

FAQ Repository

What did the study examine? In this study, we focused on tobacco use disorder, which was identified based on diagnostic codes found in electronic health records. Our aim was to investigate the genetic profile and biological mechanisms associated with tobacco use disorder. By doing so, we hoped to uncover new insights that could lead to the development of effective interventions for treating this condition. **Link to study:** <https://pubmed.ncbi.nlm.nih.gov/37034728/>

What year was it published? 2023

Key messages:

1. The study identified 88 genetic variants and uncovered 461 potential candidate genes for tobacco use disorder.
2. Tobacco use disorder shares biological processes common to many substance use disorders and is genetically correlated with other psychiatric and medical disorders.
3. The study demonstrated that electronic health records are a viable and cost-effective method for the genetic study of tobacco use disorder.
4. This study included individuals of multiple ancestries (European, African American, Latin American) to obtain a more complete picture of the genetic profile of tobacco use disorder.

Corresponding author of FAQ: Sandra Sanchez-Roige (sanchezroige@health.ucsd.edu)

Additional link:

Frequently Asked Questions (FAQ)

This document provides information about the study:

Toikumo et al. (2023). “Multi-ancestry meta-analysis of tobacco use disorders prioritizes novel candidate risk genes and reveals associations with numerous health outcomes.” *Nature Human Behavior*.

The document draws from and builds upon FAQs offered in papers written by the Social Science Genetic Association Consortium (SSGAC), the Externalizing Consortium and others available at the FAQs on Genomic Studies (FoGS) repository. We acknowledge the valuable contribution these papers have made to the field.

For clarifications or additional questions, please contact Dr. Sandra Sanchez-Roige (sanchezroige@health.ucsd.edu).

Who conducted this study?

We are members of the PsycheMERGE Network Substance Use Disorders Working Group. Our goal is to identify risk factors that make individuals more likely to develop substance use disorders. For more information, you may visit the website linked [here](#).

What is tobacco use disorder?

Tobacco use disorder is a condition that affects millions of people around the world and is the leading cause of illness and preventable death in the United States. Individuals with this condition have a persistent and excessive use of tobacco, which is found in products like cigarettes and cigars. This maladaptive use of tobacco can have a profound impact on the individual's life, causing significant personal and work life problems. Additionally, long-term use of tobacco can cause other illnesses, such as lung cancer. Individuals who want to stop tobacco use due to these problems often experience difficulties. For instance, when attempting to quit, individuals may experience withdrawal symptoms. Withdrawal symptoms can vary from person to person but commonly include cravings, irritability, restlessness, difficulty concentrating, anxiety, depressed mood, increased appetite, and sleep disturbances. These symptoms can be uncomfortable and make it challenging to quit tobacco use.

Why is it important to study tobacco use disorder?

Previous studies have found that some of the risk of an individual developing tobacco use disorder is associated with the genetic variants a person has. Genetic studies, like ours, allow researchers to identify these genetic variants to better understand the biology of tobacco use disorder. This is vital because understanding the biology can lead to new treatments to help people with tobacco use disorder.

What did this study do?

This study conducted a genome-wide association study (GWAS) of tobacco use disorder. A GWAS measures hundreds of thousands of genetic variants and correlates each one with the characteristic that is being studied. In this study, GWAS is estimating which genetic variants are more common in people with tobacco use disorder compared to people who don't have tobacco use disorder. A GWAS is a correlational study. This means that a GWAS can identify genetic variants that are more common in one group of people compared to another, but it does not necessarily identify genetic variants that *cause* the outcome that is being studied.

Because GWAS need extremely large sample sizes, we used data available from electronic health records (see below) to conduct our analyses. Additionally, we examined the potential overlap between genetic variants linked to tobacco use disorder and genetic variants associated with other conditions. Lastly, we investigated whether the genes associated with tobacco use disorder have connections to currently available pharmacological medications.

What are electronic health records, and why did you use them for this study?

Electronic health records (EHR) are recorded every time we visit the doctor, as part of the medical appointment. A medical professional will record codes that represent your diagnosis, notes about your condition, and other related information. In this study, we used this data to identify individuals with and without tobacco use disorder diagnostic codes in their records. This approach is highly beneficial because EHR data are already available for millions of patients. Some of these individuals have also enrolled in biobanks, meaning that their genetic information is also available. We can therefore combine their EHR data with their genetic data, and use it to perform genetic studies of tobacco use disorders in a very large sample. Using a large cohort is critical to discover new genes.

It is important to note that these records are often recorded for billing purposes and not for research purposes. This means that they are subject to a lower level of precision than information collected exclusively for research would be. This can lead to missing data or incorrect information. For instance, an individual may have a diagnosis not recorded in their medical record because they haven't seen a doctor in that healthcare system for that issue. On the other hand, some individuals may have a diagnostic code in their medical records because they were being evaluated for that issue, but they were then determined to have a different diagnosis. Because of this, we applied additional measures to check the quality of the phenotype we studied.

Who participated in this study, and why does it matter?

The individuals included in this study come from multiple biobanks (UK Biobank, Vanderbilt University Medical Center Biobank, Mass General Brigham Biobank, Penn Medicine Biobank, Million Veteran Program). The individuals whose samples are included in these biobanks have shared European ancestry, shared African American ancestry, and shared Latin American ancestry. It is important to include individuals from different ancestral backgrounds in genetic studies because it provides a more complete understanding of the genetic basis of tobacco use disorder, which contributes to a more inclusive and accurate understanding of the disorder. The study of only one group of genetically similar individuals (for example, individuals of shared

European ancestry) could worsen health disparities by aiding discoveries that will disproportionately benefit only that population.

What is genetic ancestry and how is it different from race?

Genetic ancestry refers to an individual's genetic background, traced through genetic testing. It seeks to identify the genetic origins of an individual, including their ancestral populations and the geographic regions where their ancestors likely came from. Genetic ancestry has a different meaning than race. Race is a social construct that categorizes people into distinct groups based on shared physical characteristics such as skin color, hair texture, and facial features.

The key difference between genetic ancestry and race is that genetic ancestry is based on an individual's actual genetic composition. Genetic ancestry recognizes the complexity and diversity within and across populations. Genetic ancestry is identified by examining genetic variations and similarities among populations, allowing for a more accurate understanding of an individual's genetic background. In contrast, race is a social construct that categorizes individuals based on perceived physical characteristics.

What are the key findings?

Our study revealed at least four major findings:

First, we identified that tobacco use disorder is heritable, meaning that variation in the risk of being diagnosed with tobacco use disorder is associated with genetic variation between individuals. Genetic variants only accounted for a small proportion of risk for tobacco use disorder. This suggests that, while we are interested in finding biological causes that contribute to behavior, the environment plays an equally (or perhaps even more important) role in shaping tobacco use behaviors.

Second, we identified 88 genetic variants associated with tobacco use disorder. When we mapped those variants onto genes, we uncovered hundreds of genes associated with tobacco use disorder. This discovery opens up new avenues for experimental testing and further investigation.

Third, when combined, genetic variants associated with tobacco use disorder were correlated with genetic variants associated with many other conditions, including other psychiatric conditions and medical diseases, such as HIV infection, heart disease, and pain. This is known as

'genetic correlation'. However, it is important to note that these are correlations and whether these correlations are causal is not yet known.

Lastly, some of the genetic variants identified were located in genes that are associated with medications that are often used for smoking cessation (e.g., varenicline), as well as medications for other conditions that tend to co-occur with tobacco use disorders, such as antipsychotics and antidepressants.

In summary, our study sheds light on the heritability of tobacco use disorder, reveals numerous genetic variants and associated genes, identifies genetic correlations with other conditions, and provides insights into potential medications for treatment.

Did you find "the gene for" tobacco use disorder?

No - tobacco use disorder, like virtually all complex traits, is not a single-gene condition. In fact, this study identified hundreds of genetic variants associated with tobacco use disorder. However, it is important to understand that genetic variants alone do not cause the disorder. Both genetic and environmental factors contribute to an individual's risk to develop tobacco use disorder.

What is a "polygenic score"? How did the study use "tobacco use disorder polygenic scores"?

A polygenic score summarizes the genetic risk of an individual for tobacco use disorder into a single numeric score. But polygenic scores are not definitive diagnostic tools, as they merely estimate genetic risk, and do not determine whether an individual will develop tobacco use disorder with certainty. Both environmental and lifestyle factors play a significant role in the development of tobacco use disorder, and these should be considered alongside polygenic scores.

Does this study show that an individual's level of tobacco use is determined and fixed at conception?

No. The environment plays an important role in determining one's levels of tobacco use and misuse. For example, in a society where tobacco was unavailable, it wouldn't matter what the genetic risk for tobacco use disorder is, as the likelihood of developing tobacco problems would be determined entirely by the environment (i.e., no one would develop tobacco use disorder because tobacco is not available). Furthermore, supportive and nurturing environments can reduce the likelihood that any individual will develop tobacco misuse. Similarly, environmental

adversities can increase the likelihood that any individual will develop tobacco misuse. This is why it is important to consider both the genes and environment when determining an individual's outcomes.

Could studying genetic influences on tobacco use disorder lead to discrimination?

It is crucial to interpret these findings responsibly and avoid using them to discriminate against individuals with specific genetic variants. In the past, genetic research has been misused to support harmful ideas, suggesting that some people are naturally inferior or that social inequality is inevitable. **More specifically to our study, discrimination might arise if individuals with certain genetic markers associated with tobacco use disorder are unfairly treated or if unfounded assumptions about their vulnerability for tobacco use disorder are made.** We strongly reject such claims. We believe that genetic influences do not determine a person's worth or abilities, and that environmental factors are just as significant in shaping who we are. It is our duty as researchers to share this knowledge openly and ensure it is not used inappropriately or to perpetuate discrimination.