

Candidate Genes from an FDA-Approved Algorithm Fail to Predict Opioid Use Disorder Risk in Over 450,000 Veterans

Christal N. Davis^{1,2}, Zeal Jinwala^{1,2}, Alexander S. Hatoum³, Sylvanus Toikumo^{1,2}, Arpana Agrawal⁴, Christopher T. Rentsch^{5,6,7}, Howard J. Edenberg⁸, James W. Baurley⁹, Emily E. Hartwell^{1,2}, Richard C. Crist^{1,2}, Joshua C. Gray¹⁰, Amy C. Justice^{5,6}, Joel Gelernter^{5,11,12}, Rachel L. Kember^{1,2}, & Henry R. Kranzler^{1,2}

¹ Mental Illness Research, Education and Clinical Center, Crescenz Veterans Affairs Medical Center, Philadelphia, PA, USA

² Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

³ Department of Psychological and Brain Sciences, Washington University School of Medicine, St. Louis, MO, USA

⁴ Department of Psychiatry, Washington University, St. Louis, MO, USA

⁵ Veterans Affairs Connecticut Healthcare System, West Haven, CT, USA

⁶ Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA

⁷ Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK

⁸ Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA

⁹ BioRealm LLC, Walnut, CA, USA

¹⁰ Department of Medical and Clinical Psychology, Uniformed Services University, Bethesda, MD, USA

¹¹ Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA

¹² Departments of Genetics and Neuroscience, Yale University School of Medicine, New Haven, CT, USA

Corresponding Author: Henry R. Kranzler, Center for Studies of Addiction, University of Pennsylvania Perelman School of Medicine, 3535 Market Street, Suite 500, Philadelphia, PA 19104; Telephone: 215-746-1943; Email: kranzler@pennmedicine.upenn.edu

Disclosures: Dr. Kranzler is a member of advisory boards for Dicerna Pharmaceuticals, Sophrosyne Pharmaceuticals, Enthion Pharmaceuticals, and Clearmind Medicine; a consultant to Sobrera Pharmaceuticals; the recipient of research funding and medication supplies for an investigator-initiated study from Alkermes; and a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which was supported in the last three years by Alkermes, Dicerna, Ethypharm, Lundbeck, Mitsubishi, Otsuka, and Pear Therapeutics, Drs. Kranzler and Gelernter hold U.S. patent 10,900,082 titled: "Genotype-guided dosing of opioid agonists," issued 26 January 2021.

Acknowledgments: This research is based on data from Million Veteran Program, Office of Research and Development, Veterans Health Administration and was supported by awards #I01 BX003341 and IK2 CX002336 (to EEH); the VISN 4 Mental Illness Research, Education, and Clinical Center; NIAAA grant K01 AA028292 (to RLK); and NIDA grant P30 DA046345. An MVP Core Acknowledgement list is provided in the supplemental materials. This publication does not represent the views of the Department of Veteran Affairs, the Department of Defense, Uniformed Services University, or the United States Government.

Abstract Word Count: 313

Text Word Count: 2941

Number of Tables: 2

Number of Figures: 3

Key Points

Question: How well do candidate genes from an algorithm designed to predict risk of opioid use disorder, which recently received pre-marketing approval by the Food and Drug Administration, perform in a large, independent sample?

Findings: In a case-control study of over 450,000 individuals, the 15 genetic variants from candidate genes collectively accounted for 0.4% of the variation in opioid use disorder risk. In this independent sample, the SNPs predicted risk at a level of accuracy near random chance (52.8%).

Meaning: Candidate genes from the approved genetic risk algorithm do not meet standards of reasonable clinical efficacy in assessing risk of opioid use disorder.

Abstract

Importance: Recently, the Food and Drug Administration gave pre-marketing approval to algorithm based on its purported ability to identify genetic risk for opioid use disorder. However, the clinical utility of the candidate genes comprising the algorithm has not been independently demonstrated.

Objective: To assess the utility of 15 variants in candidate genes from an algorithm intended to predict opioid use disorder risk.

Design: This case-control study examined the association of 15 candidate genetic variants with risk of opioid use disorder using available electronic health record data from December 20, 1992 to September 30, 2022.

Setting: Electronic health record data, including pharmacy records, from Million Veteran Program participants across the United States.

Participants: Participants were opioid-exposed individuals enrolled in the Million Veteran Program (n = 452,664). Opioid use disorder cases were identified using International Classification of Disease diagnostic codes, and controls were individuals with no opioid use disorder diagnosis.

Exposures: Number of risk alleles present across 15 candidate genetic variants.

Main Outcome and Measures: Predictive performance of 15 genetic variants for opioid use disorder risk assessed via logistic regression and machine learning models.

Results: Opioid exposed individuals (n=33,669 cases) were on average 61.15 (SD = 13.37) years old, 90.46% male, and had varied genetic similarity to global reference panels. Collectively, the 15 candidate genetic variants accounted for 0.4% of variation in opioid use disorder risk. The accuracy of the ensemble machine learning model using the 15 genes as predictors was 52.8% (95% CI = 52.1 - 53.6%) in an independent testing sample.

Conclusions and Relevance: Candidate genes that comprise the approved algorithm do not meet reasonable standards of efficacy in predicting opioid use disorder risk. Given the algorithm's limited predictive accuracy, its use in clinical care would lead to high rates of false positive and negative findings. More clinically useful models are needed to identify individuals at risk of developing opioid use disorder.

Introduction

Opioid misuse and opioid use disorder (OUD) are significant public health problems. In 2022, 6.1 million Americans aged 12 or older met criteria for OUD,¹ with 94.8% of these individuals acknowledging misuse of prescription analgesics and 40.9% reporting receipt of the medication from a healthcare provider.¹ Given a surge in opioid overdose deaths,^{2,3,4} efforts have been made to identify individuals at greatest risk of opioid misuse.

However, at a population level, common genetic variation accounts for a small proportion of differences in OUD liability.⁵ Polygenic scores (PGS) derived from common genetic risk variants (i.e., single nucleotide polymorphisms, SNPs) across the genome account for much less variance in OUD risk (3.74%) than sociodemographic factors (41.32%).⁶ Nonetheless, attempts have been made to develop and commercialize genetic algorithms to identify individuals susceptible to developing OUD.⁷⁻⁹ These models usually include a few SNPs in genes that are considered causal candidates based on their presumed effect on neural systems that underlie OUD risk. In addition to their small effects,¹⁰ few candidate variants have been substantiated in genome-wide association studies (GWAS), a more rigorous method of identifying risk variants.¹¹⁻¹³

Genetic predictive models are also vulnerable to confounding based on differences in patterns of genetic similarity, which can create the false appearance of strong prediction.¹⁴ If the prevalence of OUD varies among individuals with historically different geographic origins (either due to real differences or biases in the data used to train a machine learning (ML) model), the model will falsely attribute predictive power to SNPs

that are markers of global or local similarity rather than disorder risk. Such spurious associations arise from population stratification, i.e., differences in allele frequencies that result from historical migration and mating patterns (Supplementary Figure 1). Used in real-world settings, these ML models are likely to bias predictions and lead to the false conclusion that they are useful for predicting risk of complex traits like OUD.

The U.S. Food and Drug Administration (FDA) recently gave pre-marketing approval to an algorithm that incorporates 15 SNPs with proprietary weights to predict OUD risk.¹⁵ We assessed these SNPs individually and collectively in relation to OUD in a genetically diverse sample of over 450,000 U.S. Veterans (Table 1) using logistic regression and ML analyses. Specifically, we assessed (1) whether the SNPs were individually associated with OUD risk, (2) how much variance in OUD risk the SNPs accounted for collectively, (3) whether the SNPs were better predictors of genetic similarity than OUD risk, and (4) whether basic demographics (i.e., age and sex) were better predictors of OUD risk than the SNPs.

Methods

Participants

The Million Veteran Program (MVP), an initiative of the U.S. Department of Veterans Affairs to discover the genetic etiology of diseases relevant to Veterans, was approved by the Central Veterans Affairs Institutional Review Board (IRB) and all site-specific IRBs. We followed all relevant ethical regulations for research with human subjects. Using electronic health record (EHR) data, we identified patients enrolled in MVP who had filled an outpatient prescription for any opioid analgesic. Cases were defined based on having at

least one International Classification of Diseases (ICD)-9/10 OUD code. Controls were opioid-exposed individuals with neither an OUD diagnosis code nor any prescription fill for medications commonly used to treat OUD (i.e., buprenorphine, methadone, and buprenorphine/naloxone). This yielded 452,664 opioid-exposed individuals ($M_{\text{age}} = 61.15$, $SD = 13.37$; 90.46% male) with genotype data, including 33,669 cases. As sensitivity analyses, we removed cases who had received only one outpatient OUD diagnosis code ($n = 8,355$) and re-ran all models, which yielded similar results (Supplementary Tables). The sample's genetically inferred ancestry (GIA) composition, which was assigned based on patterns of similarity to reference genomes of individuals in the 1000 Genomes project,¹⁶ was 67.46% European (EUR), 20.90% African (AFR), 9.50% admixed American (AMR), 0.81% East Asian (EAS), and 0.07% South Asian (SAS), with 1.25% unassigned (Supplementary Figure 2).

Among opioid-exposed individuals, we also selected a subset with only short-term documented exposure to opioids (4-30 days total), as this reflects the relative opioid naivete of patients for whom genetic testing would likely be sought. This yielded 125,514 individuals (3,704 cases) whose mean age was 59.98 ($SD = 14.84$) and who were predominantly male (90.37%). This subsample's GIA composition was similar to the full sample (67.59% EUR, 19.22% AFR, 10.36% AMR, 1.25% EAS, 0.11% SAS, and 1.47% unassigned).

Genotyping and Imputation

MVP samples were genotyped using a custom Affymetrix Axiom Biobank Array (MVP Release 4). Quality control and imputation were performed by the MVP Genomics Working

Group.¹⁷ Duplicate samples and those with a sex mismatch, seven or more relatives in MVP (kinship > 0.08), excessive heterozygosity, or a genotype call rate < 98.5% were removed. Monomorphic variants and variants with high missingness (call rate < 0.8) or a Hardy–Weinberg equilibrium p -value < 1×10^{-6} were removed. Genotypes were phased with SHAPEIT4¹⁸ and imputed using Minimac4 software,¹⁹ with biallelic SNPs imputed using the African Genome Resources reference panel by the Sanger Institute.

To infer similarity to reference genomes and generate principal components (PCs) to account for global and local differences in genetic similarity, GIA composition was calculated by the MVP gwPheWAS Working Group.²⁰ Reference genomes used to infer genetic ancestry were from individuals from the Yoruba in Ibadan, Nigeria (YRI), the Luhya in Webuye, Kenya (LWK), the British in England and Scotland (GBR), the Han Chinese in Beijing, China (CHB), and Peruvians from Lima, Peru (PEL). Further details on how subgroups were derived from genetic similarity PCA coordinates by the MVP gwPheWAS Working Group can be found in Hunter-Zinck et al.¹⁷ In summary, a random forest classifier was trained on the reference dataset using the first ten PCs. The algorithm was applied to the MVP principal components analysis (PCA) data, and GIA was inferred when the classifier's predicted probability was greater than 50%. Individuals with lower assignment probabilities were retained in the full sample analyses. Although guidelines on the use of population descriptors in human genetics research²¹ were published after GIA analyses were conducted in MVP, we follow the recommendations as closely as possible.

Single-SNP Association Analyses

We conducted association analyses for each of the 15 candidate SNPs with OUD case/control status using logistic regressions in PLINK 2.0.²² Analyses were also performed within GIA groups and the subset of short-term opioid-exposed individuals. To account for non-independence, we randomly removed one individual from each pair of related individuals (N = 24,585). Three sets of analyses were conducted: (1) with no covariates, (2) including the first 10 genetic similarity PCs, and (3) including age, sex, and genetic similarity PCs as covariates.

Combined-SNP Regression Analyses

Using the `glm()` function in R, we fit logistic regressions to examine the collective association of the 15 SNPs with OUD case/control status. We followed a procedure similar to that used with the single-SNP models, performing analyses in the full sample, short-term opioid-exposed subsample, and GIA groups. We also ran three sets of models with increasing levels of adjustment, akin to the single-SNP models. We calculated Nagelkerke's R^2 to estimate the proportion of variance in OUD status accounted for by each model and the area under the receiver operating characteristic curve (AUC-ROC) to evaluate performance.

Machine Learning Models

We developed ensemble ML models in the full sample, the subset of short-term opioid-exposed individuals, and within GIA groups. We chose to use ML techniques, as they are useful for modeling non-linear relationships between variables. However, their performance is constrained by the strength of the underlying genetic effects on OUD risk.

To examine ML's ability to predict OUD vs. genetic similarity, additional models were developed using data confounded by inferred genetic ancestry, such that all cases were in one GIA group and controls in another.¹⁴ The confounded models show how biases in the patterns of population stratification within training datasets can influence model performance. All ML analyses were performed using the *caret* package in R, which allows for consistency in the data structure and preprocessing steps across models.²³

To prepare the analysis, dosage data for each of the 15 SNPs were recoded as hard calls (i.e., "0", "1", or "2"), reflecting the number of risk alleles for each individual. Of the 15 genetic variants of interest, 8 were directly genotyped and 7 were imputed, all with imputation quality > 0.8. In all models, controls were under-sampled to yield equal proportions of cases and controls, which was required to address the severe imbalance in the number of cases relative to controls. In confounded models, we sought to preserve the greatest number of cases. Thus, in all models where GIA was unbalanced, the more common GIA group was assigned as cases and the less common one as controls. We used a 75%/25% data split to obtain independent training and testing sets. To prevent overfitting, we used 10-folds cross-validation during training, which systematically validates the model across subsets ("folds") of the data. We used *caret*'s default tuning parameter search feature, with the optimal parameters selected based on ROC to provide the best balance between sensitivity and specificity.

We selected as base ML models two random forest implementations, "rf" (Breiman and Cutler's random forests) and "ranger" (recursive partitioning with random forests). We also used a linear support vector machine (SVM) implementation (*svmlinear2*, using the

e1071 library²⁴ in caret) to approximate the methodology used for the algorithm that received pre-marketing approval, as the actual algorithm is proprietary.⁹ Base model predicted probabilities were aggregated with the 15 SNPs to serve as predictors in a stacked ensemble ML model, which was trained using a binomial generalized linear model. This step allows the ensemble model to learn from the predictions of each of the base ML models, which can be used to enhance its performance. Accuracy, sensitivity, and specificity were the primary metrics used to evaluate model performance, although we also examined positive and negative predictive values as secondary metrics. These metrics were used to calculate a diagnostic odds ratio (DOR), indicating how much higher the odds are that the classifier produces a positive prediction in an individual with OUD than an individual without OUD.

Results

In single-SNP models that did not account for genetic similarity, 13 of 15 SNPs were significantly associated with OUD risk after Bonferroni correction (Table 2). However, upon inclusion of measures of global genetic similarity (i.e., PCs) that number declined to three (Table 2). Five of the ten SNPs whose associations became nonsignificant when we controlled for genetic similarity had opposite directions of effect in analyses that were uncontrolled. In analyses within GIA groups that accounted for local variations in genetic similarity by including 10 PCs, the three significant SNPs were associated with OUD risk only in individuals genetically similar to the EUR superpopulation. Similar results were obtained among short-term opioid-exposed individuals (Supplementary Tables 1 and 2).

In full sample logistic regressions, the 15 SNPs collectively accounted for 0.40% of the variance in OUD status (Figure 1), with an AUC of 0.54. In comparison, age and sex alone accounted for 3.27% of the variance and improved the model AUC to 0.66. Including PCs as covariates reduced the number of significant SNP associations from 11 to 5. In analyses conducted within GIA groups, 7 SNPs were significant in the EUR group, 2 in the AFR group, and 1 in the AMR group. The variance collectively accounted for by the SNPs was between 0.04% (AFR) and 0.16% (AMR) in models conducted within GIA groups, and AUC values ranged from 0.51 (AFR) to 0.53 (AMR). Supplementary Tables 3 – 6 provide full results of the logistic regressions.

The accuracy of the ensemble ML model predicting OUD risk in the full sample (Figure 2) was slightly greater than chance (52.83%, 95% CI = 52.07 to 53.59%). In test data, the model correctly identified 50.72% of OUD cases (sensitivity) and 54.95% of controls (specificity). Of the model's predicted cases, 52.96% were true cases (positive predictive value), and of the predicted controls, 52.72% were true controls (negative predictive value). The DOR was 1.25, suggesting that a predicted case is 1.25 times as likely to have OUD than a predicted control. In analyses conducted within GIA groups that accounted for local variations in genetic similarity by including 10 PCs, accuracy did not exceed random chance (EUR: 50.65% [95% CI = 49.67 to 51.62%]; AFR: 50.53% [95% CI = 49.09 to 51.96%]; AMR: 49.69% [95% CI = 47.21 to 52.16%]). Similarly, the DOR dropped to between 0.99 (AMR) and 1.03 (EUR). Similar results were obtained for the short-term opioid-exposed subsample (Supplementary Tables 7 and 8).

Confounded models assessing the ability of the SNPs to differentiate individuals based on their genetic similarity to global superpopulations performed better than models predicting OUD status (Figure 3 and Supplementary Table 9). Models distinguishing individuals who were genetically similar to EUR from AFR superpopulations had an accuracy of 90.22% (95% CI = 89.63 to 90.79%) and those distinguishing AFR from AMR had an accuracy of 87.53% [95% CI = 86.56 to 88.46%]. The model was less accurate at distinguishing EUR from AMR (66.07% [95% CI = 65.15 to 66.99%]), likely due to the higher prevalence of genetic similarity between EUR and AMR individuals.²⁵

In the full sample ML models, age and sex alone yielded more accurate predictions of OUD risk (59.49% [95% CI = 58.82 to 60.16%]) than the 15 SNPs. The sensitivity and specificity of the model were 0.51 and 0.68, respectively. The accuracy of models within GIA groups was comparable for individuals who were genetically similar to AFR and AMR superpopulations (AFR: 57.92% [95% CI = 56.50 to 59.33%]; AMR: 57.78% [95% CI = 55.32 to 60.21%]), and better in individuals who were genetically similar to the EUR superpopulation (63.50% [95% CI = 62.55 to 64.43%]). Full results are in Supplementary Tables 10 and 11.

Discussion

In a diverse sample of over 450,000 opioid-exposed individuals (including 33,669 OUD cases), we found no evidence to support the clinical utility of the 15 candidate SNPs purported to predict OUD risk using a proprietary algorithm. Collectively, the SNPs accounted for 0.4% of the variance in OUD risk, consistent with small individual effects of common genetic variants on complex traits.^{26,27} In a hold-out testing subsample, we

observed high rates of false positives and negatives, with approximately 47 of 100 predicted cases or controls being incorrectly classified, which is no better than random chance prediction. False positives can contribute to stigma, cause patients undue concern, and bias healthcare decisions.²⁸ For example, providers exposed to information that emphasized genetic causes of mental health conditions had less empathy toward patients than providers presented with psychosocial causes.²⁹ False negatives could give patients and prescribers a false sense of security regarding opioid use and lead to inadequate treatment plans. Notably, clinicians could better predict OUD risk using an individual's age and sex than the 15 genetic variants. In summary, although the test approved by the FDA is intended to complement standard clinical assessment, its use is unlikely to confer additional benefits and may instead give providers and patients false and potentially harmful information.

We also found consistent evidence that when variations in genetic similarity were not considered, it resulted in biased findings. In single-SNP logistic regressions, 10 of 13 significant associations (77%) with OUD risk were rendered nonsignificant when measures of global genetic similarity were included as covariates. Half of these associations showed a reversal in the direction of effect when genetic similarity covariates were included. In contrast, the ML models highlighted the SNPs' ability to differentiate individuals who varied in their genetic similarity to global superpopulations with high accuracy, implying ancestral confounding in unadjusted models.¹⁴ To avoid spurious associations and ensure the validity of results, developers of genetic risk models intended for clinical care should

carefully consider and account for bias attributable to population stratification and sociodemographic factors. Failure to do so may exacerbate existing health inequities.³⁰

Several limitations of this study should be considered. First, models were evaluated using EHR diagnosis codes, whose assignment is susceptible to bias.³¹ However, comparable results were previously obtained in a smaller sample where diagnoses were assigned using structured interviews.¹⁴ Second, the MVP sample is predominantly male, though given its size, the analyses included over 40,000 women (2,619 OUD cases), which far exceeds the full sample on which the pre-marketing approved algorithm was developed ($N_{\text{total}} = 1,762$, $N_{\text{cases}} = 653$; EM-24069 AvertD™ Package Insert as of 12-15-2023). The MVP sample also has higher rates of OUD and pain³² and is older than the general population.³³ We encourage efforts to evaluate the 15 genetic variants in additional datasets. Third, we used GIA group—the current standard for MVP research—as a population descriptor. Despite being an improvement over previous descriptors that combined GIA with self-reported race and ethnicity,³⁴ we note that GIA population descriptors do not fully align with recent guidance.²¹ Finally, MVP uses array genotyping, which is less accurate than, for example, that involving mass spectrometry,³⁵ and imputation was required for about half of the SNPs. Although this may have reduced genotyping accuracy, which is important for individual-level risk prediction, it would not be expected to impact the results in a large sample like ours.

Genetics researchers have argued against the use of candidate genes to predict OUD and other psychiatric traits.³⁶ Because genetic risk models in psychiatry will continue to emerge and could prove to be clinically useful, it is crucial that researchers and

regulatory agencies adopt rigorous standards for developing and evaluating them prior to their application in clinical settings. When regulatory agencies evaluate genetic risk algorithms that use advanced statistical methodologies (e.g., ML), it is imperative to heed the guidance of scientific advisors and independently validate the findings. By applying rigorous standards to reduce sources of bias, the potential benefits of genetic risk models can be maximized while protecting patient safety and well-being.

Table 1. Opioid use disorder case/control status across genetically inferred ancestry groups.

	Any Opioid Exposure		Short-term Opioid Exposure	
	Cases	Controls	Cases	Controls
ALL	33669	418995	3704	121810
EUR	20392	284986	1982	82859
AFR	9496	85118	1211	22910
AMR	3215	39790	411	12586

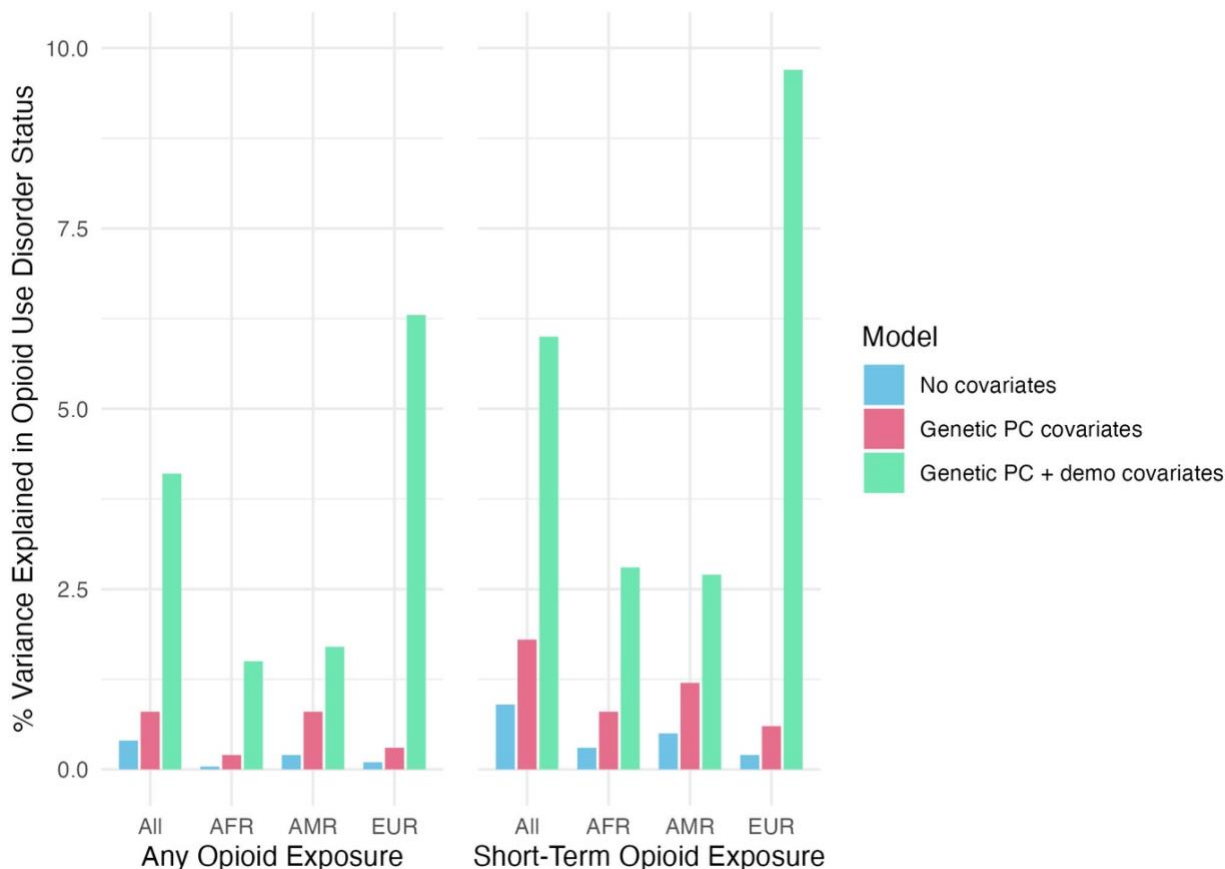
Note: ALL = full sample, including individuals genetically similar to the East Asian and South Asian superpopulations, and those missing on genetically-inferred ancestry, EUR = individuals genetically similar to the European superpopulation, AFR = individuals genetically similar to the African superpopulation, and AMR = individuals genetically similar to the admixed American superpopulation.

Table 2. Single nucleotide polymorphism (SNP) associations with opioid use disorder case/control status among opioid exposed individuals.

SNP (Gene)	Model	AFR OR (<i>p</i> -value)	AMR OR (<i>p</i> -value)	EUR OR (<i>p</i> -value)	ALL OR (<i>p</i> -value)
rs7997012 (HTR2A)	Unadjusted	1.057 (0.064)	1.005 (0.849)	0.993 (0.487)	0.893 (3e-39)
	Adjusted	1.018 (0.555)	1.003 (0.906)	0.998 (0.837)	0.998 (0.803)
rs2236861 (OPRD1)	Unadjusted	1.031 (0.301)	1.016 (0.633)	<i>1.035 (0.005)</i>	0.957 (3e-5)
	Adjusted	1.018 (0.546)	1.013 (0.696)	<i>1.032 (0.011)</i>	<i>1.028 (0.010)</i>
rs4680 (COMT)	Unadjusted	0.996 (0.801)	1.033 (0.229)	0.984 (0.126)	0.953 (4e-9)
	Adjusted	0.988 (0.474)	1.022 (0.430)	<i>0.978 (0.038)</i>	1.012 (0.173)
rs1045642 (ABCB1)	Unadjusted	0.982 (0.348)	0.993 (0.806)	1.009 (0.382)	0.902 (2e-37)
	Adjusted	0.965 (0.078)	0.990 (0.696)	1.010 (0.365)	0.985 (0.086)
rs1800497 (ANKK1)	Unadjusted	0.973 (0.102)	0.988 (0.673)	1.060 (4e-6)	1.079 (8e-16)
	Adjusted	0.978 (0.186)	1.050 (0.094)	1.056 (4e-5)	1.031 (0.002)
rs4532 (DRD1)	Unadjusted	1.044 (0.084)	0.993 (0.818)	1.020 (0.066)	0.931 (5e-16)
	Adjusted	1.025 (0.337)	0.969 (0.300)	1.019 (0.081)	1.015 (0.126)
rs948854 (GAL)	Unadjusted	0.989 (0.491)	0.996 (0.885)	<i>1.024 (0.044)</i>	1.124 (5e-43)
	Adjusted	0.976 (0.146)	0.968 (0.289)	1.020 (0.097)	1.017 (0.068)
rs211014 (GABRG2)	Unadjusted	0.988 (0.495)	0.964 (0.293)	<i>0.966 (0.009)</i>	1.042 (2e-5)
	Adjusted	0.992 (0.658)	0.960 (0.246)	<i>0.968 (0.015)</i>	<i>0.975 (0.013)</i>
rs1801133 (MTHFR)	Unadjusted	1.007 (0.776)	0.991 (0.735)	<i>1.026 (0.021)</i>	0.939 (2e-12)
	Adjusted	0.992 (0.756)	1.036 (0.209)	<i>1.023 (0.042)</i>	1.018 (0.056)
rs6347 (SLC6A3)	Unadjusted	1.000 (0.970)	1.041 (0.251)	1.004 (0.741)	1.107 (9e-27)
	Adjusted	1.009 (0.641)	0.992 (0.829)	1.004 (0.776)	1.005 (0.588)
rs1611115 (DBH)	Unadjusted	<i>0.952 (0.033)</i>	<i>0.922 (0.010)</i>	0.994 (0.672)	0.957 (4e-5)
	Adjusted	<i>0.950 (0.024)</i>	0.958 (0.179)	0.992 (0.562)	<i>0.979 (0.045)</i>
rs1051660 (OPRK1)	Unadjusted	0.967 (0.213)	1.063 (0.255)	0.980 (0.273)	0.985 (0.322)
	Adjusted	0.967 (0.223)	1.042 (0.447)	0.976 (0.192)	0.976 (0.102)
rs1799971 (OPRM1)	Unadjusted	1.032 (0.516)	0.881 (0.001)	0.885 (1e-13)	0.816 (8e-49)
	Adjusted	0.994 (0.905)	<i>0.910 (0.011)</i>	0.882 (2e-14)	0.894 (5e-15)
rs3758653 (DRD4)	Unadjusted	0.997 (0.908)	1.030 (0.375)	0.987 (0.392)	0.993 (0.532)
	Adjusted	0.998 (0.930)	1.066 (0.056)	0.984 (0.301)	0.999 (0.907)
rs9479757 (OPRM1)	Unadjusted	1.019 (0.424)	<i>1.115 (0.020)</i>	1.067 (2e-4)	1.098 (2e-12)
	Adjusted	1.022 (0.337)	<i>1.116 (0.019)</i>	1.070 (1e-4)	1.057 (3e-5)

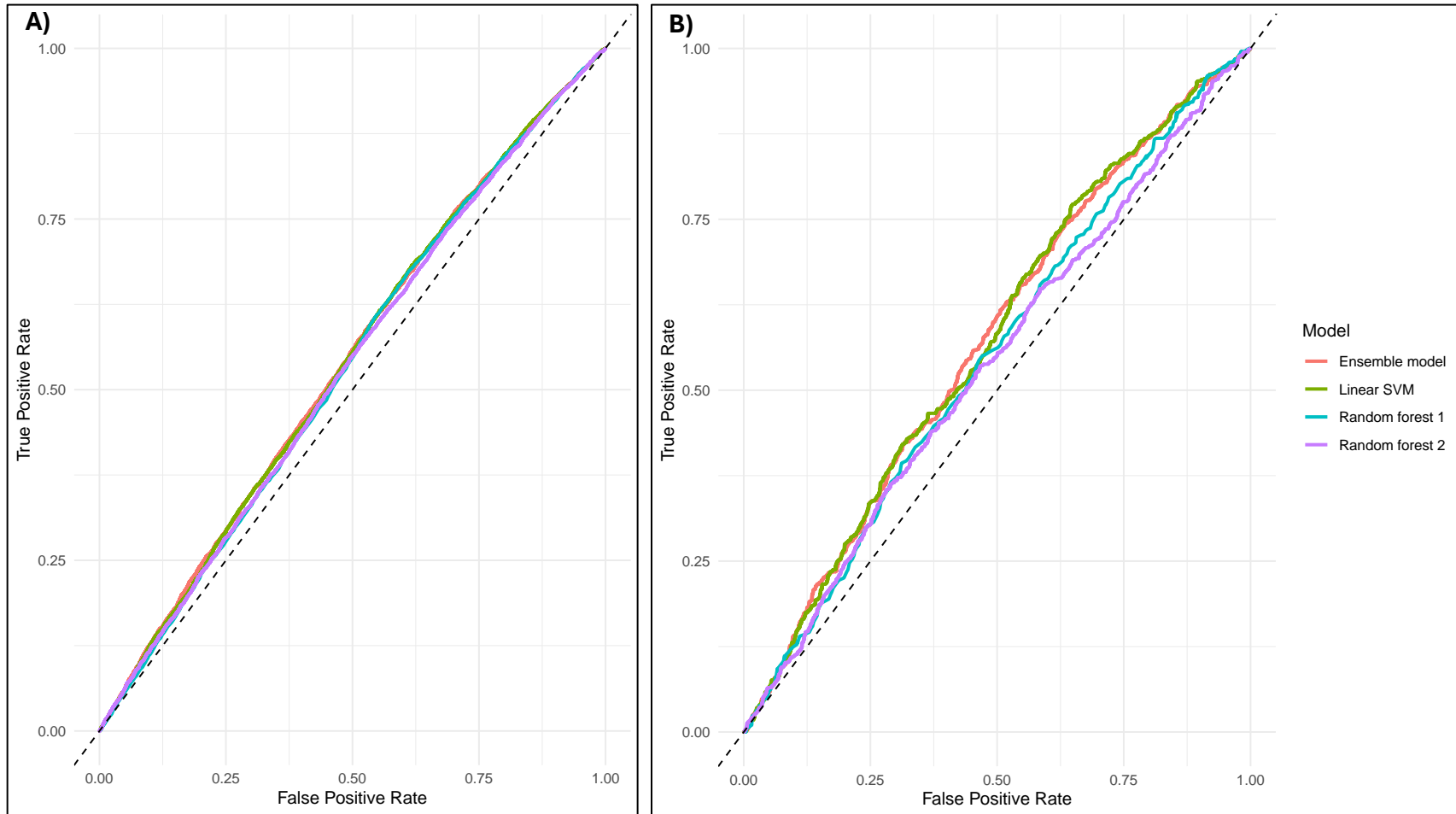
Note: Bold indicates significance at Bonferroni-adjusted *p*-value (0.05/15 = 0.003). Italics indicates significance at *p* < 0.05. Unadjusted models include the 15 SNPs as predictors, while adjusted models also include the first 10 genetic similarity PCs as covariates. OR = odds ratio, SE = standard error, SNP = single nucleotide polymorphism, AFR = African, AMR = Admixed American, EUR = European, ALL = full sample (including East Asian, South Asian, and participants missing on GIA). GIA is based on genetic similarity to global superpopulations defined by the 1000 Genomes Project.

Figure 1. Percent of variance in opioid use disorder case/control status explained by combined-SNP regression models.



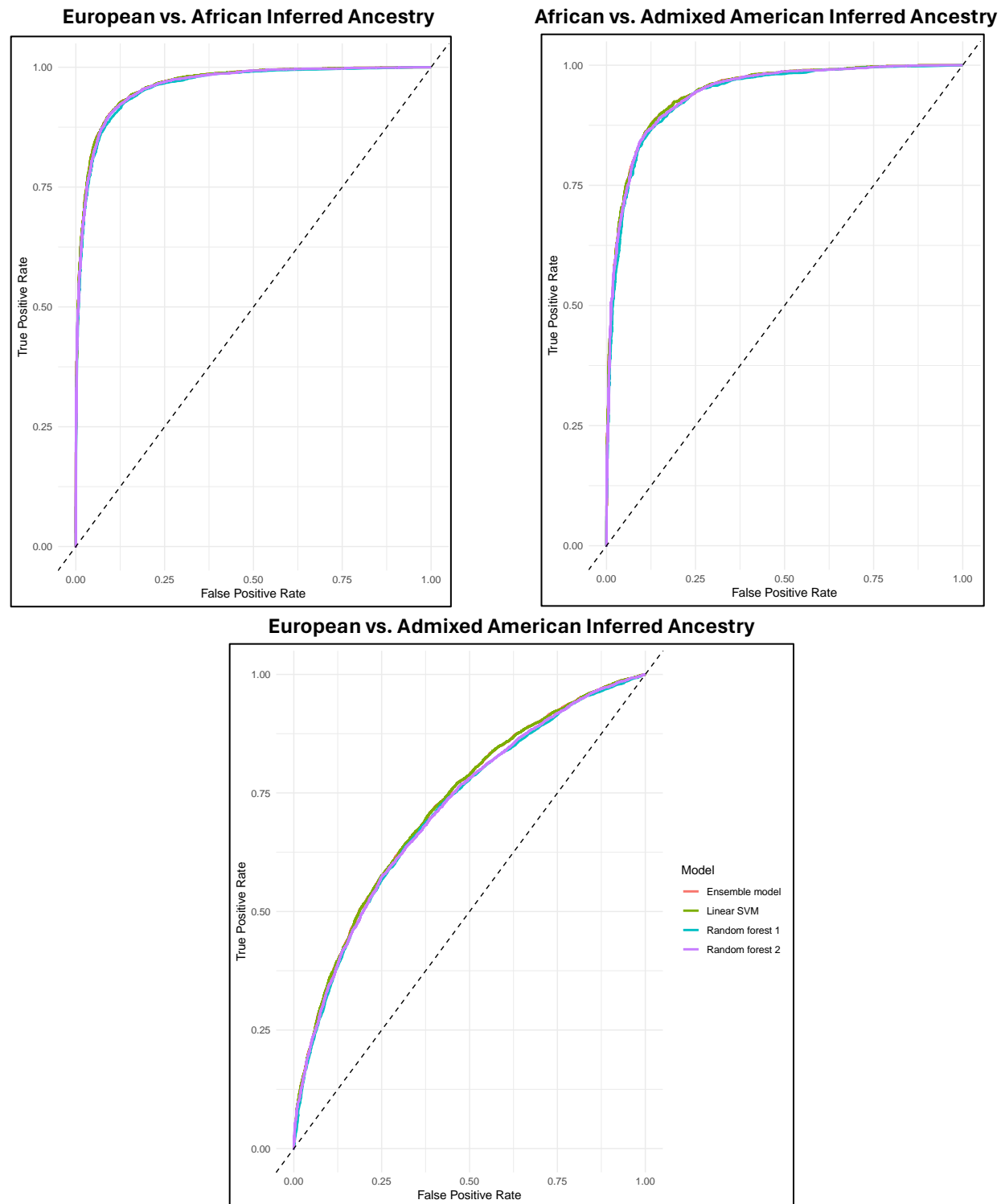
Note: AFR = individuals genetically similar to the African superpopulation, AMR = individuals genetically similar to the admixed American superpopulation, EUR = individuals genetically similar to the European superpopulation, OUD = opioid use disorder. No covariates model included only the 15 single-nucleotide polymorphisms, genetic PC covariates model included the first 10 principal components (PCs) indexing global variations in genetic similarity as covariates, and the genetic PC + demo covariates model included the first 10 PCs reflecting genetic similarity, age, and sex as covariates.

Figure 2. Receiver operating characteristic curves predicting opioid use disorder case/control status.



A) presents results from the full sample of individuals with any opioid exposure, **B)** presents results from the full sample of individuals with short term opioid exposure. Linear SVM = linear support vector machine model. The diagonal line represents a classifier model that predicts at chance levels.

Figure 3. Receiver operating characteristic (ROC) curves of models predicting genetically-inferred ancestry from 15 candidate single nucleotide polymorphisms.



Note: Linear SVM = linear support vector machine model. The diagonal line represents a classifier model that predicts at chance levels. Genetically-inferred ancestry was based on genetic similarity to global superpopulations defined by the 1000 Genomes Project.

References

1. Substance Abuse and Mental Health Services Administration. *Key substance use and mental health indicators in the United States: Results from the 2022 National Survey on Drug Use and Health*. 2023. <https://www.samhsa.gov/data/report/2022-nsduh-annual-national-report>
2. Dayer LE, Painter JT, McCain K, King J, Cullen J, Foster HR. A recent history of opioid use in the US: Three decades of change. *Substance Use & Misuse*. 2019/01/28 2019;54(2):331-339. doi:10.1080/10826084.2018.1517175
3. Tanz LJ, Gladden RM, Dinwiddie AT, et al. Routes of drug use among drug overdose deaths--United States, 2020-2022. *MMWR: Morbidity & Mortality Weekly Report*. 2024;73(6)
4. Ahmad F, Cisewski J, Rossen L, Sutton P. Provisional drug overdose death counts. <https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm>
5. Kember RL, Vickers-Smith R, Xu H, et al. Cross-ancestry meta-analysis of opioid use disorder uncovers novel loci with predominant effects in brain regions associated with addiction. *Nature Neuroscience*. 2022/10/01 2022;25(10):1279-1287. doi:10.1038/s41593-022-01160-z
6. Na PJ, Deak JD, Kranzler HR, Pietrzak RH, Gelernter J. Genetic and non-genetic predictors of risk for opioid dependence. *Psychological Medicine*. 2024:1-8. doi:10.1017/S0033291723003732
7. Moran M, Blum K, Ponce JV, et al. High genetic addiction risk score (GARS) in Chronically prescribed severe chronic opioid probands attending multi-pain clinics: an open clinical pilot trial. *Molecular Neurobiology*. 2021/07/01 2021;58(7):3335-3346. doi:10.1007/s12035-021-02312-1
8. Donaldson K, Demers L, Taylor K, Lopez J, Chang S. Multi-variant genetic panel for genetic risk of opioid addiction. *Ann Clin Lab Sci*. Aug 2017;47(4):452-456.
9. Donaldson K, Cardamone D, Genovese M, Garbely J, Demers L. Clinical performance of a gene-based machine learning classifier in assessing risk of developing opioid use disorder in subjects taking oral opioids: a prospective observational study. *Ann Clin Lab Sci*. Jul 2021;51(4):451-460.
10. Gratten J, Wray NR, Keller MC, Visscher PM. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nature Neuroscience*. 2014/06/01 2014;17(6):782-790. doi:10.1038/nn.3708
11. Crist RC, Reiner BC, Berrettini WH. A review of opioid addiction genetics. *Current Opinion in Psychology*. 2019/06/01/ 2019;27:31-35. doi:<https://doi.org/10.1016/j.copsyc.2018.07.014>
12. Johnson EC, Border R, Melroy-Greif WE, de Leeuw CA, Ehringer MA, Keller MC. No evidence that schizophrenia candidate genes are more associated with schizophrenia than noncandidate genes. *Biological Psychiatry*. 2017;82(10):702-708. doi:10.1016/j.biopsych.2017.06.033
13. Border R, Johnson EC, Evans LM, et al. No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. *American Journal of Psychiatry*. 2019/05/01 2019;176(5):376-387. doi:10.1176/appi.ajp.2018.18070881

14. Hatoum AS, Wendt FR, Galimberti M, et al. Ancestry may confound genetic machine learning: Candidate-gene prediction of opioid use disorder as an example. *Drug and Alcohol Dependence*. 2021/12/01/ 2021;229:109115. doi:<https://doi.org/10.1016/j.drugalcdep.2021.109115>
15. FDA approves first test to help identify elevated risk of developing opioid use disorder. 2023. <https://www.fda.gov/medical-devices/medical-devices-news-and-events/fda-approves-first-test-help-identify-elevated-risk-developing-opioid-use-disorder>
16. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. Oct 1 2015;526(7571):68-74. doi:10.1038/nature15393
17. Hunter-Zinck H, Shi Y, Li M, et al. Genotyping array design and data quality control in the Million Veteran Program. *The American Journal of Human Genetics*. 2020/04/02/ 2020;106(4):535-548. doi:<https://doi.org/10.1016/j.ajhg.2020.03.004>
18. Delaneau O, Zagury J-F, Robinson MR, Marchini JL, Dermitzakis ET. Accurate, scalable and integrative haplotype estimation. *Nature Communications*. 2019/11/28 2019;10(1):5436. doi:10.1038/s41467-019-13225-y
19. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nature Genetics*. 2016/10/01 2016;48(10):1284-1287. doi:10.1038/ng.3656
20. Verma A, Huffman JE, Rodriguez A, et al. diversity and scale: genetic architecture of 2,068 traits in the VA Million Veteran Program. *medRxiv*. Jun 29 2023;doi:10.1101/2023.06.28.23291975
21. National Academies of Sciences E, Medicine. *Using population descriptors in genetics and genomics research: a new framework for an evolving field*. The National Academies Press; 2023:240.
22. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7. doi:10.1186/s13742-015-0047-8
23. Kuhn M. Building predictive models in R using the caret package. *Journal of Statistical Software*. 11/10 2008;28(5):1 - 26. doi:10.18637/jss.v028.i05
24. *e1071: Misc Functions of the Department of Statistics, Probability Theory Group*. Version 1.7-14. 2023. <https://cran.r-project.org/web/packages/e1071/index.html>
25. Wang L-J, Zhang CW, Su SC, et al. An ancestry informative marker panel design for individual ancestry estimation of Hispanic population using whole exome sequencing data. *BMC Genomics*. 2019/12/30 2019;20(12):1007. doi:10.1186/s12864-019-6333-6
26. O'Connor LJ. The distribution of common-variant effect sizes. *Nature Genetics*. 2021/08/01 2021;53(8):1243-1249. doi:10.1038/s41588-021-00901-3
27. O'Connor LJ, Schoech AP, Hormozdiari F, Gazal S, Patterson N, Price AL. Extreme polygenicity of complex traits is explained by negative selection. *The American Journal of Human Genetics*. 2019/09/05/ 2019;105(3):456-476. doi:<https://doi.org/10.1016/j.ajhg.2019.07.003>
28. Haslam N, Kvaale EP. Biogenetic explanations of mental disorder: The mixed-blessings model. *Current Directions in Psychological Science*. 2015/10/01 2015;24(5):399-404. doi:10.1177/0963721415588082
29. Lebowitz MS, Ahn W-k. Effects of biological explanations for mental disorders on clinicians' empathy. *Proceedings of the National Academy of Sciences*. 2014/12/16 2014;111(50):17786-17790. doi:10.1073/pnas.1414058111

30. Chung RY-N, Freedman B. Chapter 9 - Health inequalities in AI machine learning. In: Sung JJY, Stewart C, eds. *Artificial Intelligence in Medicine*. Academic Press; 2024:119-130.
31. Vickers-Smith R, Justice AC, Becker WC, et al. Racial and ethnic bias in the diagnosis of alcohol use disorder in veterans. *American Journal of Psychiatry*. 2023/06/01 2023;180(6):426-436. doi:10.1176/appi.ajp.21111097
32. Seal KH, Shi Y, Cohen G, et al. Association of mental health disorders with prescription opioids and high-risk opioid use in US veterans of Iraq and Afghanistan. *JAMA*. 2012;307(9):940-947. doi:10.1001/jama.2012.23433.
33. America is getting older. 2023. <https://www.census.gov/newsroom/press-releases/2023/population-estimates-characteristics.html>
34. Fang H, Hui Q, Lynch J, et al. Harmonizing genetic ancestry and self-identified race/ethnicity in genome-wide association studies. *Am J Hum Genet*. Oct 3 2019;105(4):763-772. doi:10.1016/j.ajhg.2019.08.012
35. Blondal T, Waage BG, Smarason SV, et al. A novel MALDI–TOF based methodology for genotyping single nucleotide polymorphisms. *Nucleic Acids Research*. 2003;31(24):e155-e155. doi:10.1093/nar/gng156
36. Baum ML, Widge AS, Carpenter LL, McDonald WM, Cohen BM, Nemeroff CB. Pharmacogenomic clinical support tools for the treatment of depression. *American Journal of Psychiatry*. 2024:appi.ajp.20230657. doi:10.1176/appi.ajp.20230657