Polygenic Risk and the Developmental Progression to Heavy, Persistent Smoking and Nicotine Dependence

Evidence From a 4-Decade Longitudinal Study

Daniel W. Belsky, PhD; Terrie E. Moffitt, PhD; Timothy B. Baker, PhD; Andrea K. Biddle, PhD; James P. Evans, MD, PhD; HonaLee Harrington, BA; Renate Houts, PhD; Madeline Meier, PhD; Karen Sugden, PhD; Benjamin Williams, BS; Richie Poulton, PhD; Avshalom Caspi, PhD

**Importance:** Genome-wide hypothesis-free discovery methods have identified loci that are associated with heavy smoking in adulthood. Research is needed to understand developmental processes that link newly discovered genetic risks with adult heavy smoking.

**Objective:** To test how genetic risks discovered in genome-wide association studies of adult smoking influence the developmental progression of smoking behavior from initiation through conversion to daily smoking, progression to heavy smoking, nicotine dependence, and struggles with cessation.

**Design:** A 38-year, prospective, longitudinal study of a representative birth cohort.

**Setting:** The Dunedin Multidisciplinary Health and Development Study of New Zealand.

**Participants:** The study included 1037 male and female participants.

**Exposure:** We assessed genetic risk with a multilocus genetic risk score. The genetic risk score was composed of single-nucleotide polymorphisms identified in 3 meta-analyses of genome-wide association studies of smoking quantity phenotypes.

**Main Outcomes and Measures:** Smoking initiation, conversion to daily smoking, progression to heavy smoking, nicotine dependence (Fagerström Test of Nicotine Dependence), and cessation difficulties were evaluated at 8 assessments spanning the ages of 11 to 38 years.

**Results:** Genetic risk score was unrelated to smoking initiation. However, individuals at higher genetic risk were more likely to convert to daily smoking as teenagers, progressed more rapidly from smoking initiation to heavy smoking, persisted longer in smoking heavily, developed nicotine dependence more frequently, were more reliant on smoking to cope with stress, and were more likely to fail in their cessation attempts. Further analysis revealed that 2 adolescent developmental phenotypes—early conversion to daily smoking and rapid progression to heavy smoking—mediated associations between the genetic risk score and mature phenotypes of persistent heavy smoking, nicotine dependence, and cessation failure. The genetic risk score predicted smoking risk over and above family history.

**Conclusions and Relevance:** Initiatives that disrupt the developmental progression of smoking behavior among adolescents may mitigate genetic risks for developing adult smoking problems. Future genetic research may maximize discovery potential by focusing on smoking behavior soon after smoking initiation and by studying young smokers.

JAMA Psychiatry. Published online March 27, 2013. doi:10.1001/jamapsychiatry.2013.736

IGARETTE SMOKING IS A costly, prevalent public health problem. The US Centers for Disease Control and Prevention attribute more than 400,000 deaths and $95 million in lost productivity to smoking during 2000-2004. Approximately 20% of adults still smoke daily despite widespread knowledge of smoking’s health effects and increasing economic costs to smokers due to increasing taxes. Thus, more effective interventions to prevent smoking, motivate smoking cessation, and prevent relapse are needed.

Studies of twins suggest that genetic differences among individuals have an important role in smoking behavior, cessa-
analyses). An important additional step in the trans-
smoking cessation likelihood and in pharmacogenetic
readiness being used in clinical applications (eg, to predict
smokers and former smokers revealed genes that relate
with cessation in their 20s and 30s. We tested whether
the cohort as members initiated smoking during adoles-
cence.95% retention. We collected smoking behavior data at 8
assessments spanning the ages of 11 to 38 years. This ap-
plication of these GWAS findings is to test whether genetic
risk accelerates the developmental progression from smoking initiation to heavy smoking, and this, in
turn, increases the severity of adult smoking problems,
such as heavy, intractable smoking and nicotine depen-
dence. This model has relevance to public health inter-
ventions that might delay the developmental progress-
ion to heavy smoking. To put the magnitudes of genetic
effects in context and to determine whether molecu-
lar genetic measurements provided novel information
about risk, we conducted an additional analysis compar-
ing molecular genetic information to family history in-
formation. These analyses asked how large molecular
genetic effects were relative to family history effects and
whether molecular genetic effects were independent of
family history effects in predicting risk.

**METHODS**

**SAMPLE**

Participants are members of the Dunedin Multidisciplinary Health and Development Study, a longitudinal investigation of health and behavior in a complete birth cohort. Study members (N=1037, 91% of eligible births, 52% male) were all individu-
als born between April 1972 and March 1973 in Dunedin, New Zealand, who were eligible for the longitudinal study based on residence in the province at age 3 years and who participated in the first follow-up assessment at age 3 years. The cohort repre-
sents the full range of socioeconomic status in the general popu-
lation of New Zealand’s South Island and is primarily white.22
Assessments were performed at birth and at the ages of 3, 5, 7, 9, 11, 13, 15, 18, 21, 26, 32, and, most recently, 38 years, when
1007 study members were still alive, with 95% retention. At each
assessment wave, study members are brought to the Dunedin
research unit for a full day of interviews and examinations. The
Otago Ethics Committee approved each phase of the study, and
informed consent was obtained from all study members.

**MEASURES**

**Genetic Risk Score**

A challenge for developmental research following up GWAS
discoveries is that effect sizes for individual single-nucleotide polymorphisms (SNPs) are small; the largest effects for smok-
ing quantity approach a change of 1 cigarette per day per risk allele. Moreover, many of the longitudinal studies24 with data
necessary to investigate developmental phenotypes are under-
powered to test individual SNP effects. However, evidence shows
that smoking-associated loci make additive contributions to risk,
recommending aggregating risk alleles.24-27 Summing risk alleles
across GWAS-identified SNPs to compute a genetic risk score (GRS) yields a quantitative index of genetic risk with a normal
distribution28 and a potentially larger effect size.

We derived the GRS from 3 recent meta-analyses of GWAS
discoveries that used as their phenotype cigarettes smoked per day.7-9 To
construct the GRS, we considered SNPs from regions with ge-
nome-wide significant associations in at least 2 meta-analyses.
All 3 meta-analyses identified SNPs in the q23.1 region of chro-
mosome 15 containing the CHRNA5-CHRNA3-CHRNB4 gene
cluster. Two meta-analyses identified SNPs in the q13.2 re-
gion of chromosome 19 containing the gene CYP2A6. These

![Image](https://via.placeholder.com/150)

**Figure 1.** Genetic risk and the developmental progression of smoking behavior. In the hypothesized model, genetic risk influences the mature phenotypes of heavy smoking persistence, nicotine dependence, and cessation failure through a pathway mediated by 3 developmental phenotypes: smoking initiation, conversion to daily smoking, and progression to heavy smoking.

smoked more heavily as adults, if they were more nico-
tine dependent, and if they were more likely to fail in their
cessation attempts. Finally, we tested the hypothesis that
genetic risk accelerates the developmental progression
from smoking initiation to heavy smoking, and this, in
turn, increases the severity of adult smoking problems,
such as heavy, intractable smoking and nicotine depen-
dence. This model has relevance to public health inter-
ventions that might delay the developmental progress-
ion to heavy smoking. To put the magnitudes of genetic
effects in context and to determine whether molecu-
lar genetic measurements provided novel information
about risk, we conducted an additional analysis compar-
ing molecular genetic information to family history in-
formation. These analyses asked how large molecular
genetic effects were relative to family history effects and
whether molecular genetic effects were independent of
family history effects in predicting risk.
genes influence nicotine response and nicotine metabolism, have been linked with nicotine dependence, and are candidate genes in research into the development of smoking behavior. Therefore, we focused our inquiry on the top GWAS SNPs in these 2 regions (eMethods; http://www.jamapsych.com). In 15q25.1, we selected the SNPs rs16969698, rs6495308, rs8032771, and rs12595358. The SNPs rs16969698 and rs6495308, which fall within the CHRNA5-CHRNA3-CHRNB4 gene cluster, were reported previously to have independent associations with smoking quantity. The SNPs rs8032771 and rs12595358, which are located downstream of the CHRNA5-CHRNA3-CHRNB4 gene cluster, were in weak linkage disequilibrium with rs16969698 and rs6495308 ($R^2 \leq 0.10$) and were genome-wide significant in the largest meta-analysis7 ($P<1 \times 10^{-16}$ for both; $P$ values for these SNPs were not published in the other 2 meta-analyses). In 19q13.2, we selected the SNPs rs7937 and rs4105144. Following 2 previous studies25,27 using the other 2 meta-analyses). In 19q13.2, we selected the SNPs rs7937 and rs4105144. Following 2 previous studies25,27 using the other 2 meta-analyses. In 19q13.2, we selected the SNPs rs7937 and rs4105144. Following 2 previous studies25,27 using the other 2 meta-analyses. In 19q13.2, we selected the SNPs rs7937 and rs4105144. Following 2 previous studies25,27 using the other 2 meta-analyses.

To validate this GRS, we used independent data from the Atherosclerosis Risk in the Communities database and the Study of Addiction: Genetics and Environment database, accessed through the National Institutes of Health Database of Genotypes and Phenotypes. When a GRS SNP was unavailable in one of these databases, we selected the closest linkage disequilibrium proxy for that SNP to include in the GRS. Among European-descent Atherosclerosis Risk in the Communities participants (n=8293), each SD increase in the GRS predicted a 1.45-pack-year increase in lifetime cigarette consumption among individuals who had ever smoked ($P<.001$) and a 1.02-cigarette increase in daily consumption among these ever smokers ($P<.001$). Replication of the GRS–smoking quantity association in the Study of Addiction: Genetics and Environment database and additional validation analyses testing versions of the GRS that exclude the SNPs rs16969698 and rs6495308 are presented in eTable 1.

Dunedin cohort genotyping was conducted with a commercially available array (Illumina, Inc) using DNA extracted from whole blood (93% of the sample) or buccal swabs (7% of the sample). The GRS SNPs or proxies (linkage $R^2 \geq 0.85$) were called successfully in 95% of European-descent study members (eTable 2). These 880 individuals formed the analysis sample. Cohort members carried a mean (SD) of 7.06 (2.27) of 12 possible risk alleles. Cohort members’ sex and socioeconomic status were unrelated to their genetic risk (Pearson $r \leq 0.01$). The GRS was standardized to have a mean (SD) of 0 (1) for analyses (GRS).

### Family History of Smoking

Family histories of smoking were available for 99% of the cohort. The family history consisted of reports of smoking history provided by study members and both parents for study members’ siblings, parents, and grandparents. The family history was summarized as the proportion of family members in the pedigree who were ever regular smokers, adjusted to account for differences in genetic relatedness to the proband of first- and second-degree relatives.

### Smoking Behavior

The developmental progression of smoking behavior in the Dunedin cohort is shown in Figure 2A. Measurement of adolescent developmental phenotypes and mature phenotypes of smoking behavior is shown in Figure 2B.

### RESULTS

Data analysis was divided into 3 parts. First, we analyzed associations between the GRS and developmental phenotypes of smoking behavior. Second, we analyzed associations between the GRS and mature phenotypes. Third, we tested whether developmental phenotypes mediated associations between the GRS and mature phenotypes. We used different statistical models to analyze outcome data as required by the outcome’s distribution. We analyzed continuously distributed outcome data (eg, lifetime cigarette consumption in pack-years) using ordinary least squares. We analyzed dichotomous outcome data (eg, daily smoker by age 15 years) using Poisson regression models because this is a standard method to derive relative risks. We analyzed count outcome data (eg, the number of assessments at which the study member met criteria for nicotine dependence) using negative binomial regression models to account for the overdispersion of many of the count measures. We analyzed hazards of smoking initiation, progression to heavy smoking, becoming nicotine dependent, and relapsing from a quit attempt using Cox proportional hazards regression models. To account for differences in the frequency with which study members attempted cessation, we constructed panel data sets that included one observation per study member per assessment for the data for ages 18–32 years and one observation per study member per quit attempt for the data for life-history calendars. We analyzed these panel data sets to analyze the genetic effect on smokers’ risks of cessation failure during ages 18 to 32 years and on their hazards of relapse during ages 32 to 38 years. We accounted for nonindependence of repeated observations of individuals using generalized estimating equation models of risks and conditional risk-set models of hazards.

We tested whether genetic effects on the mature phenotypes of persistent heavy smoking, nicotine dependence, and relapse were mediated by adolescent developmental phenotypes using the structural equation described by McKinnon and Dwyer and the methods described by Preacher et al. To allow for a single test of mediation, we conducted a principal components analysis of the mature phenotypes of persistent heavy smoking (pack-years smoked at age 38 years), nicotine dependence (total number of symptoms across all assessments), and cessation failure (number of assessments with relapse). This analysis indicated that the mature phenotypes were positively and significantly correlated (eTable 3) and could be summarized in a single component that explained 78% of the variance in the 3 measures (factor loading = 0.61 for persistent heavy smoking, 0.60 for nicotine dependence, and 0.52 for cessation failure). We used this component as the dependent variable in our mediation analysis. Analyses were adjusted for sex and conducted using STATA statistical software, version 11.0 (StataCorp LP). Panel-data models were fitted to longitudinal repeated-measures data using the XT and ST commands in STATA statistical software, version 11.0. Unless otherwise noted, effect sizes are presented for 1-SD increase in genetic risk.

### GENETIC RISK AND SMOKING INITIATION

The GRS was not associated with whether individuals initiated smoking or with the timing of initiation (relative risk [RR] for smoking initiation = 0.98; 95% CI, 0.95–1.02; cumulative hazard ratio [HR] for initiation = 1.01;
Figure 2. Smoking behavior in the Dunedin cohort. A, Developmental progression of smoking behavior in the Dunedin cohort. Study members reported their smoking status during in-person assessments at the ages of 11 (percentage of ever-smokers = 7%), 13 (13%), 15 (62%), 18 (66%), 21 (70%), 26 (70%), 32 (71%), and 38 (71%) years and their daily cigarette consumption at the ages of 13 (percentage of daily smokers = 1%), 15 (14%), 18 (31%), 21 (34%), 26 (35%), 32 (30%), and 38 (20%) years. We assessed cessation failure using study members’ reports of quit attempts and outcomes at the ages of 18, 21, 26, 32, and 38 years. B, Measurements of developmental and mature smoking phenotypes. Data are number (percentage) of study members unless otherwise indicated.

95% CI, 0.94-1.09; based on a 1-SD increase in genetic risk; Table). Subsequent analyses focused on the 627 Dunedin cohort members who initiated smoking at some point during follow-up (Figure 2).

GENETIC RISK AND THE PROGRESSION OF SMOKING BEHAVIOR

Individuals at higher genetic risk were more likely to progress to smoking 20 or more cigarettes per day and did so more rapidly (HR = 1.35; 95% CI, 1.14-1.58). Figure 3A shows the cumulative hazards for smoking 20 cigarettes or more per day for individuals at low, average, and high genetic risk. An unexpected finding was that individuals who initiated smoking but who did not progress to daily smoking or to heavy smoking, so-called chippers, were at the lowest genetic risk of any group in the cohort (Figure 3B).

Among ever-smokers, 19% converted to daily smoking by age 15 years (early conversion) and 10% progressed to smoking 20 cigarettes or more per day by age 18 years (rapid progression to heavy smoking). Adolescents at higher genetic risk were more likely to convert to daily smoking early (RR = 1.24; 95% CI, 1.06-1.45) and to progress rapidly from smoking initiation to heavy smoking (RR = 1.43; 95% CI, 1.10-1.86).
To cope with stress. Among the 277 study members at higher genetic risk were more reliant on smoking as a coping strategy (B = 0.22; 95% CI, 0.11-0.32).

GENETIC RISK AND SMOKING CESSATION

Assessment of cessation failure is challenging. Therefore, we looked for convergent evidence across 2 approaches to testing genetic associations with cessation failure. We first analyzed study members’ reports of cessation failure between the ages of 18 and 32 years. Across 14 years of follow-up, 405 cohort members smoked daily. A total of 90% of this group made at least one quit attempt, and 51% reported a cessation failure at 1 or more assessments. Cohort members at higher genetic risk were more likely to experience cessation failure in their quit attempts (RR = 1.11; 95% CI, 1.01-1.22).

We next used the month-to-month life-history calendars to look closely at cohort members’ smoking behavior during their 30s, when cessation was most common. Across 72 months of follow-up, 277 cohort members smoked daily, and 53% of these smokers made a quit attempt lasting 1 month or more. Relapse was common (occurring in 62% of quitters). Quitters at higher genetic risk were more likely to relapse and did so sooner after quitting (HR = 1.22; 95% CI, 1.02-1.45). Only 20% of daily smokers achieved successful cessation (abstinent for ≥1 month).

GENETIC RISK AND NICOTINE DEPENDENCE

Through age 38 years, 27% of ever-smokers developed nicotine dependence. Individuals at higher genetic risk were more likely to become nicotine dependent compared with individuals at lower genetic risk and were nicotine dependent at more assessments (HR for nicotine dependence = 1.27; 95% CI, 1.09-1.47; IRR for assessments with nicotine dependence = 1.22; 95% CI, 1.06-1.41) (Figure 4B).

In addition to testing genetic associations with nicotine dependence, we also asked whether cohort members at higher genetic risk were more reliant on smoking to cope with stress. Among the 277 study members who smoked daily during ages 32 to 38 years, those at higher genetic risk relied more heavily on smoking as a coping strategy (B = 0.22; 95% CI, 0.11-0.32).

### Table. Effect Sizes for Genetic and Family History Associations With Developmental and Clinical Phenotypes of Smoking Behavior

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect Size Measure</th>
<th>Genetic Risk Score</th>
<th>Family History Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoker status</td>
<td>RR</td>
<td>0.98 (0.95 to 1.02)</td>
<td>1.12 (1.07 to 1.17)</td>
</tr>
<tr>
<td>Ever-smoker status</td>
<td>HR^b</td>
<td>1.51 (0.94 to 1.09)</td>
<td>1.06 (0.98 to 1.15)</td>
</tr>
<tr>
<td>Lifetime hazard for smoking initiation</td>
<td>RR</td>
<td>1.24 (1.06 to 1.45)</td>
<td>1.52 (1.27 to 1.83)</td>
</tr>
<tr>
<td>Smoking to cope with stress (ages 32-38 years, among 277 daily smokers)</td>
<td>RR</td>
<td>1.43 (1.10 to 1.86)</td>
<td>1.68 (1.26 to 2.24)</td>
</tr>
<tr>
<td>Smoking to cope for ≥20 cigarettes per day</td>
<td>HR^b</td>
<td>1.35 (1.14 to 1.58)</td>
<td>1.47 (1.23 to 1.76)</td>
</tr>
<tr>
<td>Heavy smoking persistence (among 627 ever-smokers)</td>
<td>B</td>
<td>1.05 (0.36 to 1.73)</td>
<td>2.49 (1.80 to 3.19)</td>
</tr>
<tr>
<td>Count of assessments smoking ≥20 cigarettes per day (among 627 ever-smokers)</td>
<td>IRR</td>
<td>1.26 (1.07 to 1.49)</td>
<td>1.49 (1.24 to 1.80)</td>
</tr>
<tr>
<td>Nicotine dependence (among 627 ever-smokers)</td>
<td>RR</td>
<td>1.27 (1.09 to 1.47)</td>
<td>1.53 (1.29 to 1.80)</td>
</tr>
<tr>
<td>Count of assessments with nicotine dependence</td>
<td>IRR</td>
<td>1.22 (1.06 to 1.41)</td>
<td>1.50 (1.28 to 1.75)</td>
</tr>
<tr>
<td>Smoking to cope score</td>
<td>B</td>
<td>0.22 (0.11 to 0.32)</td>
<td>0.09 (−0.06 to 0.24)</td>
</tr>
<tr>
<td>Risk of cessation failure</td>
<td>RR^b</td>
<td>1.11 (1.01 to 1.22)</td>
<td>1.11 (1.00 to 1.23)</td>
</tr>
<tr>
<td>Ages of 18-32 years (405 daily smokers, 364 who attempted cessation)</td>
<td>RR</td>
<td>1.11 (1.01 to 1.22)</td>
<td>1.11 (1.00 to 1.23)</td>
</tr>
<tr>
<td>Hazard of relapse after quit attempts lasting ≥1 mo (among 146 who quit for ≥1 mo)</td>
<td>HR^b</td>
<td>1.22 (1.02 to 1.45)</td>
<td>0.96 (0.79 to 1.17)</td>
</tr>
<tr>
<td>Likelihood of successful cessation (among daily smokers)</td>
<td>RR</td>
<td>0.73 (0.57 to 0.93)</td>
<td>0.94 (0.73 to 1.20)</td>
</tr>
</tbody>
</table>

Abbreviations: HR, hazard ratio; IRR, incident rate ratios; RR, relative risk.

^a The correlation between the genetic risk score and the family history score was r = 0.011 (P = .76).

^b Effect sizes were estimated from longitudinal data sets that included repeated observation of individuals over time.
study members reported cessation failure. The adult smoking problems factor explained 78% of the variance in the 3 indicators. Individuals at higher genetic risk developed more smoking problems in adulthood ($r = 0.10$, $P = .01$). We next tested whether this association was accounted for by the more rapid developmental progression of smoking behavior among individuals at higher genetic risk. A total of 81% of this association was accounted for by the 2 adolescent developmental phenotypes of early conversion to daily smoking and rapid progression to smoking 20 or more cigarettes per day (eTable 4). As a further attempt to address the question of whether preventing rapid progression from smoking initiation to heavy smoking could mitigate genetic risks, we conducted a utopian control analysis.\(^7\) We asked whether genetic risks continued to predict adult smoking problems in the subset of individuals who initiated smoking but who did not exhibit either of the rapid progression phenotypes ($n = 454$). In this subgroup, genetic risk was uncoupled from the development of smoking problems in adulthood ($r = 0.05$, $P = .18$).

**OVERLAP OF MOLECULAR GENETIC RISK AND FAMILY HISTORY OF SMOKING BEHAVIOR**

The family history score and the GRS were uncorrelated ($r = 0.011$). Both family history and the GRS predicted study members’ smoking phenotypes (Table). When family history and the GRS were standardized and included in regression models simultaneously, the GRS and family history coefficients were unchanged and remained statistically significant (ie, genetic risk and family history were independent and additive predictors of smoking phenotypes). In the mediation analyses, adjustment for family history did not change results. Thus, the GRS contained different information about risk for developmental and mature phenotypes of smoking behavior compared with family history.

**COMMENT**

Etiologic research on substance abuse highlights the importance of progression from initiation to heavy use during adolescence in the development of dependence in adulthood.\(^58,59\) In this study, we linked the developmental progression of smoking behavior to genetic risk. We derived a GRS from GWASs of smoking quantity. This GRS was not related to smoking initiation. In fact, daily smokers who did not progress to heavy use were at lower genetic risk than individuals who never smoked. Among individuals who initiated smoking, those at higher genetic risk progressed more rapidly to heavy smoking and nicotine dependence, were more likely to become persistent heavy smokers and persistently nicotine dependent, and had more difficulty quitting. Critically, high genetic risk led individuals to become persistent heavy smokers and persistently nicotine dependent, and unable to quit only to the extent that they progressed rapidly from smoking initiation to heavy smoking during adolescence.

The GWASs from which we derived our measure of genetic risk were designed to discover genetic correlates of smoking quantity. Therefore, the fact that genetic risk was uncoupled from the development of smoking problems in adulthood suggests a more refined analysis of genetic risk related smoking quantity and adult smoking problems.
Figure 4. Genetic risk predicts mature phenotypes of smoking behavior. A, Among individuals who initiated smoking, those at higher genetic risk smoked more cigarettes by 38 years of age. Ever-smokers were all individuals who initiated smoking by 38 years of age (n = 627). The bars of the histogram graph the percentages of the sample carrying 1 to 12 risk alleles. The dots and SE bars reflect mean lifetime cigarette consumption (in pack-years) for ever-smokers carrying 1 to 3, 4, 5, 6, 7, 8, 9, 10, and 11 to 12 risk alleles. The regression line shows the association between the genetic risk score (GRS) and pack-years smoked by 38 years of age (Pearson correlation r = 0.12; P = .003). B, Ever-smokers at higher genetic risk were more likely to become nicotine dependent. The bars of the chart graph the proportion of ever-smokers at low (n = 157), average (n = 292), and high (n = 178) genetic risk who became nicotine dependent (≥4 Fagerström symptoms) by 38 years of age and who were nicotine dependent at 2 or more assessments. C, Smokers at higher genetic risk were more likely to experience cessation failure during their 30s. The bars of the chart graph the proportions of daily smokers at low, average, and high genetic risk who experienced relapse after a quit attempt lasting 1 month or longer and who achieved successful cessation (abstinence ≥1 year) through 38 years of age. Percentage with relapse was calculated from cohort members who quit smoking for 1 month or longer during 32 to 38 years of age (n = 36 for the low genetic risk group, n = 61 for the average genetic risk group, and n = 34 for the high genetic risk group). Percentage with successful cessation was calculated for cohort members who smoked daily during their 30s (n = 65 for the low genetic risk group, n = 120 for the average genetic risk group, and n = 77 for the high genetic risk group). B and C, Low genetic risk individuals had GRSs more than 0.5 SD below the cohort mean, average genetic risk individuals had GRSs within 0.5 SD of the cohort mean, and high genetic risk individuals had GRSs more than 0.5 SD above the cohort mean. Error bars reflect SEs.

Genetic risks discovered by these GWAs do not predict smoking initiation is not entirely unexpected. Nevertheless, that so-called chippers (light but persistent smokers) in our cohort had below average genetic risk is consistent with the theory that the genetic risks captured in our score influence response to nicotine, not the propensity to initiate smoking. Thus, our result affirms the value of using former and light smokers as a comparison group to heavy and nicotine dependent smokers in discovery analyses targeting these risks.

Previous research has related polymorphisms in the genes included in our genetic risk score to developmental phenotypes of smoking behavior and to mature phenotypes of adult smoking problems. To our knowledge, ours is the first study to track the relations of particular genetic risk variants with the development of smoking behavior from initiation through conversion to daily smoking and progression to heavy smoking and on to the mature phenotypes of persistent heavy smoking, nicotine dependence, and struggles with cessation through midlife. Moreover, this extended follow-up allowed us to find, for the first time, that GWAS-identified variation in 15q25.1 and 19q13.2 influences adult smoking problems through a pathway mediated by adolescent progression from smoking initiation to heavy smoking. Our study is also the first, to our knowledge, to find that GWAS-identified SNPs provide information about smoking risks that cannot be ascertained from a family history, including information about risk for cessation failure.

These findings should be considered in light of 3 limitations. First, although the Dunedin Study sample consisted of European-descent individuals, as did the samples analyzed in the GWASs used to develop the GRS, we cannot rule out the possibility of population stratification. Further, replication in other populations is needed. Second, our analyses of cessation were subject to censored data. Third, the 4 decades of follow-up in the Dunedin Study coincided with major secular events, such as bans against smoking in the workplace. Comparisons of cohorts born at different times might elucidate gene-policy interactions in smoking behavior and speak to the generalizability of the current findings. Despite these limitations, this study has implications for etiologic research and public health. With respect to etiology, our study makes 3 contributions. First, next-generation sequencing studies and other efforts to ascertain causal variants responsible for GWAS signals may maximize their discovery potential by focusing on samples of young people strategically selected to reflect important developmental transitions. Such work could use experimental designs to test hypotheses about mechanisms of genetic risk on postinitiation phenotypes. Second, we demonstrated that a GRS based on the assumption of additive risks can be used to follow up GWAS results in a birth cohort far smaller than the original discovery samples. Future etiologic research can use GRPs to apply GWAS results to longitudinal studies. Third, results are consistent with the hypothesis in pediatric medicine that some adolescents, after only experimental use, are prone to quickly become heavy users and dependent. This finding suggests that gene-environment interaction analyses of smoking and nicotine dependence may...
profit from a focus on environments that coincide with or immediately precede the adolescent period and influence the propensity of children at high genetic risk to initiate smoking. Smoking by peers is one such environment. Tobacco control policies targeting youth may be another.

Turning to public health, our research adds a genetic dimension to long-standing arguments that early prevention could be a critical strategy in reducing cigarette consumption. Specifically, our findings and others' suggest that initiatives that disrupt the developmental progression of smoking behavior, such as surtaxes and age restrictions on tobacco purchases, may ameliorate some genetic risks. Moving beyond population-level prevention, we found that information about smoking risk captured in a score composed of GWAS-identified variants was independent of information that could be derived from a family history of smoking behavior. This novel finding suggests that genetic information could be used to identify high-risk youngsters for targeted prevention. However, the associations we detected between the GRS and smoking phenotypes were small in magnitude. Small effect sizes do not preclude public health relevance, but they caution against the use of genetic information to evaluate risk in individuals; children who our study would classify at high genetic risk are not guaranteed to become addicted if they try smoking, and, even more importantly, children we would classify at low genetic risk are not immune to addiction. The public health use of the current findings must be tempered with recognition that most risk-associated genetic variation does not determine poor health outcomes, and, correspondingly, its absence does not guarantee protection.

Submitted for Publication: March 9, 2012; final revision received July 12, 2012; accepted September 17, 2012. Psychiatry Online: March 27, 2013. doi:10.1001/jamapsychiatry.2013.736

Author Affiliation: Department of Health Policy and Management, Gillings School of Public Health (Dr Belsky and Biddle), and Department of Genetics, School of Medicine (Dr Evans), University of North Carolina–Chapel Hill; Center for the Study of Aging and Human Development (Dr Belsky) and Department of Psychiatry and Behavioral Sciences (Dr Belsky, Moffitt, Houts, Meier, Sugden, and Caspi; Ms Harrington; and Mr Williams), Duke University Medical Center, and Institute for Genome Sciences and Policy (Dr Belsky, Moffitt, Houts, Meier, Sugden, and Caspi; Ms Harrington; and Mr Williams) and Department of Psychology and Neuroscience (Dr Belsky, Moffitt, Houts, Meier, Sugden, and Caspi; Ms Harrington; and Mr Williams), Duke University, Durham, North Carolina; Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, King’s College London, London, United Kingdom (Dr Moffitt, Sugden, and Caspi and Mr Williams); Center for Tobacco Research and Intervention, University of Wisconsin, and Department of Medicine, University of Wisconsin-Madison School of Medicine and Public Health, Madison (Dr Baker); and Dunedin Multidisciplinary Health and Development Research Unit, University of Otago, Otago, New Zealand (Dr Poulton).

Correspondence: Avshalom Caspi, PhD, Duke University/ Duke University Medical Center, Center for Aging and the Study of Human Development, 2020 W Main St, Ste 201, Durham, NC 27708 (ac1156@duke.edu).

Funding/Support: This research received support from grant AG032282 from the US National Institute on Aging, grant MH077874 from the US National Institute on Mental Health, and grant G0101483 from the UK Medical Research Council. The Dunedin Multidisciplinary Health and Development Research Unit is supported by The New Zealand Health Research Council. Dr Belsky was supported in part by grant 1R36HS020524-01 from the US Agency for Healthcare Research and Quality and grant T32-AI000029 from the National Institute on Aging. Dr Baker was supported in part by grant 1K05CA139871 from the US National Cancer Institute. Dr Meier was supported in part by grant P30DA023026 from the US National Institute on Drug Abuse. Additional support was provided by the Jacobs Foundation.


Additional Contributions: We thank the Dunedin Study members, their families, unit research staff, and study founder Phil Silva.

REFERENCES
