], 2007). Research from population based studies and placebo-controlled nicotine replacement trials have also found women evince more quit attempts and have higher rates of

This article was published online on 20 June 2019. An error was subsequently identified. This notice is included in the online and print versions to indicate that both have been corrected 28 June 2019.

relapse than men (Hammond, 2009; Perkins, 2001; Perkins & Scott, 2008; Pogun & Yararbas, 2009; Reynoso, Susabda, & Cepeda-Benito, 2005), suggesting that biological factors, in addition to social norms, may contribute to sex differences in smoking behavior.

1.1 | Sex differences in biopsychosocial predictors of smoking

Increasingly, sex differences in nicotine dependence have been linked to biopsychosocial predictors of smoking, including sex differences in depression (Perkins, 2001; Torres & O'Dell, 2016) and hypothalamic–pituitary–adrenocortical (HPA) axis functioning. Clinically, major depressive disorder (MDD) is more prevalent in women than in men (Hankin & Abramson, 2001; McLaughlin, Xuan, Subramanian, & Koenen, 2011), and women with a history of smoking are at higher risk of past and current MDD than men (Husky, Mazure, Paliwal, & McKee, 2008). Women also report greater perceived risks from quitting smoking, including greater negative affect—a key feature of MDD (McKee, O'Malley, Salovey, Krishnan-Sarin, & Mazure, 2005). Moreover, though concurrent depression and cigarette consumption are comorbid across sex (John, Meyer, Rumpf, & Hapke, 2004), longitudinal data suggests that only among women are there bidirectional associations between smoking behavior and depression over time (Bares, 2014).

One of the most well studied biological processes linked to depression is functioning of the HPA-axis, with documented sex differences in how this process unfolds. The HPA-axis forms one component of the physiological stress response by coordinating the release of glucocorticoids (i.e., cortisol in humans) from the adrenal gland (Gunnar & Quevedo, 2007). As the principle tobacco alkaloid (Benowitz, Hukkanen, & Jacob, 2009), nicotine stimulates greater secretion of HPA-axis hormones (e.g., cortisol), processes that are mediated by nicotine binding to nicotinic acetylcholinergic receptors expressed widely in the central nervous system (Rohleder & Kirschbaum, 2006; Tweed, Hsia, Lutfy, & Friedman, 2012). Though the molecular mechanisms of nicotine *initiation* and HPA-axis functioning may be similar across males and females, females show relatively greater circulating stress-related HPA-axis hormones during nicotine *withdrawal* (Hogle & Curtin, 2006). Moreover, negative mood states such as anxiety and depression that accompany such HPA-axis changes are also more prevalent during nicotine withdrawal for females than males (Hogle & Curtin, 2006; Soyster, Anzai, Fromont, & Prochaska, 2016). Together, this research suggests that compared to males, females are more susceptible to smoking as a strategy to cope with increased anxiety and depression, and have more difficulty quitting because of concomitant changes in HPA-axis function during withdrawal states (Torres & O'Dell, 2016), both of which may contribute to sex differences in smoking behavior and cessation.

1.2 | Do pleiotropic effects underlie sex-differences in smoking behavior?

Results from LD score regression analyses of genome-wide association study (GWAS) studies have shown evidence of genetic correlation between smoking initiation (SI), smoking intensity or cigarettes smoked per day (CPD), and depressive symptoms, or evidence that the effects of genetic variants for these traits are correlated (Bulik-Sullivan et al., 2015; Zheng et al., 2017). Table 1 reports cross-trait LD score regression estimates of genetic correlation from LD Hub that were calculated using the most recent GWAS summary statistics of depressive symptoms, SI, and CPD (Zheng et al., 2017). Depressive symptoms are positively correlated with both SI and CPD, but are statistically more significant for SI (r_g = 0.249; p = 9.96E−06) than CPD (r_g = 0.253; p = .005). In females, results from twin studies indicate a stronger genetic basis for smoking behavior (Li, Cheng, Ma, & Swan, 2003), and the presence of an underlying genetic basis for a common predisposition to smoking and depression (Dierker, Avenevoli, Stolar, & Merikangas, 2002; Kendler et al., 1993).

The genetic correlation between depression and smoking behavior observed from twin studies or LD score regression could arise through multiple mechanisms, but the most common interpretation is that they arise as a result of pleiotropy—i.e., that alleles affecting one trait on average also affect a second trait. However, despite the strong biological links in the literature between smoking, depression, and HPA-axis function in females, few studies have examined the existence of sex-specific pleiotropic effects between genetic variants for depression and smoking. Pleiotropic effects between depression and smoking could manifest as biological or mediated pleiotropy (Figure 1). Biological pleiotropy occurs when a genetic variant or gene has a direct biological influence on more than one phenotypic trait, whereas mediated pleiotropy occurs when one phenotype is itself causally related to a second phenotype so that a variant or gene associated with the first phenotype is indirectly associated with the second (Solovieff, Cotsapas, Lee, Purcell, & Smoller, 2013). Both are considered real forms of pleiotropy; however, it is important to distinguish between the two in order to accurately identify the etiological mechanisms of the two phenotypes (Solovieff et al., 2013).

Genetic variation underlying HPA-axis functioning is a plausible candidate neurobiological system in which to examine pleiotropic effects between smoking and depression (Rovaris, Mota, & Bau, 2016; Torres & O'Dell, 2016). Hyperactivation of the HPA-axis response, associated with depression and nicotine withdrawal, is thought to reflect inefficient feedback inhibition by endogenous cortisol (Pariante & Lightman, 2008). Many genes contribute to the initiation and regulation of the HPA-axis (Arnett, Muglia, Laryea, & Muglia, 2016), including *NR3C1*, *NR3C2*, *FKBP5*, and *CRHR1* (Figure 2). Though several studies have linked SNP-level variation within these genes to depression (see reviews by Arnett

TABLE 1 LD score regression estimates of genetic correlation for depressive symptoms and smoking behaviors

Phenotype 1	Phenotype 2	r_g (*SE*)	*p*-value
Depressive symptoms	CPD	0.253 (0.09)	.005
Depressive symptoms	SI	0.249 (0.056)	9.96E−06
Depressive symptoms	Former smoker	−0.158 (0.096)	.101

Note: LD score regression estimates were downloaded from LD Hub (Zheng et al., 2017), retrieved from http://ldsc.broadinstitute.org/lookup/. Abbreviations: CPD, cigarettes per day; LD, linkage disequilibrium; r_g, genetic correlation; *SE*, standard error; SI, smoking initiation.

et al., 2016; Gillespie, Phifer, Bradley, & Ressler, 2009), only a handful of studies have examined associations with smoking behavior. For example, Rogausch, Kochen, Meineke, and Hennig (2007) found that G allele carriers at rs41423247 of *NR3C1* were more likely to become smokers and had significantly higher daily cigarette consumption than C homozygotes. Other SNPs within *NR3C2*, *FKBP5*, and *CRHR1* have also been examined in relation to smoking outcomes (dos Santos et al., 2012; Jensen et al., 2015; Koopmann et al., 2016; Rovaris et al., 2013; Tang et al., 2015), with mixed results. Interestingly, a small GWAS of daily cigarette use reported associations at SNPs within *NR3C2*, providing additional evidence that genes underlying HPA-axis function may also play a role in smoking behavior. Moreover, research suggests that sex differences in corticotropin-releasing hormone signaling may underlie greater female vulnerability for stress-related psychiatric disorders (Bangasser et al., 2010). Thus, in the literature it is not clear whether genetic variants in HPA-related genes influence smoking via increased stress-related cortisol levels in individuals (particularly females) with depressive symptoms, or whether there is an additional effect of these variants on smoking independent of their effects on depression.

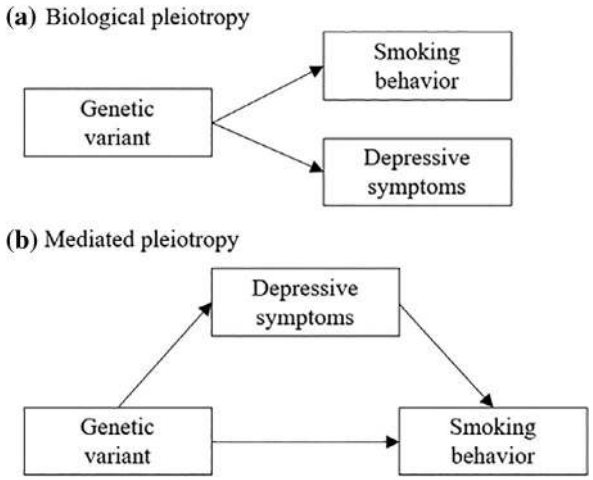

FIGURE 1 Biological and mediated pleiotropy. (a) Biological pleiotropy: A variant or gene region affects depressive symptoms and smoking behavior. (b) Mediated pleiotropy: A variant or gene region affects depressive symptoms, which in turn affect smoking behavior. As a result, an association is observed between the variant or gene region and both phenotypes

1.3 | Statistical approaches to testing pleiotropy with molecular genetic data

More recently, studies have begun using GWAS findings to identify pleiotropic effects at the genome-wide or gene-region level. At the genome-wide level, genetic overlap can be assessed by testing whether a polygenic score (PGS) for the first phenotype is significantly associated with the second phenotype. A PGS applies weights from a GWAS to genotype data to construct a weighted sum of genetic risk for a phenotype. Using a PGS increases power to detect cross-phenotype associations because it combines the cumulative effect sizes of all genetic variants across the genome for an outcome into a single scalar of genetic propensity (Dudbridge, 2013). Thus, PGSs can be easily incorporated into a multiple regression framework to simultaneously test for biological or mediated pleiotropy and/or moderation by sex. Past studies have used this approach to examine common genetic effects that underlie schizophrenia and bipolar disorder (Purcell et al., 2009), type 2 diabetes and hypertension (Lee, Yang, Goddard, Visscher, & Wray, 2012), and MDD and risk of alcohol dependence (Andersen et al., 2017). A downside to this approach is it does not

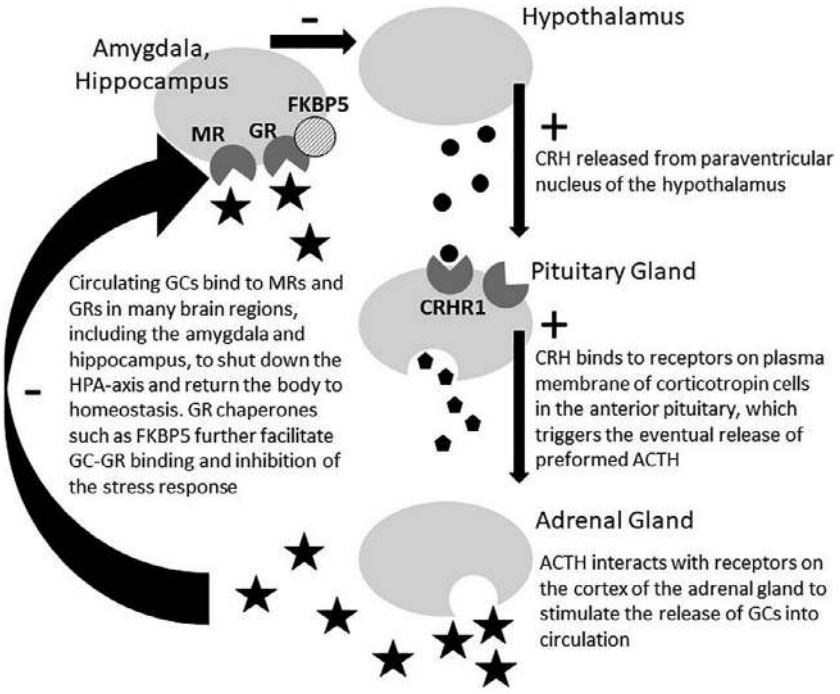

FIGURE 2 Propagation of the hypothalamic–pituitary–adrenocortical (HPA) axis. The CRH receptor type 1 is encoded by *CRHR1*. *NR3C1* and *NR3C2* encode GRs and MRs, respectively. FKBP5 is encoded by *FKBP5*. ACTH, adrenocorticotropic hormone; CRH, corticotropin-releasing-hormone; FKBP5, FK506-binding protein 51; GCs, glucorticoids (i.e., cortisol in humans); GR, glucocorticoid receptors; MRs, mineralocorticoid receptors

implicate a particular region of the genome and any related biological processes that may be driving cross-phenotype associations.

One alternative to the polygenic approach is the use of sequence kernel association testing (SKAT) or other region-based tests, which use gene aggregation techniques to test for associations between gene regions and a phenotype. SKAT improves upon the limited power of single-marker association studies by testing for associations between common and rare variants within a gene region, which increases the power to detect true effects while also maintaining biological specificity (Wu et al., 2011). This approach requires a priori knowledge of potential biologically relevant gene regions, which may be challenging to determine in the context of cross-phenotype associations of complex polygenic traits. As a result, the use of both aggregation methods may be desirable; polygenic approaches are well powered to identify the existence of pleiotropy between two phenotypes on a global level, while regional methods can provide a more detailed mapping of specific genes that may be driving these associations.

1.4 | Current study

The current study utilized both polygenic and gene region aggregation techniques to assess (a) the presence of biological or mediated pleiotropy between genetic risk factors for depression or genes implicated in HPA-axis functioning and smoking behavior, and (b) whether pleiotropic effects vary by sex. To accomplish these aims, we used data on 8,086 participants from the Health and Retirement Study (HRS)—a large, population-representative study with detailed genetic and sociodemographic data. To assess pleiotropy at the genome-wide level, we first constructed a PGS using results from a recent GWAS of depressive symptoms (Okbay et al., 2016) and tested associations between the PGS and smoking behavior as well as moderation by sex. Gene ontology analysis from the GWAS of depressive symptoms implicated SNPs in genomic regions related to enrichment of the central nervous system and the adrenal/pancreas (Okbay et al., 2016, p. 628), both of which may capture stress-related HPA-axis function in smokers. We used SKAT and interaction-SKAT (iSKAT) to examine independent and sex-specific associations between smoking behavior and gene regions that have been directly linked to HPA-axis functioning in both animal and human models, including *NR3C1*, *NR3C2*, *FKBP5*, and *CRHR1* (Figure 2).

Importantly, literature on the prospective associations between smoking, depression, and anxiety in longitudinal studies has been inconsistent in terms of the direction of association (for a review, see Fluharty, Taylor, Grabski, & Munafò, 2016). Since the HRS is a representative sample of older adults, we could not assess whether the onset of depression preceded tobacco use earlier in the life course. Therefore, we also tested for pleiotropy in the reciprocal direction, or for significant associations between genetic risk for smoking and a phenotype for depressive symptoms. For these analyses we used PGSs constructed from a GWAS of SI and CPD (Furberg et al., 2010), and SKAT and iSKAT analyses of the *BDNF* and *CHRNA5* gene regions, which have been implicated in previous studies of SI and nicotine dependence (Furberg et al., 2010; Liu et al., 2010; Thorgeirsson et al., 2010).

2 | METHODS

2.1 | Study sample

The HRS is a nationally representative, longitudinal panel study of individuals over the age of 50 and their spouses (Juster & Suzman, 1995; Sonnega et al., 2014) that is sponsored by the National Institute on Aging (NIA U01AG009740) and conducted by the University of Michigan. Launched in 1992, the HRS introduces a new cohort of participants every 6 years and interviews around 20,000 participants every 2 years.

Genotype data on ~15,000 HRS participants was collected from a random subset of the ~26,000 total participants that were selected to participate in enhanced face-to-face interviews and saliva specimen collection for DNA in 2006, 2008, and 2010. Since the HRS respondents are from various ancestral backgrounds, and we used results from GWAS of European ancestry (EA) to construct our PGS, we report results from the HRS EA sample in the main text, because the PGS will not have the same predictive power in non-European populations (Carlson et al., 2013; Martin et al., 2017; see Section 2.2 for details on population assignment). Restricting our analyses to one ancestral group is also important in that SNPs within regions of interest may tag different causal variants if the underlying linkage disequilibrium (LD) structure varies across ancestral groups (Martin et al., 2017; Rosenberg et al., 2010). However, for completeness, we report corresponding methods and results from exploratory, cross-ancestry analyses in the HRS African ancestry (AA) sample in Data S1 Supporting Information for this study (n = 1,984). We also excluded participants born before 1930 due to documented mortality selection, or increased survival among low risk smoking genotypes, in earlier birth cohorts (Domingue et al., 2017), and spouses born after 1959, since these individuals are not part of the core population-representative HRS sample. Our final EA sample includes 8,086 respondents born between 1930 and 1959.

2.2 | Genotyping and quality control

Genotyping was conducted by the Center for Inherited Disease Research (CIDR) in 2011, 2012, and 2015 (RC2 AG0336495, RC4 AG039029). Full quality control details can be found in the Quality Control Report (*Quality Control Report for Genotypic Data*, 2013). Genotype data on over 15,000 HRS participants was obtained using the Illumina HumanOmni2.5 BeadChips (HumanOmni2.5-4v1, HumanOmni2.5-8v1), which measures ~2.4 million SNPs. Genotyping quality control was performed by the Genetics Coordinating Center at the University of Washington, Seattle, WA. Individuals with missing call rates >2%, SNPs with call rates <98%, HWE p-value <.0001, chromosomal anomalies, and first-degree relatives in the HRS were removed. Imputation to 1000G Phase I v3 (released March 2012) was performed using SHAPEIT2 followed by IMPUTE2. The worldwide reference panel of all 1,092 samples from the phase I integrated variant set was used. These imputation analyses were performed and documented by the Genetics Coordinating Center at the University of Washington, Seattle, WA. All positions and names are aligned to build GRCh37/hg19.

Principal component (PC) analysis was performed on a selected set of independent SNPs to identify population group outliers and to

provide sample eigenvectors as covariates in the statistical model to adjust for possible population stratification, and were provided by the HRS. The EA sample included all respondents that had PC loadings within ±1 standard deviations (*SD*s) for eigenvectors one and two in the PC analysis of all unrelated study subjects and who self-identified as White on survey data. A second set of PCs was then generated for the analytical EA sample to further account for any population stratification within the EA sample. The EA genotype sample has been defined by the HRS and is available on dbGaP (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000428.v2.p2).

2.3 | Measures

2.3.1 | Smoking phenotypes

We examined two elements of smoking behavior: SI and smoking intensity. For SI, we used the classification of "ever smoker" from the CDC, which defines a smoker as someone who reports smoking 100 cigarettes or more in their lifetime (CDC, 2008). We assigned a value of "1" for SI if a respondent reports ever being a smoker at baseline when they entered the HRS or in subsequent waves of the HRS. For smoking intensity, we used number of CPD. If a respondent currently smokes, the HRS asks how many cigarettes they currently smoke per day on average. If the respondent no longer smokes, they are asked how many cigarettes they smoked per day when they were smoking the most. Past studies in other longitudinal cohorts have found a high overall correlation between these variables over time, supporting the idea of using either value as a general assessment of CPD (Furberg et al., 2010).

2.3.2 | Depressive symptoms phenotype

We pooled all available waves of HRS data and used the mean score respondents received on the Center for Epidemiological Studies-Depression (CES-D) scale eight item short form (Radloff, 1977; Watson, 1988; Watson, Wiese, Vaidya, & Tellegen, 1999). Participants indicated if they experienced each item much of the time during the past week using (1 = yes, 0 = no). We summed negative indicators (depressed, activities were an effort, sleep is restless, felt lonely, felt sad, and unmotivated) and reverse-coded positive indicators (felt happy and enjoyed life) to construct an overall score (range [0, 8]). Mean CES-D scores included up to 11 waves of data.

2.3.3 | Covariates

We included the first 10 PCs of the genetic data (see Section 2.2), educational attainment (1 = GED/HS degree or higher; 0 = no degree), age (respondent mean across all waves), and birth cohort (indicator variables for 6-year time spans, or 1930–1935, 1936–1941, 1942–1947, 1948–1953, and 1954–1959). In models with CPD as the outcome, we also controlled for current versus former smoker status. The association between educational attainment and smoking behavior has been well-documented (e.g., de Walque, 2010; Pampel, 2009; Pampel, Krueger, & Denney, 2010), and evidence also suggests a strong association between education and depressive symptoms (Adler et al., 1994; Lorant et al., 2003). Smoking behavior and self-reported depressive symptoms have also been shown to vary significantly by birth cohort (Jeuring et al., 2018; Weinberger et al., 2018). As depressive symptoms tend to increase with age, we controlled for the mean age of the respondent across all HRS waves to ensure any differences we observed by sex were not related to average age at reporting (Blazer, Burchett, Service, & George, 1991).

2.4 | Polygenic scores

Linear PGSs for depressive symptoms were constructed using effect sizes from the largest GWAS of depressive symptoms to date conducted by the Social Science and Genetics Association Consortium (SSGAC) on EA individuals. The SSGAC meta-analyzed publicly available results from a study performed by the Psychiatrics Genomics Consortium (PGC; Ripke et al., 2013) with GWAS results from the initial release of UK Biobank genetic data (total N = 180,666; Okbay et al., 2016). A replication analysis was performed using data from 23andMe (N = 368,890). Since the HRS was part of the original GWAS meta-analysis discovery sample, weights were re-estimated by the SSGAC to exclude the HRS. Due to privacy constraints, these weights do not contain data from 23andMe.

Linear PGSs for SI and CPD were constructed using effect sizes from a GWAS meta-analyses of EA individuals conducted by the Tobacco and Genetics Consortium (TAG; Furberg et al., 2010). For SI and CPD, a total of 74,053 participants were included in the discovery phase of the analysis; in a follow up meta-analysis of the 15 most significant regions, 143,023 participants were included for SI and 73,853 for CPD.

Genotyped SNPs in the HRS genetic database were matched to SNPs with reported results in the GWAS. In the HRS EA genetic data, 1,126,742 genotyped SNPs were available to construct the depressive symptoms PGS, 710,288 SNPs were available to construct the SI PGS, and 767,171 SNPs were available to construct the CPD PGS. To increase the power of its predictive capacity, SNPs included in the PGSs were not trimmed for LD and a p-value threshold or cut-off was not imposed (Andersen et al., 2017; Stein et al., 2017; Ware, Schmitz, Gard, & Faul, 2018). The PGSs were calculated as a weighted sum of the number of disease-associated alleles (zero, one, or two) at each SNP multiplied by the effect size for that SNP estimated from the GWAS meta-analysis. All SNPs were coded to be associated with increasing disease risk. To simplify interpretation, all PGSs were standardized to have a mean of zero and *SD* of one.

Importantly, to test if adjustments for LD affected our genome-wide tests for pleiotropy, we also estimated results with a depressive symptoms PGS constructed in the software LDpred (Vilhjálmsson et al., 2015). LDpred uses a Bayesian method to calculate PGSs that estimates posterior mean effect sizes from GWAS summary statistics by assuming a prior for the genetic architecture and LD information from a reference panel. We used the EA HRS sample as the reference panel with an LD window of 180 and the fraction of SNPs with non-zero effects assumed to be one.

2.5 | Identification of HPA-axis and smoking gene regions

We examined four genes that support the propagation of the HPA-axis: *NR3C1*, *NR3C2*, *FKBP5*, and *CRHR1* (see Data S1 Supporting Information for gene information). Figure 2 depicts how the proteins that each of these genes encode support the HPA-axis response, both in terms of initiation and regulation (Gunnar & Quevedo, 2007; Lupien, McEwen, Gunnar, & Heim, 2009). Though many genes support the HPA-axis response (Arnett et al., 2016), we chose genes that have been robustly linked to function of the HPA-axis in animal and human models and studied in humans with regards to stress-related psychiatric disorders. Using gene knockout models in mice and human studies of genetic variation, *NR3C1*, *NR3C2*, *FKBP5*, and *CRHR1* have each been linked to the production of glucorticoids (i.e., cortisol in humans, corticosterone in mice; Arnett et al., 2016; Gillespie et al., 2009; Grad & Picard, 2007; Laryea, Arnett, & Muglia, 2012; Schmidt et al., 2003). In addition, genetic variation in these genes has been repeatedly linked to individual variability in susceptibility for depression (Binder et al., 2004; Bradley et al., 2008; de Kloet et al., 2016; Liu et al., 2006; Schatzberg et al., 2014; Velders et al., 2011).

For SI, we examined the gene encoding brain-derived neurotrophic factor (*BDNF*). Identification of variants in *BDNF* have replicated in multiple GWAS of SI (Furberg et al., 2010; Liu et al, 2019). *BDNF* regulates synaptic plasticity and survival of cholinergic and dopaminergic neurons (Zhang & Poo, 2001), and is thought to play a role in modulation of dopamine reward circuits that promote continued use of nicotine after initial exposure (Furberg et al., 2010). For smoking intensity, we examined the *CHRNA5* gene, which codes for the alpha-5 subunit of the nicotinic receptors. *CHRNA5*, along with nicotinic receptor genes *CHRNA3* and *CHRNA4*, has been identified as a risk factor for heaviness of smoking (as defined by CPD), and the development of lung cancer in GWAS (Furberg et al., 2010; Liu et al., 2010; Thorgeirsson et al., 2010). We focused specifically on *CHRNA5* because it contains SNP rs16969968 (i.e., "Mr. Big"), which is widely believed to be the causal variant underlying the GWAS signal in the *CHRNA5/CHRNA3/CHRNB4* regions. In particular, it is known to cause an amino acid change in the alpha-5 subunit of the nicotinic receptors, and experiments have found this change alters the responsiveness of the nicotinic receptors to nicotine (Bierut et al., 2008).

2.6 | Statistical analyses

2.6.1 | PGS analysis

To facilitate genome-wide identification of pleiotropic effects, we estimated associations between PGSs and smoking phenotypes using a linear regression model

$$Y_i = \beta_0 + \beta_1 \mathrm{PGS}_i + \mathrm{PC}_i'\beta_2 + \epsilon_i, \quad (1)$$

where Y is the SI or CPD status of individual i, and PC_i is a vector that includes the first 10 EA genetic PCs. To test for mediated pleiotropy we ran the following additional regressions

$$\mathrm{CESD}_i = \delta_0 + \delta_1 \mathrm{PGS}_i + \mathrm{PC}_i'\delta_2 + \varepsilon_i, \quad (2)$$

$$Y_i = \theta_0 + \theta_1 PGS_i + \theta_2 \mathrm{CESD}_i + \mathrm{PC}_i'\theta_3 + X_i'\theta_4 + \mu_i, \quad (3)$$

where CESD_i is the CES-D score for individual i and X is a matrix of covariates that we include in our fully specified model (sex, educational attainment, age, and birth cohort). For mediation to hold, there must be (a) an association between the depressive symptoms PGS and smoking phenotype (SI or CPD) in Equation (1); (b) an association between the depressive symptoms PGS and the intermediate or mediator phenotype (CES-D) in Equation (2); and (c) an association between the smoking and CES-D phenotypes in Equation (3) (Baron & Kenny, 1986). Additionally, if these conditions hold, then the association between the depressive symptoms PGS and smoking phenotype must be less in Equation (3) than Equation (1).

To test if pleiotropic effects vary by gender we interacted sex (1 = female; 0 = male) with the depressive symptoms PGS in our fully specified model

$$Y_i = \gamma_0 + \gamma_1 \mathrm{PGS}_i + \gamma_2 CESD + \gamma_3 \mathrm{PGS}_i \times \mathrm{Female}_i + \mathrm{PC}_i'\gamma_4 + X_i'\gamma_5 + \sigma_i. \quad (4)$$

If γ_3 is significant even after adjusting for CES-D, this is evidence of sex-specific biological pleiotropy. In additional specifications, we also tested for sex-specific mediated pleiotropy by including a three-way interaction between the depressive symptoms PGS, sex, and CES-D. For ease of interpretation, we used a linear probability model (LPM) to estimate results for the dichotomous SI phenotype, since marginal effects or corresponding odds ratios for interaction terms in logit models are difficult to interpret (Karaca-Mandic, Norton, & Dowd, 2012). Finally, we also ran the same models in the reciprocal direction using PGSs for SI or CPD as independent variables, SI or CPD phenotypes as mediators, and the CES-D phenotype as the outcome. CPD models were analyzed in the sample of current/former smokers and included an additional control for current/former smoker status. Regression analyses were carried out using Stata 15 (StataCorp, 2017).

2.6.2 | SKAT analysis

We performed gene-region analysis (SKAT) on selected gene regions for HPA-axis function (*NR3C1*, *NR3C2*, *FKBP5*, *CRHR1*) and smoking behavior (*BDNF* and *CHRNA5*; Lee, 2013; Lee, Miropolsky, & Wu, 2013; Lee, Teslovich, Boehnke, & Lin, 2013; Wu et al., 2011). SKAT aggregates genetic information across the region using a kernel function and uses a computationally efficient variance component test to test for association. SKAT assumes the following genetic main effect model for the SI phenotype:

$$\mathrm{logit}P(y_{i=1}) = \alpha_0 + X_i'\alpha_1 + G_i'\alpha_2, \quad (5)$$

where the phenotype is dichotomous (0 = never smoker, 1 = ever smoker). Here, α_0 is an intercept term, X_i is a matrix of nongenetic covariates (first 10 EA genetic PCs, CES-D, age, educational

attainment, and birth cohort), and $G_i = (g_{i1}, \ldots, g_{ip})$ is a matrix of genotypes (0, 1, 2). The vector of regression coefficients for the covariates is represented by α_1, and α_2 is a vector of regression coefficients for the p observed genetic variants in the region. A primary assumption of SKAT is that each α_{2j}, $j = 1, \ldots, p$, follows an arbitrary distribution with mean zero and variance $w_j^2\tau$. The weights, w_j, are specified based on minor allele frequency (MAF). We weighted variants using the default $\beta(1, 25)$ weighting scheme to up-weight minor alleles. Testing $H_0: \tau = 0$ is equivalent to testing $H_0: \alpha_2 = 0$. SKAT extends to the linear model for CPD and CES-D, where the outcome is $E(y_i)$.

2.6.3 | iSKAT analysis

The iSKAT analysis is a gene or region based G × E interaction test (Lin, Lee, Christiani, & Lin, 2013). Suppose n subjects are genotyped in a region with p SNPs. For iSKAT, the interaction model for the SI phenotype is:

$$\text{logit}P(y_{i=1}) = \alpha_0 + X_i'\alpha_1 + G_i'\alpha_2 + \alpha_3\ \text{Female}_i + G \times \text{Female}_i'\alpha_4, \qquad (6)$$

where all symbols are as described above with the addition of the environmental factor (Female_i) and its effect estimate (α_3), as well as a vector of effect estimates for $G \times \text{Female}_i$ (α_4), which is an $n \times p$ matrix of $G \times E$ interactions in the region. This model assumes that each of the α_{4j}, $j = 1, \ldots, p$, independently follows an arbitrary distribution with mean zero and common variance τ^2. Testing H_0: $\tau^2 = 0$ is equivalent to testing H_0: $\alpha_4 = 0$, which tests whether at least one of the interaction terms is nonzero. iSKAT is robust to the proportion of causal variants in the region, the signs and magnitudes of the rare variants, and also controls for main effects of the rare variants (Lin et al., 2013). Both SKAT and iSKAT analyses were performed using R (Lee, Miropolsky, & Wu, 2013).

3 | RESULTS

3.1 | Descriptive statistics

Our analytic sample consisted of 4,597 females and 3,489 males. Descriptive statistics by sex are reported in Table 2. Males were 13% more likely to smoke than females ($p < .001$), and smoked 6.62 more CPD, consistent with population estimates for similar birth cohorts (CDC, 2008). Women in the HRS sample had significantly higher average CES-D scores (1.38) than men (1.04; $p < .001$), consistent with U.S population prevalence estimates for depression by gender (Kessler et al., 1994). In keeping with national trends for white men and women from these birth cohorts, 89% completed at least a GED/HS degree (Escobedo & Peddicord, 1996).

Table 3a,b presents zero order correlations between all study variables for the full sample and for males and females separately. In the full sample, the depressive symptoms, SI, and CPD PGSs were all significantly correlated with their intended phenotypes (all $p < .001$). These correlations replicated in the sex-specific samples, though some correlations (i.e., SI PGS and SI, SI PGS and current smoker, depressive symptoms PGS and SI) were somewhat stronger in females than males. In all three samples, the depressive symptoms PGS was not correlated with the CPD phenotype. Finally, completion of a GED/high school degree, age, and birth cohort (not shown) were correlated with smoking and CES-D, justifying their inclusion as potential confounders in the analysis.

3.2 | PGS results

3.2.1 | Depressive symptoms PGS to smoking phenotypes

Table 4 presents results from the multiple regression PGS analyses that test for pleiotropic effects at the genome-wide level for the depressive symptoms PGS and smoking phenotypes. Due to the null zero order correlations between the depressive symptoms PGS and the CPD phenotype reported in Table 3a,b, we focus on results for SI in the main text and report results for CPD in the Appendix (Table A1). For all PGS analyses, we used a Bonferroni adjusted alpha level of 0.006 (0.05/8 = 0.006). To see if accounting for LD in PGS construction affected our estimates, analogous results using a PGS for depressive symptoms constructed with LDpred are presented in Table A2.

We found pleiotropic effects in both sexes, as evidenced by the significant association between the depressive symptoms PGS regression coefficient (β) and the SI phenotype in Model 1, Column 2 ($p = .005$). The coefficient was positive, suggesting that greater genetic propensity for depressive symptoms was associated with an increased risk of SI. However, following adjustment for the CES-D phenotype, the depressive symptoms PGS coefficient was no longer statistically significant after multiple comparison correction (Model 2; $p = .064$). This, in combination with the highly significant association between the depressive symptoms PGS and CES-D in Model 1, Column 1 ($p = 7.90E{-}21$), is evidence of mediated pleiotropy—i.e., the depressive symptoms PGS was associated with both phenotypes when tested separately but appears to be more directly related to the CES-D phenotype (Figure 3). We did not find evidence of sex-specific biological pleiotropy from the regression coefficient on the two-way interaction term in Model 5 ("DS PGS × Female", $p = .444$), or sex-specific mediated pleiotropy from the three-way interaction term coefficient in Model 6 ("DS PGS × CES-D × Female", $p = .673$). There was no evidence of biological, mediated, or sex-specific mediated pleiotropy between the depressive symptoms PGS and the CPD phenotype (Table A1). Accounting for LD in PGS construction did not affect our results (Table A2).

3.2.2 | Smoking PGSs to depressive symptoms phenotype

In Tables 5 and 6, we present results in the reciprocal direction that tested for genome-wide pleiotropy between the SI and CPD PGSs and the CES-D phenotype. Evidence of mediated pleiotropy was found for both PGSs: the SI PGS was associated with CES-D (Table 5, Model

TABLE 2 Descriptive statistics of study variables for the full sample, males, and females, European ancestry

	All	Males	Females	Difference *p*-value
N	8,086	3,489	4,597	
Smoking initiation	4,657 (58)	2,284 (65)	2,373 (52)	4.91E−36
Cigarettes per day (CPD)[a]	24.21 (16.77)	27.58 (18.05)	20.96 (14.73)	5.81E−42
Current smoker	1,981 (24)	870 (25)	1,111 (24)	4.27E−01
Mean CES-D	1.23 (1.41)	1.04 (1.29)	1.38 (1.47)	3.93E−28
Depressive symptoms PGS	0 (1.00)	0.01 (0.99)	−0.01 (1.01)	4.03E−01
Smoking initiation PGS	0 (1.00)	−0.01 (1.00)	0.01 (1.00)	3.06E−01
CPD PGS[a]	0 (1.00)	−0.02 (0.99)	0.02 (1.01)	1.02E−01
Age[b]	61.62 (6.20)	62.15 (5.92)	61.21 (6.37)	1.31E−11
Education				
No degree	858 (11)	355 (10)	503 (11)	2.67E−01
GED or HS degree	4,547 (56)	1,818 (52)	2,729 (59)	6.87E−11
College degree	2,681 (33)	1,316 (38)	1,365 (30)	2.86E−14
Birth cohort				
1930–1935	1,772 (22)	790 (23)	982 (21)	1.68E−01
1936–1941	2,173 (27)	989 (28)	1,184 (26)	9.25E−03
1942–1947	1,529 (19)	602 (17)	927 (20)	9.28E−04
1948–1953	1,631 (20)	732 (21)	899 (20)	1.14E−01
1954–1959	981 (12)	376 (11)	605 (13)	1.14E−03

Note: Data are in *n* (%) or mean (*SD*).
[a]CPD statistics are calculated for the sample of respondents who are current or former smokers.
[b]Sample mean of respondents' mean age across their observations (all: *N* = 4,641; males: *N* = 2,275; females: *N* = 2,366).

1, Column 2; $p = .002$) and partially mediated by the SI phenotype (Table 5, Model 2; $p = .031$; Figure 3); the CPD PGS was associated with CES-D (Table 6, Model 1, Column 2; $p = .034$) and partially mediated by the CPD phenotype (Table 6, Model 2; $p = .046$), though the results for CPD were not as robust as those for SI and do not pass tests for multiple comparisons. In both models, sex-specific mediated pleiotropy was not observed.

Results from an exploratory cross-ancestry replication in the AA sample are reported in Tables S3–S6 in Data S1 Supporting Information. We caution that *these results are not directly comparable with EA results* because the PGSs were constructed using results from an EA GWAS. The depressive symptoms PGS was associated with CES-D in the AA sample (Table S3, Model 1, Column 1; $p = .005$), but apart from this, findings were null for the AA sample and the EA PGS results did not replicate.

3.3 | SKAT and iSKAT results

Pleiotropic effects at the gene-region level for the SI, CPD, and CES-D phenotypes are reported in Tables 7 and 8. *p*-Values for the joint effect of all SNPs within each HPA-axis gene region (*NR3C1*, *NR3C2*, *FKBP5*, and *CRHR1*) on the SI phenotype are reported in the top half of the table, and *p*-values for the joint effect of SNPs within gene regions for smoking on the CES-D phenotype (*BDNF* and *CHRNA5*) are reported in the bottom half of the table. A significant *p*-value indicates that the joint variance of the SNP effect sizes within a gene region is statistically different from zero (i.e., one or more SNPs within the gene region have a statistically significant association with the phenotype). Models 1–4 test the association between each gene region with and without adjustments for covariates (SKAT), and Model 5 adds a sex-specific gene region interaction to the fully adjusted model (iSKAT).

3.3.1 | HPA-axis gene regions to smoking phenotypes

SKAT and iSKAT tests for the four HPA-axis gene regions were conducted using Bonferroni-adjusted alpha levels of 0.003 (0.05/20 = 0.003). We observed some preliminary evidence of biological pleiotropy in males and females in *FKBP5*, and evidence of sex-specific biological pleiotropy in females in *NR3C2* (Table 7). Pleiotropy was not observed in *NR3C1* (Model 4, $p = .955$; Model 5, $p = .321$) or *CRHR1* (Model 4, $p = .809$; Model 5, $p = .161$). Within *FKBP5*, the joint variance of the SNP-set ($n_{SNP} = 286$) was significantly different from zero for the SI phenotype (Model 1, $p = .017$), and the *p*-value remained significant at $p < .05$ after controlling for CES-D (Model 2, $p = .019$), and after adjusting for covariates (Model 4, $p = .028$). Thus, in contrast to the PGS results, we did not find any evidence of mediated pleiotropy by the CES-D phenotype—i.e., the

TABLE 3 (a) Correlations between study variables; (b) correlations between study variables by sex

(a) European ancestry sample, *N* = 8,086

	SI	CPD[a]	Current smoker	Mean CES-D	DS PGS	SI PGS	CPD PGS	Mean age	GED/HS degree
SI	1								
CPD[a]	NA	1							
Current smoker	0.489***	−0.078***	1						
Mean CES-D	0.103***	0.034**	0.193***	1					
DS PGS	0.023**	−0.010	0.049***	0.101***	1				
SI PGS	0.122***	−0.001	0.077***	0.053***	0.067***	1			
CPD PGS	0.026**	0.057***	0.045***	0.033***	−0.013	0.031***	1		
Mean age	0.027**	0.116***	−0.081***	−0.078***	0.000	−0.008	0.023**	1	
GED/HS degree	−0.086***	−0.076***	−0.119***	−0.187***	−0.042***	−0.032***	−0.048***	−0.132***	1

(b) European ancestry males (*N* = 3,489) and females (*N* = 4,597)

		Females								
		SI	CPD[a]	Current smoker	Mean CES-D	DS PGS	SI PGS	CPD PGS	Mean age	GED/HS degree
Males	SI	1	NA	0.547***	0.141***	0.030**	0.148***	0.020	−0.016	−0.095***
	CPD[a]	NA	1	−0.015	0.078***	−0.002	0.017	0.062***	0.035*	−0.081***
	Current smoker	0.419***	−0.105***	1	0.200***	0.046***	0.108***	0.059***	−0.080***	−0.131***
	Mean CES-D	0.093***	0.057***	0.189***	1	0.104***	0.054***	0.035**	−0.056***	−0.215***
	DS PGS	0.011	−0.017	0.052***	0.101***	1	0.076***	−0.003	0.005	−0.050***
	SI PGS	0.092***	0.003	0.036**	0.049***	0.055***	1	0.035**	−0.002	−0.041***
	CPD PGS	0.044***	0.065***	0.028*	0.021	−0.026	0.025	1	0.026*	−0.071***
	Mean age	0.064***	0.158***	−0.085***	−0.092***	−0.009	−0.013	0.025	1	−0.129***
	GED/HS degree	−0.079***	−0.087***	−0.104***	−0.143***	−0.032*	−0.019	−0.016	−0.139***	1

Abbreviations: CES-D, Center for Epidemiological Studies-Depression 8 item scale; DS, depressive symptoms; GED, general education degree; HS, high school degree; PGS, polygenic score; SI, smoking initiation.

[a]CPD column is calculated for the population of respondents who report ever smoking (all: *N* = 4,641; males: *N* = 2,275; females: *N* = 2,366).

*p < .10; **p < .05; ***p < .01.

association between *FKBP5* and the SI phenotype was not significantly attenuated after controlling for CES-D. These results suggest biological pleiotropy, or one gene predicting several phenotypes. We did not observe evidence of sex-specific pleiotropy in *FKBP5* (Model 5, p = .545), a finding that is consistent with the genome-wide PGS analyses. Conversely, in *NR3C2*, the interaction term in the iSKAT model for the SNP-set (n_{SNP} = 1,138) was significantly different from zero (Model 5; p = .030), suggesting the presence of sex-specific biological pleiotropy. In keeping with the PGS results, evidence of pleiotropy between HPA-axis related genes and the CPD phenotype was not observed (Table 8).

Overall, results are suggestive as they are not significant after adjusting for multiple testing. Moreover, results from iSKAT do not allow us to test the direction of the sex-specific interaction effect or determine which variants within *NR3C2* may be driving the interaction. As a result, future studies in larger samples are needed to determine whether SNPs in this region may have protective or deleterious effects on SI in males versus females.

3.3.2 | Smoking gene regions to depressive symptoms phenotype

SKAT and iSKAT tests for the two smoking gene regions were conducted using Bonferroni-adjusted alpha levels of 0.006 (0.05/8 = 0.006). We found no evidence of biological or mediated pleiotropy between *BDNF* or *CHRNA5* and the CES-D phenotype in models that controlled for the SI phenotype (Table 7). In models that adjusted for the CPD phenotype, the iSKAT coefficient for the *CHRNA5* SNP-set (n_{SNP} = 107) was significantly different from zero (Table 8, Model 5; p = .025), suggesting sex-specific biological pleiotropy between *CHRNA5* and the CES-D

TABLE 4 Biological and mediated pleiotropy depressive symptoms polygenic score regression results for smoking initiation and CES-D, European ancestry, N = 8,086

	CES-D		SI											
	Model 1		Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Outcome	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value
DS PGS (std)	0.153*	7.90E−21	0.016*	.005	0.011	.064	0.010	.094	0.008	.164	0.003	.717	0.006	.577
	(0.016)		(0.006)		(0.006)		(0.006)		(0.006)		(0.008)		(0.011)	
CES-D					0.035*	3.15E−19	0.032*	4.21E−16	0.039*	5.00E−23	0.039*	5.21E−23	0.032*	7.13E−07
					(0.004)		(0.004)		(0.004)		(0.004)		(0.006)	
Female									−0.159*	1.28E−46	−0.159*	1.26E−46	−0.173*	1.02E−32
									(0.011)		(0.011)		(0.014)	
DS PGS × Female											0.008	.444	0.003	.849
											(0.011)		(0.014)	
DS PGS × CES-D													−0.002	.770
													(0.006)	
CES-D × Female													0.011	.158
													(0.008)	
DS PGS × CES-D × Female													0.003	.673
													(0.008)	
Age							−0.006	0.007	−0.010*	9.25E−06	−0.010*	9.28E−06	−0.010*	9.93E−06
							(0.002)		(0.002)		(0.002)		(0.002)	
GED/HS degree							−0.095*	2.19E−07	−0.093*	2.49E−07	−0.093*	2.57E−07	−0.092*	3.14E−07
							(0.018)		(0.018)		(0.018)		(0.018)	
R^2	.015		.008		.018		.026		.051		.051		.051	

Note: All models adjust for 10 European ancestry PCs and birth cohort. Bonferroni-adjusted alpha level: *p < .006.

Abbreviations: CES-D, Center for Epidemiological Studies-Depression 8 item scale; DS, depressive symptoms; GED, general education degree; HS, high school degree; PGS, polygenic score; *SE*, standard errors; SI, smoking initiation.

M1: CES-D/SI = PGS + PCS.

M2: SI = PGS + CES-D + PCS.

M3: SI = PGS + CES-D + PCS + AGE + GED/HS + COHORT.

M4: SI = PGS + CES-D + FEMALE + PCS + AGE + GED/HS + COHORT.

M5: SI = PGS + CES-D + FEMALE + FEMALE × PGS + PCS + AGE + GED/HS + COHORT.

M6: SI = PGS + CES-D + FEMALE + FEMALE × PGS + PGS × CES-D + CES-D × FEMALE + PGS × CES-D × FEMALE + PCS + AGE + GED/HS + COHORT.

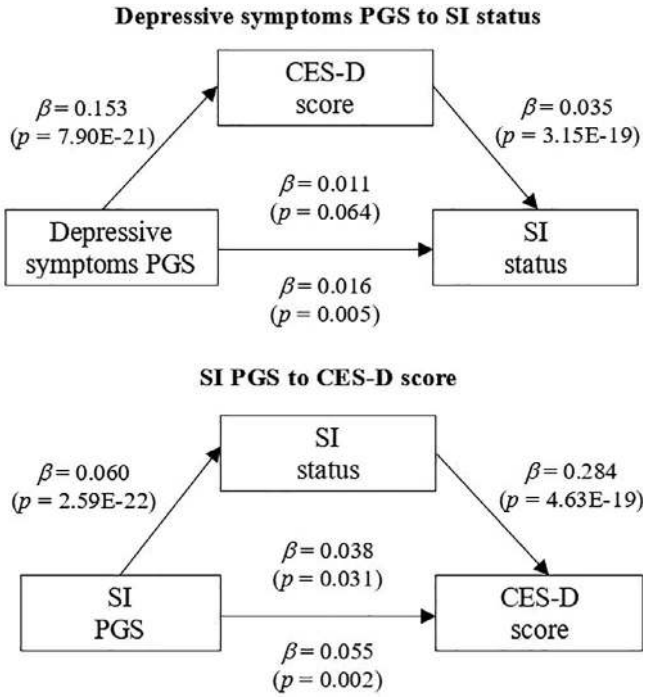

FIGURE 3 Mediated pleiotropy results from polygenic score (PGS) analyses. Results are from Tables 4 and 5, Models 1 and 2. All models adjust for 10 European ancestry genetic PCs. CES-D, Center for Epidemiological Studies-Depression 8 item scale; PGS, polygenic score; SI, smoking initiation

phenotype. However, results did not pass tests for multiple comparisons and are therefore suggestive.

SKAT and iSKAT results for individuals of AA are reported in Tables S7 and S8. Unlike the PGS analysis, results from gene region analysis are more directly comparable across ancestral populations because the level of inference is the gene region, which is defined equivalently across all ancestral populations and (presumably) has the same biological function. However, reduced sample size coupled with shorter LD blocks, greater haplotype diversity, and genotyping chips designed to tag European variants means we were likely underpowered to detect effects in the HRS AA sample. In general, SKAT and iSKAT findings were null in the AA sample, and EA findings for *FKBP5* and *NR3C2* did not replicate.

4 | DISCUSSION

We used two statistical approaches to test for biological and mediated pleiotropy between genetic risk factors for depressive symptoms and smoking behavior. Using a polygenic approach, we found evidence of mediated pleiotropy in both males and females. Results were significant in both directions—i.e., the PGS for depressive symptoms was associated with increased risk of SI, and the PGS for SI was associated with a higher CES-D score. Results from SKAT and iSKAT showed preliminary evidence of biological pleiotropy for *FKBP5* in the combined sample, and sex-specific biological pleiotropy in *NR3C2*. In contrast to the PGS results, we found no evidence of mediated pleiotropy, and no evidence of pleiotropy in the reciprocal direction, or between smoking genes (*BDNF* and *CHRNA5*) and CES-D in the full sample. Conversely, in the sample of current or former smokers, *CHRNA5* was associated with the CPD phenotype, and there was preliminary evidence for sex-specific pleiotropy between *CHRNA5* and CES-D.

Together, these results suggest that bidirectional associations between depressive symptoms and SI may be partially accounted for by shared genetic factors, and that on average these pleiotropic effects do not vary by sex on the genome-wide level. At finer levels of observation, results for females suggest that genes related specifically to HPA-axis functioning may contribute to SI, and following initiation, genetic factors related to nicotine dependence may further contribute to an increase in depressive symptoms. However, after Bonferroni correction, the SKAT and iSKAT results we report for *FKBP5* and *NR3C2* were not significant at $p < .05$. Thus, we caution these results are suggestive, and further analyses in larger samples are needed to confirm the associations we report. In addition, because the HRS is a representative sample of older adults, we cannot discern when in the life course the onset of depressive symptoms occurred and whether it proceeded tobacco use (Fluharty et al., 2016). Therefore, longitudinal analyses in younger cohorts, and/or different methodologies that can draw stronger conclusions regarding causality, are needed to confirm the direction of the association between genetic risk for depression and smoking behavior.

4.1 | Sex-specific pleiotropy in *NR3C2*

Preliminary evidence of sex-specific pleiotropic effects within *NR3C2* in the HRS is consistent with research that reports sex-differences in basal levels of HPA-axis functioning among older adults with depression. *NR3C2* encodes the mineralocorticoid receptor (MR), of which cortisol has a higher affinity for and thus, in part, determines basal HPA-axis functioning (De Kloet, Vreugdenhil, Oitzl, & Joels, 1998; Gunnar & Quevedo, 2007). In a meta-analysis of 20 studies of adults older than 60 years, significantly larger effect sizes for the association between morning basal cortisol and depression were reported in women compared to men (Murri et al., 2014). Moreover, though the current sample was composed of older postmenopausal women, a recent study in premenopausal women found that genetic variation within the MR gene moderated the impact of progesterone and estradiol on markers of negative affect (i.e., anxiety, emotion recognition; Hamstra et al., 2017). Thus, more research is needed to investigate the interplay of genetic risk for HPA-axis function dysregulation, ovarian hormones, and both depression and smoking behaviors among women across the lifespan.

In addition to biological reasons why women may be at greater risk for depression-related phenotypes and downstream compensatory behaviors (e.g., smoking), some research suggests that women are more likely to rate negative life events as stressful (Kessler & McLeod, 1984) and to report more negative affect than men (Hankin & Abramson, 2001; Watson, Clark, & Tellegen, 1988). Given that other research has challenged the assumption of reporting differences between men and

TABLE 5 Biological and mediated pleiotropy smoking initiation polygenic score regression results for smoking initiation and CES-D, European ancestry, N = 8,086

	Smoking initiation (SI)		Depressive symptoms (CES-D)											
	Model 1		Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Outcome	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value
SI PGS (std)	0.060*	2.59E−22	0.055*	.002	0.038	.031	0.033	.058	0.027	.117	0.022	.364	0.026	.508
	(0.006)		(0.018)		(0.018)		(0.017)		(0.017)		(0.024)		(0.039)	
SI					0.284*	4.63E−19	0.253*	5.96E−16	0.308*	7.49E−23	0.308*	8.48E−23	0.243*	6.41E−07
					(0.032)		(0.031)		(0.031)		(0.031)		(0.049)	
Female									0.373*	5.88E−33	0.373*	5.91E−33	0.307*	4.15E−10
									(0.031)		(0.031)		(0.049)	
SI PGS × Female											0.008	.783	0.006	.904
											(0.030)		(0.048)	
SI PGS × SI													−0.001	.987
													(0.048)	
SI × Female													0.110	.082
													(0.063)	
SI PGS × SI × Female													−0.007	.909
													(0.062)	
Age							−0.010	.107	−0.001	.921	−0.001	.925	−0.001	.925
							(0.006)		(0.006)		(0.006)		(0.006)	
GED/HS degree							−0.880*	8.42E−68	−0.863*	2.06E−66	−0.863*	2.18E−66	−0.862*	3.29E−66
							(0.050)		(0.050)		(0.050)		(0.050)	
R^2	.019		.005		.015		.059		.076		.076		.076	

Note: All models adjust for 10 European ancestry PCs and birth cohort. Bonferroni-adjusted alpha level: *p < .006.

Abbreviations: CES-D, Center for Epidemiological Studies-Depression 8 item scale; GED, general education degree; HS, high school degree; PGS, polygenic score; *SE*, standard errors; SI, smoking initiation.

M1: SI/CES-D = PGS + PCS.

M2: CES-D = PGS + SI + PCS.

M3: CES-D = PGS + SI + PCS + AGE + GED/HS + COHORT.

M4: CES-D = PGS + SI + FEMALE + PCS + AGE + GED/HS + COHORT.

M5: CES-D = PGS + SI + FEMALE + FEMALE × PGS + PCS + AGE + GED/HS + COHORT.

M6: CES-D = PGS + SI + FEMALE + FEMALE × PGS + PGS × SI + CES-D × FEMALE + PGS × SI × FEMALE + PCS + AGE + GED/HS + COHORT.

TABLE 6 Biological and mediated pleiotropy cigarettes per day (CPD) polygenic score regression results for CPD and CES-D, European ancestry, N = 4,641

	CPD		Depressive symptoms (CES-D)											
	Model 1		Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Outcome	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value
CPD PGS (std)	0.920*	4.55E−04	0.050	0.034	0.047	.046	0.028	.214	0.020	.369	0.036	.250	0.011	.839
	(0.262)		(0.023)		(0.023)		(0.022)		(0.022)		(0.031)		(0.055)	
CPD					0.003	.017	0.005*	5.58E−05	0.008*	2.43E−09	0.008*	2.60E−09	0.007*	7.34E−05
					(0.001)		(0.001)		(0.001)		(0.001)		(0.002)	
Female									0.430*	1.65E−23	0.430*	1.75E−23	0.370*	6.61E−07
									(0.043)		(0.043)		(0.074)	
CPD PGS × Female											−0.030	.468	0.031	.670
											(0.041)		(0.072)	
CPD PGS × CPD													0.001	.569
													(0.002)	
CPD × Female													0.003	.292
													(0.003)	
CPD PGS × CPD × Female													−0.003	.271
													(0.003)	
Age							0.001	.944	0.013	.154	0.013	.155	0.013	.142
							(0.009)		(0.009)		(0.009)		(0.009)	
Current smoker							0.540*	5.19E−35	0.517*	6.15E−33	0.518*	5.22E−33	0.516*	9.39E−33
							(0.043)		(0.043)		(0.043)		(0.043)	
GED/HS degree							−0.833*	1.67E−38	−0.806*	6.56E−37	−0.808*	5.19E−37	−0.808*	6.33E−37
							(0.064)		(0.063)		(0.063)		(0.063)	
R^2	.006		.007		.009		.098		.118		.118		.118	

Note. All models adjust for 10 European ancestry PCs and birth cohort. Sample includes current or former smokers only. Bonferroni-adjusted alpha level: *p < .006.

Abbreviations: CPD, cigarettes smoked per day; CES-D, Center for Epidemiological Studies-Depression 8 item scale; GED, general education degree; HS, high school degree; PGS, polygenic score; *SE*, standard errors.

M1: CPD/CES-D = PGS + PCS.

M2: CES-D = PGS + CPD + PCS.

M3: CES-D = PGS + CPD + PCS + AGE + GED/HS + COHORT.

M4: CES-D = PGS + CPD + FEMALE + PCS + AGE + GED/HS + COHORT.

M5: CES-D = PGS + CPD + FEMALE + FEMALE × PGS + PCS + AGE + GED/HS + COHORT.

M6: CES-D = PGS + CPD + FEMALE + FEMALE × PGS + PGS × SI + CES-D × FEMALE + PGS × SI × FEMALE + PCS + AGE + GED/HS + COHORT.

TABLE 7 Gene region marginal and joint effects *p*-values for smoking initiation and CES-D, sequence kernel association testing, European ancestry (*N* = 8,086)

		CES-D	Smoking initiation (SI)				
		Gene region *p*-value, SKAT	Gene region *p*-value, SKAT				$G \times E$ *p*-value, iSKAT (optimal test rho)
	Region location	Model 1	Model 1	Model 2	Model 3	Model 4	Model 5
NR3C1 (n_{SNP} = 211)	5q31.3	0.327	0.968	0.966	0.966	0.955	0.321 (1)
NR3C2 (n_{SNP} = 1,147)	4q31.23	0.383	0.258	0.180	0.163	0.103	0.030 (0)
FKBP5 (n_{SNP} = 286)	6p21.31	0.007	0.017	0.019	0.019	0.028	0.545 (0)
CRHR1 (n_{SNP} = 1,185)	17q21.31	0.336	0.927	0.894	0.892	0.809	0.161 (0.2)
		SI	Depressive symptoms (CES-D)				
		Gene region *p*-value, SKAT	Gene region *p*-value, SKAT				$G \times E$ *p*-value, iSKAT (optimal test rho)
	Region location	Model 1	Model 1	Model 2	Model 3	Model 4	Model 5
CHRNA5 (n_{SNP} = 107)	15q25.1	0.346	0.343	0.348	0.159	0.145	0.073 (0)
BDNF (n_{SNP} = 152)	11p14.1	0.598	0.281	0.244	0.180	0.248	0.760 (0)

Note: We weighted variants using the default $\beta(1, 25)$ weighting scheme to up-weight minor alleles. Optimal test rho: the value of rho to maximize statistical power resulting in the best linear combination of SKAT and burden tests. Rho = 1 is equivalent to a burden test while rho = 0 is equivalent to a SKAT test. Bonferroni-adjusted alpha level for SI results: $^*p < .003$. Bonferroni-adjusted alpha level for CES-D results: $^*p < .006$.

Abbreviations: CES-D, Center for Epidemiological Studies-Depression 8 item scale; $G \times E$, gene by environment; SKAT, sequence kernel association testing; iSKAT, interaction SKAT.

M1: CES-D/SI = GENE + PCS.

M2: SI/CES-D = GENE + CES-D/SMOKE + PCS.

M3: SI/CES-D = GENE + CES-D/SMOKE + PCS + AGE + GED/HS + COHORT.

M4: SI/CES-D = GENE + FEMALE + CES-D/SMOKE + PCS + AGE + GED/HS + COHORT.

M5: SI/CES-D = GENE + FEMALE + FEMALE × GENE + CES-D/SMOKE + PCS + AGE + GED/HS + COHORT.

women (Kendler, Thornton, & Prescott, 2001; Martin, Neighbors, & Griffith, 2013), more research is needed to understand the biological and social factors that contribute to gender-based discrepancies in psychopathology.

In contrast to previous research (e.g., Rogausch et al., 2007; Tang et al., 2015), genetic variation within the *NR3C1* and *CRHR1* was not associated with smoking behavior independent of depressive symptoms in the current sample. As neither of these studies accounted for depressive symptoms in their analysis, it may be that genetic variation within these two genes are associated with smoking behavior insofar as nicotine use is comorbid with depression (Kessler et al., 1994; Torres & O'Dell, 2016). These results highlight the relevance of considering pleiotropic effects within the psychiatric literature, as comorbidity is extremely common (Kessler et al., 1994).

Notably, we did not find evidence of pleiotropy between depression/HPA-axis related PGS/genes and CPD, suggesting genetic risk factors for depression are more related to smoking onset and persistence as opposed to smoking intensity. SI is thought to be a downstream consequence of depression and negative affect (Kassel et al., 2003; Torres & O'Dell, 2016), which may explain why genetic risk for depression and HPA-axis dysregulation was more strongly related to SI versus CPD. Though further investigation is needed, evidence from behavioral and molecular genetic studies suggest that genetic risk factors for SI and CPD may be partially independent (Heath & Martin, 1993; Wang & Li, 2010).

4.2 | Strengths and limitations

To our knowledge, this is the first study to use both PGSs and gene region aggregation methods to test for pleiotropic effects. PGSs allow us to estimate whether pleiotropy persists on a genome-wide level, increasing our power to detect effects. However, this approach does not elucidate the mechanisms of shared genetic liability for depression and smoking. In addition, GWAS weights used to construct the PGS for depressive symptoms did not condition on smoking behavior, and may therefore capture genetic risk for depression as well as genetic risk for endophenotypes like smoking that are associated with depressive symptoms. This may in part explain why on a genome-wide level we found strong evidence of mediated pleiotropy as opposed to biological pleiotropy. In addition, collapsing across all variants—some of which may have weaker pleiotropic effects—introduces substantial noise into the aggregated index, potentially attenuating evidence for association. Using a PGS may also make it difficult to observe sex differences in pleiotropy, since a genome-wide average might dilute signals at the gene level that are related to specific pathological functions. Conversely, with the SKAT and iSKAT analyses, we were able to pinpoint specific gene regions and associated biological

TABLE 8 Gene region marginal and joint effects *p*-values for cigarettes per day and CES-D, sequence kernel association testing, European ancestry (*N* = 4,641)

		CES-D	Cigarettes smoked per day (CPD)				
		Gene region *p*-value, SKAT	Gene region *p*-value, SKAT				$G \times E$ *p*-value, iSKAT (optimal test rho)
	Region location	Model 1	Model 1	Model 2	Model 3	Model 4	Model 5
NR3C1 (n_{SNP} = 211)	5q31.3	0.838	0.106	0.068	0.085	0.154	0.500 (0)
NR3C2 (n_{SNP} = 1,147)	4q31.23	0.209	0.466	0.453	0.452	0.709	0.305 (0)
FKBP5 (n_{SNP} = 286)	6p21.31	0.029	0. 784	0.783	0.824	0.822	0.673 (0)
CRHR1 (n_{SNP} = 1,185)	17q21.31	0.579	0. 607	0.725	0.689	0.689	0.765 (0)
		CPD	Depressive symptoms (CES-D)				
		Gene region *p*-value, SKAT	Gene region *p*-value, SKAT				$G \times E$ *p*-value, iSKAT (optimal test rho)
	Region location	Model 1	Model 1	Model 2	Model 3	Model 4	Model 5
CHRNA5 (n_{SNP} = 107)	15q25.1	0.011	0.663	0.641	0.593	0.580	0.025 (0)
BDNF (n_{SNP} = 152)	11p14.1	0.301	0.770	0.789	0.815	0.692	0.540 (0)

Note: We weighted variants using the default $\beta(1, 25)$ weighting scheme to up-weight minor alleles. Optimal test rho: the value of rho to maximize statistical power resulting in the best linear combination of SKAT and burden tests. Rho = 1 is equivalent to a burden test while rho = 0 is equivalent to a SKAT test. Sample includes current or former smokers only. Bonferroni-adjusted alpha level for CPD results: $^*p < .003$. Bonferroni-adjusted alpha level for CES-D results: $^*p < .006$.

Abbreviations: CES-D, Center for Epidemiological Studies-Depression 8 item scale; $G \times E$, gene-by-environment; SKAT, sequence kernel association testing; iSKAT, interaction SKAT.

M1: CES-D/CPD = GENE + PCS.

M2: CPD/CES-D = GENE + CES-D/CPD + PCS.

M3: CPD/CES-D = GENE + CES-D/CPD + PCS + AGE + CURRENT SMOKER + GED/HS + COHORT.

M4: CPD/CES-D = GENE + FEMALE + CES-D/CPD + PCS + AGE + CURRENT SMOKER + GED/HS + COHORT.

M5: CPD/CES-D = GENE + FEMALE + FEMALE × GENE + CES-D/CPD + PCS + AGE + CURRENT SMOKER + GED/HS + COHORT.

processes that may contribute to sex differences in SI and CPD. This approach requires a priori knowledge of potential candidate gene regions and decisions on whether or not to weight certain (e.g., rare) variants in the region. Both approaches were not able to detect epistatic effects across genes, which may be particularly relevant in the case of *NR3C1* and *NR3C2*, since these receptors necessarily need to act together to regulate different stages of the stress response (Rovaris et al., 2016).

The use of a large, population representative cohort of individuals from the same ancestral group is an advantage of this study in that it both increases our power to detect pleiotropic effects while also minimizing the presence of ascertainment bias. Ascertainment bias can induce spurious cross-phenotype correlations in clinical studies if, for example, patients suffering from depression and smoking (or a third related phenotype) are more likely to seek treatment than those suffering from only one condition (Smoller, Lunetta, & Robins, 2000). However, population-based studies are susceptible to biases in measurement error (Liao et al., 2014). CES-D in particular may not adequately capture more proximal biological processes involved in HPA-axis regulation, reducing our power to detect effects. Studies with in-depth clinical and multi-informant measures of psychopathology, HPA-axis functioning, and smoking are needed to further refine the associations we observed. Moreover, because the HRS is a sample of older individuals, results may be subject to mortality selection, which would bias the effects we observe downwards if individuals who survived to older ages were less likely to smoke and/or report symptoms of depression. To reduce the potential of mortality selection, we limited our analyses to individuals born after 1930.

Finally, a significant limitation of this study is that were limited to conducting analyses in individuals of European decent. Although we report findings for individuals of AA, we did not include these results in the main text because comparable GWAS in other ancestral populations are currently unavailable. Estimates from an EA GWAS are not necessarily accurate or valid in other ancestral populations, and PGSs constructed from EA GWAS summary statistics will not have the same predictive power for individuals from other ancestral backgrounds (Carlson et al., 2013; Martin et al., 2017). Thus, we caution that our EA PGS results cannot be generalized to other ancestral populations. Although SKAT and iSKAT results are more directly comparable across ancestral groups, the relatively small sample size of AA individuals in the HRS in addition to shorter LD blocks, greater haplotype diversity, and the use of genotyping chips that were designed to tag European variants means that we were likely underpowered to draw meaningful conclusions at the gene level.

4.3 | Conclusions

From a public health perspective, understanding the degree to which genetic risk for depression contributes to sex differences in SI, maintenance, and relapse, has important implications for smoking cessation

therapy. In particular, while nicotine replacement therapy alone is the most common treatment for smoking cessation (Burton, Gitchell, & Shiffman, 2000), if genetic risk for depression plays a larger role in female nicotine dependence, then tailored interventions for smoking cessation that include nonpharmacological treatments may be necessary (Reynoso et al., 2005). Our findings suggest that common genetic factors contribute to comorbidity between depressive symptoms and smoking behavior with some suggestive evidence of female-specific pleiotropy in genes that have been linked to HPA-axis function and smoking intensity. As a result, future GWAS studies of behavioral and mental health phenotypes should consider reporting summary statistics by sex, particularly if prevalence rates differ dramatically between males and females. Overall, further research is needed to assess the replicability of our findings and, more broadly, the degree to which sex-specific dysregulation of the HPA-axis and depression-related genes play a fundamental role in nicotine dependence.

ACKNOWLEDGMENTS

The research reported in this article was supported by grants from the National Institute on Aging to Schmitz (T32 AG000221; P30 AG012846; K99 AG056599), Ware (R25 AG053227; R01 AG055406), and from the National Institute of Child Health and Human Development to Gard (T32 HD007109-36). The views expressed are those of the authors and not necessarily those of the NIA or NICHD. We thank Colter Mitchell, Jessica Faul, Benjamin Domingue, and Kelly Bakulski for their valuable feedback. We are grateful to the work of the staff of the Health and Retirement Study for their many years of service, and to our study participants for sharing their lives with us and making this research possible.

CONFLICT OF INTEREST

The authors report no biomedical financial interests or potential conflicts of interest.

ORCID

Lauren L. Schmitz https://orcid.org/0000-0002-1641-289X

REFERENCES

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994). Socioeconomic status and health. The challenge of the gradient. *The American Psychologist*, *49*(1), 15–24.

Andersen, A. M., Pietrzak, R. H., Kranzler, H. R., Ma, L., Zhou, H., Liu, X., … Han, S. (2017). Polygenic scores for major depressive disorder and risk of alcohol dependence. *JAMA Psychiatry*, *74*(11), 1153–1160. https://doi.org/10.1001/jamapsychiatry.2017.2269

Arnett, M. G., Muglia, L. M., Laryea, G., & Muglia, L. J. (2016). Genetic approaches to hypothalamic–pituitary–adrenal axis regulation. *Neuropsychopharmacology*, *41*(1), 245–260. https://doi.org/10.1038/npp.2015.215

Bangasser, D. A., Curtis, A., Reyes, B. A., Bethea, T. T., Parastatidis, I., Ischiropoulos, H., … Valentino, R. J. (2010). Sex differences in corticotropin-releasing factor receptor signaling and trafficking: Potential role in female vulnerability to stress-related psychopathology. *Molecular Psychiatry*, *15*(9), 896–904. https://doi.org/10.1038/mp.2010.66

Bares, C. B. (2014). Gender, depressive symptoms, and daily cigarette use. *Journal of Dual Diagnosis*, *10*(4), 187–196. https://doi.org/10.1080/15504263.2014.961852

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Benowitz, N. S., Hukkanen, J., & Jacob, P. (2009). Nicotine chemistry, metabolism, kinetics and biomarkers. *Handb Exp Pharmacol*. *192*: 29–60. https://doi.org/10.1007/978-3-540-69248-5_2

Bierut, L. J., Stitzel, J. A., Wang, J. C., Hinrichs, A. L., Grucza, R. A., Xuei, X., … Horton, W. J. (2008). Variants in nicotinic receptors and risk for nicotine dependence. *American Journal of Psychiatry*, *165*(9), 1163–1171. https://doi.org/10.1176/appi.ajp.2008.07111711

Binder, E. B., Salyakina, D., Lichtner, P., Wochnik, G. M., Ising, M., Putz, B., … Muller-Myhsok, B. (2004). Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nature Genetics*, *36*(12), 1319–1325. https://doi.org/10.1038/ng1479

Blazer, D., Burchett, B., Service, C., & George, L. K. (1991). The association of age and depression among the elderly: An epidemiologic exploration. *Journal of Gerontology*, *46*(6), M210–M215. https://doi.org/10.1093/geronj/46.6.M210

Bradley, R. G., Binder, E. B., Epstein, M. P., Tang, Y., Nair, H. P., Liu, W., … Ressler, K. J. (2008). Influence of child abuse on adult depression: Moderation by the corticotropin-releasing hormone receptor gene. *Archives of General Psychiatry*, *65*(2), 190–200. https://doi.org/10.1001/archgenpsychiatry.2007.26

Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P. R., … Daly, M. J. (2015). An atlas of genetic correlations across human diseases and traits. *Nature Genetics*, *47*(11), 1236–1241. https://doi.org/10.1038/ng.3406

Burton, S. L., Gitchell, J. G., & Shiffman, S. (2000). Use of FDA-approved pharmacologic treatments for tobacco dependence—United States, 1984-1998. *MMWR. Morbidity and Mortality Weekly Report*, *49*(29), 665.

Carlson, C. S., Matise, T. C., North, K. E., Haiman, C. A., Fesinmeyer, M. D., Buyske, S., … Consortium, P. (2013). Generalization and dilution of association results from European GWAS in populations of non-European ancestry: The PAGE study. *PLoS Biology*, *11*(9), e1001661. https://doi.org/10.1371/journal.pbio.1001661

Centers for Disease Control and Prevention [CDC]. (2002). Trends in cigarette smoking among high school students--United States, 1991–2001. *MMWR. Morbidity and Mortality Weekly Report*, *51*(19), 409.

Centers for Disease Control and Prevention [CDC]. (2008). Cigarette smoking among adults—United States, 2007. *MMWR. Morbidity and Mortality Weekly Report*, *57*(45), 1221.

Centers for Disease Control and Prevention [CDC]. (2014). *The health consequences of smoking—50 years of progress: A report of the surgeon general*. Atlanta, GA: US Department of Health and Human Services.

Cheng, K.-W., & Kenkel, D. S. (2010). US cigarette demand: 1944–2004. *The BE Journal of Economic Analysis & Policy*, *10*(1), 1–21.

de Kloet, E. R., Otte, C., Kumsta, R., Kok, L., Hillegers, M. H., Hasselmann, H., … Joels, M. (2016). Stress and depression: A crucial role of the mineralocorticoid receptor. *Journal of Neuroendocrinology*, *28*(8), 1–12. https://doi.org/10.1111/jne.12379

De Kloet, E. R., Vreugdenhil, E., Oitzl, M. S., & Joels, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews*, *19*(3), 269–301. https://doi.org/10.1210/edrv.19.3.0331

de Walque, D. (2010). Education, information, and smoking decisions evidence from smoking histories in the United States, 1940–2000. *Journal of Human Resources*, *45*(3), 682–717.

Dierker, L. C., Avenevoli, S., Stolar, M., & Merikangas, K. R. (2002). Smoking and depression: An examination of mechanisms of

comorbidity. *American Journal of Psychiatry*, *159*(6), 947–953. https://doi.org/10.1176/appi.ajp.159.6.947

Domingue, B. W., Belsky, D. W., Harrati, A., Conley, D., Weir, D. R., & Boardman, J. D. (2017). Mortality selection in a genetic sample and implications for association studies. *International Journal of Epidemiology*, *46*(4), 1285–1294. https://doi.org/10.1093/ije/dyx041

dos Santos, V. A., Chatkin, J. M., Bau, C. H., Paixao-Cortes, V. R., Sun, Y., Zamel, N., & Siminovitch, K. (2012). Glutamate and synaptic plasticity systems and smoking behavior: Results from a genetic association study. *PLoS One*, *7*(6), e38666. https://doi.org/10.1371/journal.pone.0038666

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genetics*, *9*(3), e1003348. https://doi.org/10.1371/journal.pgen.1003348

Escobedo, L. G., & Peddicord, J. P. (1996). Smoking prevalence in US birth cohorts: The influence of gender and education. *American Journal of Public Health*, *86*(2), 231–236. https://doi.org/10.2105/AJPH.86.2.231

Fluharty, M., Taylor, A. E., Grabski, M., & Munafò, M. R. (2016). The association of cigarette smoking with depression and anxiety: A systematic review. *Nicotine & Tobacco Research*, *19*(1), 3–13. https://doi.org/10.1093/ntr/ntw140

Furberg, H., Kim, Y., Dackor, J., Boerwinkle, E., Franceschini, N., Ardissino, D., … Sullivan, P. F. (2010). Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature Genetics*, *42*(5), U134–441. https://doi.org/10.1038/ng.571

Gillespie, C. F., Phifer, J., Bradley, B., & Ressler, K. J. (2009). Risk and resilience: Genetic and environmental influences on development of the stress response. *Depression and Anxiety*, *26*(11), 984–992. https://doi.org/10.1002/da.20605

Grad, I., & Picard, D. (2007). The glucocorticoid responses are shaped by molecular chaperones. *Molecular and Cellular Endocrinology*, *275*(1–2), 2–12. https://doi.org/10.1016/j.mce.2007.05.018

Gunnar, M., & Quevedo, K. (2007). The neurobiology of stress and development. *Annual Review of Psychology*, *58*, 145–173. https://doi.org/10.1146/annurev.psych.58.110405.085605

Hammond, S. K. (2009). Global patterns of nicotine and tobacco consumption. *Handb Exp Pharmacol*, *192*, 3–28. https://doi.org/10.1007/978-3-540-69248-5_1

Hamstra, D. A., de Kloet, E. R., Quataert, I., Jansen, M., & Van der Does, W. (2017). Mineralocorticoid receptor haplotype, estradiol, progesterone and emotional information processing. *Psychoneuroendocrinology*, *76*, 162–173. https://doi.org/10.1016/j.psyneuen.2016.11.037

Hankin, B. L., & Abramson, L. Y. (2001). Development of gender differences in depression: An elaborated cognitive vulnerability–transactional stress theory. *Psychological Bulletin*, *127*(6), 773–796. https://doi.org/10.1037//O033-2909.127.6.773

Heath, A. C., & Martin, N. G. (1993). Genetic models for the natural history of smoking: Evidence for a genetic influence on smoking persistence. *Addictive Behaviors*, *18*(1), 19–34. https://doi.org/10.1016/0306-4603(93)90005-T

Hogle, J. M., & Curtin, J. J. (2006). Sex differences in negative affective response during nicotine withdrawal. *Psychophysiology*, *43*(4), 344–356. https://doi.org/10.1111/j.1469-8986.2006.00406.x

Husky, M. M., Mazure, C. M., Paliwal, P., & McKee, S. A. (2008). Gender differences in the comorbidity of smoking behavior and major depression. *Drug and Alcohol Dependence*, *93*(1–2), 176–179. https://doi.org/10.1016/j.drugalcdep.2007.07.015

Jamal, A., King, B. A., Neff, L. J., Whitmill, J., Babb, S. D., & Graffunder, C. M. (2016). Current cigarette smoking among adults—United States, 2005–2015. *MMWR. Morbidity and Mortality Weekly Report*, *65*(44), 1205–1211. https://doi.org/10.15585/mmwr.mm6544a2

Jensen, K. P., Herman, A. I., Morean, M. E., Kranzler, H. R., Gelernter, J., & Sofuoglu, M. (2015). FKBP5 variation is associated with the acute and chronic effects of nicotine. *Pharmacogenomics Journal*, *15*(4), 340–346. https://doi.org/10.1038/tpj.2014.76

Jeuring, H. W., Comijs, H. C., Deeg, D. J. H., Stek, M. L., Huisman, M., & Beekman, A. T. F. (2018). Secular trends in the prevalence of major and subthreshold depression among 55-64-year olds over 20 years. *Psychol Med*. *48*(11), 1824–1834. https://doi.org/10.1017/S0033291717003324

John, U., Meyer, C., Rumpf, H.-J., & Hapke, U. (2004). Smoking, nicotine dependence and psychiatric comorbidity—A population-based study including smoking cessation after three years. *Drug & Alcohol Dependence*, *76*(3), 287–295.

Juster, F. T., & Suzman, R. (1995). An overview of the Health and Retirement Study. *Journal of Human Resources*, *30*, S7–S56. https://doi.org/10.2307/146277

Karaca-Mandic, P., Norton, E. C., & Dowd, B. (2012). Interaction terms in nonlinear models. *Health Services Research*, *47*(1 Pt. 1), 255–274. https://doi.org/10.1111/j.1475-6773.2011.01314.x

Kassel, J. D., Stroud, L. R., & Paronis, C. A. (2003). Smoking, stress, and negative affect: Correlation, causation, and context across stages of smoking. *Psychological Bulletin*, *129*(2), 270–304. http://dx.doi.org/10.1037/0033-2909.129.2.270

Kendler, K. S., Neale, M. C., MacLean, C. J., Heath, A. C., Eaves, L. J., & Kessler, R. C. (1993). Smoking and major depression: A causal analysis. *Archives of General Psychiatry*, *50*(1), 36–43. https://doi.org/10.1001/archpsyc.1993.01820130038007

Kendler, K. S., Thornton, L. M., & Prescott, C. A. (2001). Gender differences in the rates of exposure to stressful life events and sensitivity to their depressogenic effects. *American Journal of Psychiatry*, *158*(4), 587–593. https://doi.org/10.1176/appi.ajp.158.4.587

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., … Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry*, *51*(1), 8–19. https://doi.org/10.1001/archpsyc.1994.03950010008002

Kessler, R. C., & McLeod, J. D. (1984). Sex differences in vulnerability to undesirable life events. *American Sociological Review*, *49*, 620–631. https://doi.org/10.2307/2095420

Koopmann, A., Bez, J., Lemenager, T., Hermann, D., Dinter, C., Reinhard, I., … Kiefer, F. (2016). The effect of nicotine on HPA axis activity in females is modulated by the FKBP5 genotype. *Annals of Human Genetics*, *80*(3), 154–161. https://doi.org/10.1111/ahg.12153

Laryea, G., Arnett, M. G., & Muglia, L. J. (2012). Behavioral studies and genetic alterations in corticotropin-releasing hormone (CRH) neurocircuitry: Insights into human psychiatric disorders. *Behavioral Sciences*, *2*(2), 135–171. https://doi.org/10.3390/bs2020135

Lee, S. (2013). *MetaSKAT: Meta analysis for SNP-set (sequence) kernel association test*. R package version 0.27. Retrieved from http://CRAN.R-project.org/package=MetaSKAT

Lee, S., Miropolsky, L., & Wu, M. (2013). *SKAT: SNP-set (sequence) kernel association test*. R package version 0.82. Retrieved from http://CRAN.R-project.org/package=SKAT

Lee, S., Teslovich, T. M., Boehnke, M., & Lin, X. (2013). General framework for meta-analysis of rare variants in sequencing association studies. *American Journal of Human Genetics*, *93*(1), 42–53. https://doi.org/10.1016/j.ajhg.2013.05.010

Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M., & Wray, N. R. (2012). Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics*, *28*(19), 2540–2542. https://doi.org/10.1093/bioinformatics/bts474

Li, M. D., Cheng, R., Ma, J. Z., & Swan, G. E. (2003). A meta-analysis of estimated genetic and environmental effects on smoking behavior in male and female adult twins. *Addiction*, *98*, 23–31. https://doi.org/10.1046/j.1360-0443.2003.00295.x

Liao, J., Li, X., Wong, T. Y., Wang, J. J., Khor, C. C., Tai, E. S., … Cheng, C. Y. (2014). Impact of measurement error on testing genetic association with quantitative traits. *PLoS One*, *9*(1), e87044. https://doi.org/10.1371/journal.pone.0087044

Lin, X., Lee, S., Christiani, D. C., & Lin, X. (2013). Test for interactions between a genetic marker set and environment in generalized linear models. *Biostatistics*, *14*(4), 667–681. https://doi.org/10.1093/biostatistics/kxt006

Liu, J. Z., Tozzi, F., Waterworth, D. M., Pillai, S. G., Muglia, P., Middleton, L., … Vollenweider, P. (2010). Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nature Genetics*, *42*(5), 436–440. https://doi.org/10.1038/ng.572

Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., … Zhan, X. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature Genetics*, *51*(2), 237–244. https://doi.org/10.1038/s41588-018-0307-5

Liu, Z., Zhu, F., Wang, G., Xiao, Z., Wang, H., Tang, J., … Li, W. (2006). Association of corticotropin-releasing hormone receptor1 gene SNP and haplotype with major depression. *Neuroscience Letters*, *404*(3), 358–362. https://doi.org/10.1016/j.neulet.2006.06.016

Lorant, V., Deliege, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, M. (2003). Socioeconomic inequalities in depression: A meta-analysis. *American Journal of Epidemiology*, *157*(2), 98–112.

Lupien, S. J., McEwen, B. S., Gunnar, M. R., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews. Neuroscience*, *10*(6), 434–445. https://doi.org/10.1038/nrn2639

Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., … Kenny, E. E. (2017). Human demographic history impacts genetic risk prediction across diverse populations. *American Journal of Human Genetics*, *100*(4), 635–649. https://doi.org/10.1016/j.ajhg.2017.03.004

Martin, L. A., Neighbors, H. W., & Griffith, D. M. (2013). The experience of symptoms of depression in men vs women: Analysis of the National Comorbidity Survey Replication. *JAMA Psychiatry*, *70*(10), 1100–1106. https://doi.org/10.1001/jamapsychiatry.2013.1985

McKee, S. A., O'Malley, S. S., Salovey, P., Krishnan-Sarin, S., & Mazure, C. M. (2005). Perceived risks and benefits of smoking cessation: Gender-specific predictors of motivation and treatment outcome. *Addictive Behaviors*, *30*(3), 423–435. https://doi.org/10.1016/j.addbeh.2004.05.027

McLaughlin, K. A., Xuan, Z., Subramanian, S., & Koenen, K. C. (2011). State-level women's status and psychiatric disorders among US women. *Social Psychiatry and Psychiatric Epidemiology*, *46*(11), 1161–1171.

Murri, M. B., Pariante, C., Mondelli, V., Masotti, M., Atti, A. R., Mellacqua, Z., … Amore, M. (2014). HPA axis and aging in depression: Systematic review and meta-analysis. *Psychoneuroendocrinology*, *41*, 46–62. https://doi.org/10.1016/j.psyneuen.2013.12.004

Okbay, A., Baselmans, B. M., Neve, J.-E. d., Turley, P., Nivard, M. G., Fontana, M. A., … Derringer, J. (2016). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics*, *48*(6), 624–633.

Pampel, F. C. (2009). The persistence of educational disparities in smoking. *Social Problems*, *56*(3), 526–542. https://doi.org/10.1525/sp.2009.56.3.526

Pampel, F. C., Krueger, P. M., & Denney, J. T. (2010). Socioeconomic disparities in health behaviors. *Annual Review of Sociology*, *36*, 349–370. https://doi.org/10.1146/annurev.soc.012809.102529

Pariante, C. M., & Lightman, S. L. (2008). The HPA axis in major depression: Classical theories and new developments. *Trends in Neurosciences*, *31*(9), 464–468. https://doi.org/10.1016/j.tins.2008.06.006

Perkins, K. A. (2001). Smoking cessation in women. *CNS Drugs*, *15*(5), 391–411.

Perkins, K. A., & Scott, J. (2008). Sex differences in long-term smoking cessation rates due to nicotine patch. *Nicotine & Tobacco Research*, *10*(7), 1245–1251. https://doi.org/10.1080/14622200802097506

Peters, S. A. E., Huxley, R. R., & Woodward, M. (2014). Do smoking habits differ between women and men in contemporary Western populations? Evidence from half a million people in the UK Biobank study. *BMJ Open*, *4*(12), e005663.

Pogun, S., & Yararbas, G. (2009). Sex differences in nicotine action. *Handb Exp Pharmacol*, *192*, 261–291. https://doi.org/10.1007/978-3-540-69248-5_10

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., … Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, *460*(7256), 748–752. https://doi.org/10.1038/nature08185

Quality Control Report for Genotypic Data. (2013). Retrieved from Survey Research Center, Institute for Social Research, University of Michigan.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401.

Reynoso, J., Susabda, A., & Cepeda-Benito, A. (2005). Gender differences in smoking cessation. *Journal of Psychopathology and Behavioral Assessment*, *27*(3), 227–234. https://doi.org/10.1007/s10862-005-0638-2

Ripke, S., Wray, N. R., Lewis, C. M., Hamilton, S. P., Weissman, M. M., Breen, G., … Heath, A. C. (2013). A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*, *18*(4), 497–511. https://doi.org/10.1038/mp.2012.21

Rogausch, A., Kochen, M. M., Meineke, C., & Hennig, J. (2007). Association between the BclI glucocorticoid receptor polymorphism and smoking in a sample of patients with obstructive airway disease. *Addiction Biology*, *12*(1), 93–99. https://doi.org/10.1111/j.1369-1600.2006.00045.x

Rohleder, N., & Kirschbaum, C. (2006). The hypothalamic–pituitary–adrenal (HPA) axis in habitual smokers. *International Journal of Psychophysiology*, *59*(3), 236–243.

Rosenberg, N. A., Huang, L., Jewett, E. M., Szpiech, Z. A., Jankovic, I., & Boehnke, M. (2010). Genome-wide association studies in diverse populations. *Nature Reviews. Genetics*, *11*(5), 356–366. https://doi.org/10.1038/nrg2760

Rovaris, D. L., Mota, N. R., & Bau, C. H. D. (2016). Genetic findings on the relationship between smoking and the stress system. In V.R. Preedy (Ed), Neuropathology of Drug Addictions and Substance Misuse, Volume 1: Foundations of Understanding, Tobacco, Alcohol, Cannabinoids and Opiods (pp. 209–220). London, UK: Academic Press.

Rovaris, D. L., Mota, N. R., de Azeredo, L. A., Cupertino, R. B., Bertuzzi, G. P., Polina, E. R., … Bau, C. H. D. (2013). MR and GR functional SNPs may modulate tobacco smoking susceptibility. *Journal of Neural Transmission*, *120*(10), 1499–1505. https://doi.org/10.1007/s00702-013-1012-2

Schatzberg, A. F., Keller, J., Tennakoon, L., Lembke, A., Williams, G., Kraemer, F. B., … Murphy, G. M. (2014). HPA axis genetic variation, cortisol and psychosis in major depression. *Molecular Psychiatry*, *19*(2), 220–227. https://doi.org/10.1038/mp.2013.129

Schmidt, M., Oitzl, M. S., Müller, M. B., Ohl, F., Wurst, W., Holsboer, F., … De Kloet, E. R. (2003). Regulation of the developing hypothalamic–pituitary–adrenal axis in corticotropin releasing hormone receptor 1-deficient mice. *Neuroscience*, *119*(2), 589–595. https://doi.org/10.1016/S0306-4522(03)00097-6

Smoller, J. W., Lunetta, K. L., & Robins, J. (2000). Implications of comorbidity and ascertainment bias for identifying disease genes. *American Journal of Medical Genetics*, *96*(6), 817–822.

Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M., & Smoller, J. W. (2013). Pleiotropy in complex traits: Challenges and strategies. *Nature Reviews. Genetics*, *14*(7), 483–495. https://doi.org/10.1038/nrg3461

Sonnega, A., Faul, J. D., Ofstedal, M. B., Langa, K. M., Phillips, J. W., & Weir, D. R. (2014). Cohort profile: The Health and Retirement Study (HRS). *International Journal of Epidemiology*, *43*(2), 576–585. https://doi.org/10.1093/ije/dyu067

Soyster, P., Anzai, N. E., Fromont, S. C., & Prochaska, J. J. (2016). Correlates of nicotine withdrawal severity in smokers during a smoke-free psychiatric hospitalization. *Preventive Medicine*, *92*, 176–182. https://doi.org/10.1016/j.ypmed.2016.01.026

StataCorp. (2017). *Stata statistical software: Release 15*. College Station, TX: StataCorp LLC.

Stein, M. B., Ware, E. B., Mitchell, C., Chen, C. Y., Borja, S., Cai, T., … Jain, S. (2017). Genomewide association studies of suicide attempts in US soldiers. *American Journal of Medical Genetics. Part B: Neuropsychiatric Genetics*, *174*(8), 786–797. https://doi.org/10.1002/ajmg.b.32594

Substance Abuse Mental Health Services [SAMHSA]. (2007). *2006 National Survey on Drug Use and Health*. Rockville, MD: US Department of Health and Human Services.

Tang, X., Zhan, S., Yang, L., Cui, W., Ma, J. Z., Payne, T. J., & Li, M. D. (2015). Ethnic-specific genetic association of variants in the corticotropin-releasing hormone receptor 1 gene with nicotine dependence. *BioMed Research International*, *2015*, 1–7. https://doi.org/10.1155/2015/263864

Thorgeirsson, T. E., Gudbjartsson, D. F., Surakka, I., Vink, J. M., Amin, N., Geller, F., … Gieger, C. (2010). Sequence variants at CHRNB3–CHRNA6 and CYP2A6 affect smoking behavior. *Nature Genetics*, *42*(5), 448–453. https://doi.org/10.1038/ng.573

Torres, O. V., & O'Dell, L. E. (2016). Stress is a principal factor that promotes tobacco use in females. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *65*, 260–268. https://doi.org/10.1016/j.pnpbp.2015.04.005

Tweed, J. O., Hsia, S. H., Lutfy, K., & Friedman, T. C. (2012). The endocrine effects of nicotine and cigarette smoke. *Trends in Endocrinology and Metabolism*, *23*(7), 334–342. https://doi.org/10.1016/j.tem.2012.03.006

Velders, F. P., Kuningas, M., Kumari, M., Dekker, M. J., Uitterlinden, A. G., Kirschbaum, C., … Tiemeier, H. (2011). Genetics of cortisol secretion and depressive symptoms: A candidate gene and genome wide association approach. *Psychoneuroendocrinology*, *36*(7), 1053–1061. https://doi.org/10.1016/j.psyneuen.2011.01.003

Vilhjálmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindström, S., Ripke, S., … Hayeck, T. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *The American Journal of Human Genetics*, *97*(4), 576–592. https://doi.org/10.1016/j.ajhg.2015.09.001

Wang, J., & Li, M. D. (2010). Common and unique biological pathways associated with smoking initiation/progression, nicotine dependence, and smoking cessation. *Neuropsychopharmacology*, *35*(3), 702–719. https://doi.org/10.1038/npp.2009.178

Ware, E. B., Schmitz, L. S., Gard, A., & Faul, J. (2018). *HRS Polygenic Scores, 2006–2012 Genetic Data*.

Watson, D. (1988). Intraindividual and interindividual analyses of positive and negative affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality and Social Psychology*, *54*(6), 1020–1030.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*(5), 820–838. https://doi.org/10.1037/0022-3514.76.5.820

Weinberger, A. H., Gbedemah, M., Martinez, A. M., Nash, D., Galea, S., & Goodwin, R. D. (2018). Trends in depression prevalence in the USA from 2005 to 2015: Widening disparities in vulnerable groups. *Psychological Medicine*, *48*(8), 1308–1315. https://doi.org/10.1017/S0033291717002781

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., & Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *American Journal of Human Genetics*, *89*(1), 82–93. https://doi.org/10.1016/j.ajhg.2011.05.029

Zhang, L. I., & Poo, M. M. (2001). Electrical activity and development of neural circuits. *Nature Neuroscience*, *4*(11s), 1207–1214. https://doi.org/10.1038/nn753

Zheng, J., Erzurumluoglu, A. M., Elsworth, B. L., Kemp, J. P., Howe, L., Haycock, P. C., … Warrington, N. M. (2017). LD hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*, *33*(2), 272–279. https://doi.org/10.1093/bioinformatics/btw613

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Schmitz LL, Gard AM, Ware EB. Examining sex differences in pleiotropic effects for depression and smoking using polygenic and gene-region aggregation techniques. *Am J Med Genet Part B*. 2019;180B:448–468. https://doi.org/10.1002/ajmg.b.32748

APPENDIX

TABLE A1 Biological and mediated pleiotropy depressive symptoms polygenic score regression results for mean CES-D and cigarettes per day, European ancestry, N = 4,641

	CES-D		CPD											
	Model 1		Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Outcome	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value
DS PGS (std)	0.174*	5.93E−14	−0.103	.694	−0.176	.502	−0.115	.656	−0.198	.431	−0.299	.393	−0.247	.587
	(0.023)		(0.261)		(0.262)		(0.257)		(0.252)		(0.350)		(0.455)	
Mean CES-D					0.423	.011	0.704*	3.65E−05	1.024*	1.21E−09	1.023*	1.25E−09	1.166*	7.31E−06
					(0.166)		(0.170)		(0.168)		(0.168)		(0.260)	
Female									−6.918*	2.25E−45	−6.922*	2.15E−45	−6.636*	2.39E−24
									(0.484)		(0.484)		(0.648)	
DS PGS × Female											0.198	.678	0.376	.558
											(0.477)		(0.642)	
DS PGS × CES-D													−0.062	.809
													(0.257)	
CES-D × Female													−0.211	.525
													(0.332)	
DS PGS × CES-D × Female													−0.075	.819
													(0.327)	
Age							−0.399*	0.0001	−0.576*	1.32E−08	−0.575*	1.36E−08	−0.575*	1.39E−08
							(0.103)		(0.101)		(0.101)		(0.101)	
Current smoker							−3.134*	6.86E−10	−2.815*	1.53E−08	−2.814*	1.56E−08	−2.823*	1.43E−08
							(0.507)		(0.497)		(0.497)		(0.497)	
GED/HS degree							−2.859*	0.0001	−2.889*	7.92E−05	−2.883*	8.20E−05	−2.894*	7.84E−05
							(0.747)		(0.731)		(0.731)		(0.732)	
R^2	.019		.004		.005		.045		.085		.085		.086	

Note. All models adjust for 10 European ancestry PCs and birth cohort. Sample includes current or former smokers only. Bonferroni-adjusted alpha level: *p < .006.

Abbreviations: CES-D, Center for Epidemiological Studies-Depression 8 item scale; CPD, cigarettes per day; DS, depressive symptoms; GED, general education degree; HS, high school degree; PGS, polygenic score; *SE*, standard errors.

M1: CES-D/CPD = PGS + PCS.

M2: CPD = PGS + CES-D + PCS.

M3: CPD = PGS + CES-D + PCS + CURRENT SMOKER + AGE + GED/HS + COHORT.

M4: CPD = PGS + CES-D + FEMALE + PCS + CURRENT SMOKER + AGE + GED/HS + COHORT.

M5: CPD = PGS + CES-D + FEMALE + FEMALE × PGS + PCS + CURRENT SMOKER + AGE + GED/HS + COHORT.

M6: CPD = PGS + CES-D + FEMALE + FEMALE × PGS + PGS × CES-D + CES-D × FEMALE + PGS × CES-D × FEMALE + PCS + CURRENT SMOKER +AGE + GED/HS + COHORT.

TABLE A2 Biological and mediated pleiotropy depressive symptoms LDpred polygenic score regression results for CES-D and smoking initiation, European ancestry, N = 8,086

	CES-D		SI											
	Model 1		Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Outcome	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value	Beta (*SE*)	*p*-value
DS PGS (std)	0.149*	6.60E−21	0.016*	.005	0.011	.060	0.010	.088	0.008	.154	0.005	.580	0.007	.541
	(0.016)		(0.006)		(0.006)		(0.006)		(0.006)		(0.008)		(0.011)	
CES-D					0.035*	3.26E−19	0.032*	4.36E−16	0.039*	5.21E−23	0.039*	5.35E−23	0.032*	9.01E−07
					(0.004)		(0.004)		(0.004)		(0.004)		(0.006)	
Female									−0.159*	1.29E−46	−0.159*	1.29E−46	−0.173*	9.20E−33
									(0.011)		(0.011)		(0.014)	
DS PGS × Female											0.006	.604	0.0005	.974
											(0.011)		(0.014)	
DS PGS × CES-D													−0.001	.895
													(0.006)	
CES-D × Female													0.012	.150
													(0.008)	
DS PGS × CES-D × Female													0.003	.730
													(0.008)	
Age							−0.006	.007	−0.010*	9.42E−06	−0.010*	9.46E−06	−0.010*	1.04E−05
													(0.002)	
GED/HS degree							−0.095*	2.18E−07	−0.093*	2.48E−07	−0.093*	2.54E−07	−0.092*	3.08E−07
													(0.018)	
R^2	.015		.008		.018		.026		.051		.051		.051	

Note: All models adjust for 10 European ancestry PCs and birth cohort. Bonferroni-adjusted alpha level: *p < .006.

Abbreviations: CES-D, Center for Epidemiological Studies-Depression 8 item scale; DS, depressive symptoms; GED, general education degree; HS, high school degree; PGS, polygenic score; *SE*, standard errors; SI, smoking initiation.

M1: CES-D/SI = PGS + PCS.

M2: SI = PGS + CES-D + PCS.

M3: SI = PGS + CES-D + PCS + AGE + GED/HS + COHORT.

M4: SI = PGS + CES-D + FEMALE + PCS + AGE + GED/HS + COHORT.

M5: SI = PGS + CES-D + FEMALE + FEMALE × PGS + PCS + AGE + GED/HS + COHORT.

M6: SI = PGS + CES-D + FEMALE + FEMALE × PGS + PGS × CES-D + CES-D × FEMALE + PGS × CES-D × FEMALE + PCS + AGE + GED/HS + COHORT.