GENETIC CONTRIBUTION TO CANNABIS USE AND OPIOID TREATMENT

GENETIC CONTRIBUTION TO CANNABIS USE AND OPIOID USE DISORDER TREATMENT OUTCOMES

By ALANNAH HILLMER, B.Sc., M.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Doctor of Philosophy

Descriptive Note

DOCTOR OF PHILOSOPHY (2022)
(Neuroscience Graduate Program)

McMaster University
Hamilton, Ontario

TITLE: Genetic contribution to cannabis use and opioid use disorder treatment outcomes

AUTHOR: Alannah Hillmer, H.B.Sc. (McMaster University), M.A. (Yorkville University)

SUPERVISOR: Dr. Zainab Samaan

COMMITTEE MEMBERS: Dr. Lehana Thabane and Dr. Luciano Minuzzi

NUMBER OF PAGES: xiv, 195

Lay Abstract

Cannabis use rates in Canada are increasing, with Opioid Use Disorder (OUD) patients having high rates of cannabis use despite inconsistent findings on the impacts. To combat the opioid crisis, Methadone Maintenance Treatment (MMT) is utilized to reduce opioid cravings and use. However, individuals on MMT are likely to use other substances, including cannabis. This thesis explores the genetic literature on cannabis use and conducts a Genome-Wide Association Study (GWAS) and a Polygenetic Risk Score (PRS). The GWAS investigates genetic variants throughout the whole genome associated with a trait, while the PRS creates a genetic weight risk score. GWAS and PRS methods were used to investigate cannabis use and MMT outcomes within Europeans with OUD. While no significant GWAS results were found, a statistically significant PRS was found for regular cannabis use and methadone dose, suggesting each respective score can estimate an individual's risk of that trait.

Abstract

Background: Canada continues to face an opioid epidemic with 5,368 opioid apparent related deaths occurring between January and September of 2021. Methadone Maintenance Treatment (MMT), a form of Medication Assisted Treatment used to treat Opioid Use Disorder (OUD), has been reported to decrease opioid cravings and opioid use, however, individual differences exist in the effective dose of methadone. Further, individuals living with an OUD have higher rates of substance use including cannabis. A genetic component has been suggested to exist for both cannabis use and MMT outcomes, however inconsistent findings have been reported.

Methods: Knowledge synthesis and primary genetic association studies were conducted. A protocol was prepared for the planning of a systematic review for Genome-Wide Association Studies (GWASs) of cannabis use. The full systematic review was then conducted, providing an assessment of the literature and a description of studies quality. A GWAS and Polygenic Risk Score (PRS) was then conducted for cannabis use and MMT outcomes, separately, in Europeans only. The top Single Nucleotide Polymorphisms (SNPs) were then analyzed separately by sex and sex interactions were conducted.

Results: The systematic review included 6 studies, identifying 96 genetic variants associated with cannabis use. The GWASs for both cannabis use and MMT outcomes did not identify any significant results. A significant PRS was found for regular cannabis use and methadone dose. No sex-specific results were identified.

Discussion: This thesis summarised the evidence on the genetics of cannabis use as well as employed GWASs and PRSs to investigate cannabis use and MMT outcomes within a European population. We were able to highlight gaps within the genetic literature of cannabis and MMT outcomes as well as identify areas of interest for future research.

Acknowledgements

First and foremost, I would like to thank my mentor, teacher, and supervisor Dr. Zena Samaan, without whom this work would not have been possible. Thank you for always challenging me and pushing me to learn, grow and reach a goal I did not know was possible. The time, dedication, and kindness you share with each and every one of your students does not go unnoticed or unappreciated. Thank you for always encouraging me to learn new methods and software's and always making sure I had the resources and support to succeed.

A sincere thank you to my supervisory committee, Drs. Thabane, Kapczinski, and Minuzzi, for their continued support and insight into my thesis projects. I would like to thank each one of you for always being available to meet, providing feedback on my statistical analyses and manuscripts. Thank you for never missing an opportunity to teach me.

To Dr. Andrew Paterson, thank you for all of the expertise, guidance and insight you have provided into genetic statistical analysis. You have helped me progress my knowledge on genetics and genetic analyses to a different level. Your kindness, expertise, quick responses, and continual meetings were a huge support in completing my thesis and expanding my skill set.

Dr. Amel Lamri, thank you for being a mentor and friend throughout my Ph.D. work. You provided me the base knowledge and foundation necessary for me to grow and learn. Thank you for the countless meetings, always being a call away and never failing to be patient with me as I learned and expanded my knowledge of PLINK and R.

To Ms. Caroul Chawar, I am thankful that I not only was able to call you a colleague, but also a friend. Having you with me every step of the way for the first 2 years was not only instrumental in my learning, writing, decision-making and presenting, but you also made sure that every step was fun. I could not have asked for a better team-mate either. Thank you for continuing to be my sounding board, even post-graduation, and always being willing to help when you can.

To the MEND research team and incredible colleagues, it has been wonderful working with all of you. I would like to thank Ms. Jackie Hudson for not only being a mentor and the go-to-person in the lab, but for also being a great friend. Thank you for always supporting me, the endless trips to clinics and always being there for me. To Dr. Nitika Sanger, Balpreet Panesar and Myanca Rodrigues, thank you for making this experience unforgettable and for being the best support team I could ask for. To Alessia D'Elia, thank you for being by my side every step of the way. From starting in the lab at the same time, transferring together and working towards our defence, you have never failed to provide me with support, encouragement, expertise, and advice. Thank you for being an amazing friend and colleague.

I would also like to thank my parents, for their support and unconditional love. Thank you for always encouraging me to explore my limitless potential. Your support and belief that I can accomplish whatever I set my mind to has been instrumental in my success. To my siblings, thank you for always supporting me, encouraging me, and being understanding. To my husband, thank you for being by my side every step of the way and never failing to encourage, support or believe in me even if it meant another year of school.

I would also like to thank all of the past and present members of the GENOA and POST team for their contribution to this research project. The time and commitment put into the design and implementation of this project has been instrumental for my work. Thank you to the Canadian Addition Treatment Centers and staff for their contribution to this project. Finally, I would like to thank the Neuroscience Graduate Program, McMaster University, and Dr. Zainab Samaan for their generous financial support.

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

ADHD: Attention-Deficit Hyperactivity Disorder
Canadian Addition Treatment Centre (CATC)
Cannabidiol (CBD)
CUD: Cannabis Use Disorder
Confidence Interval (CI)
Detla-9-tetrahydrocannabinol (THC)
DSM: Diagnostic and Statistical Manual
DSM-5: Diagnostic and Statistical Manual 5th edition
Genetics of Opioid Addiction (GENOA)
GRADE: Grading of Recommendations Assessment, Development and Evaluation
GWAS: Genome-wide Association Study
GWASs: Genome-wide Association Studies
HuGeNET: The Human Genome Epidemiology Network
ICC: International Cannabis Consortium
ICD-10: International Statistical Classification of Diseases and Related Health Problems - 10
ICGHD: International Consortium on the Genetics of Heroin Dependence
iPSYCH: Integrative Psychiatric Research
Marijuana Craving Questionnaire-Short form (MCQ-SF)
Maudsley Addition Profile (MAP)
Medication-Assisted Treatment (MAT)
Methadone Maintenance Treatment (MMT)
Odds Ratio (OR)
Opioid Use Disorder (OUD)
Pharmacogenetics of Opioid Substitution Treatment Response (POST)
Polygenic Risk Score (PRS)
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
PROSPERO: International Prospective Register of Systematic Reviews
Q-Genie: The quality of genetic association studies
SAGE: Study of Addiction: Genetics and Environment
Single Nucleotide Polymorphism (SNP)
Standard Error (SE)
Strengthening the Reporting of Genetic Association studies (STREGA)
Strengthening the Reporting of Observational studies in Epidemiology (STROBE)
Urine Drug Screens (UDS)

Declaration of Academic Achievement

I, Alannah, am the primary author of all presented studies and manuscripts. I have made significant contributions to the studies by determining the research questions, conducting the statistical analyses, interpreting the results, and writing the manuscripts. Detailed lists of author contributions are provided on the first page of each manuscript, with their contributions detailed at the end of the manuscript.

1 CHAPTER 1: Introduction

1.1 Background and Significance

Trends in Cannabis use have been increasing, where 14% of Canadians 15 years or older reported using cannabis in the first quarter of 2018, compared to 16.7% reporting cannabis use in the last quarter of 2019 (1). Additionally, individuals living with an Opioid Use Disorder (OUD) have higher rates of substance use than the general population with rates of cannabis use reported to be more than 50% (1–4). Concerningly, the short and long-term impacts of cannabis use in OUD are inconclusive, with some studies showing its potential as substitute in keeping with a harm reduction approach, some showing no association, and others identifying worse outcomes for individuals receiving Medication Assisted Treatment (MAT) for an OUD (4–16).

Further, Canada, along with many other nations, continues face an opioid epidemic, with approximately 62 million people reporting opioid use worldwide (17). The likelihood of developing an OUD following opioid use is high, with many individuals who develop OUD experiencing a chronic remitting course of the disorder with a heightened risk of serious adverse outcomes such as increased risk of overdose (18). In 2017, 115,000 people died from an opioid overdose worldwide and in Canada specifically, 5,368 apparently opioid related deaths occurred between January to September of 2021 (17,19). As of 2016 the number of patients enrolled in a MAT program was over 40,000 patients (20). MATs are a critical part of the strategy to address the opioid epidemic and include the controlled administration of opioid agonist or antagonists (20,21). Currently, the World Health Organization recommends both methadone and buprenorphine/naloxone (also known as suboxone) as MATs for the treatment of OUD (17,20).

Methadone Maintenance Treatment (MMT) has been reported to decrease opioid cravings and opioid use, with the treatment target aiming to help individuals control opioid use and regain stability within their life (22–24). While methadone has been shown to be effective, individual level differences exist in the effective dose of methadone; individuals given too low of a dose may experience withdrawal symptoms and individuals given too high of a dose may experience drowsiness, confusion and mental impairment (25). Due to the chronic relapsing nature of OUD, inappropriate dosing of methadone may result in relapse or increased risk of overdose due to the direct or interacting effects with other illicitly used opioids (26–28).

A genetic component has been suggested in both response to MMT and cannabis use. Genes of interested in cannabis use outcomes (including cannabis dependence and Cannabis Use Disorder (CUD)), genes of interest include the cannabinoid receptors CB1 (*CNR1)* and CB2 *(CNR2)*, *FAAH*, *DRD2*, *ANKK1*, *ABCB1*, *CSMD1*, *ACSS1*, *SCN9A, CADM2*, and *FOXP2* however, replication of these associations has been inconsistent (29–35). For MMT outcomes (including but not limited opioid addiction, treatment response, methadone dose) have been identified as *OPRM1*, *OPRD1*, *ABCB1*, and

CYP2B6, however many of the studies investigating MMT outcomes are candidate gene studies with small sample sizes and show inconsistent findings (36–38).

This thesis aims to address the need for a comprehensive and current literature search for genome wide significant results investigating the genetics of cannabis and conduct a Genome-Wide Association Study (GWAS) and Polygenic Risk Score (PRS) to inform clinical applications and the direction of future research for both cannabis use and MMT outcomes.

1.2 Objectives

The specific objectives of this thesis addressed within the four included manuscripts are the following:

1. To systematically and methodologically search the literature on genetic study findings regarding cannabis use
2. To summarize these findings and assess the quality of the published literature
3. To examine if novel genetic variants are associated with cannabis use within the OUD population and create a PRS based on a previously reported GWAS on cannabis use
4. To examine if novel genetic variants are associated with MMT outcomes and create a PRS based on a previously reported GWAS on methadone dose

1.3 Coherence of Thesis Chapters

This thesis is comprised of four manuscripts, focused on genetics within the OUD population. Chapter 2 and 3 provided a background on the current literature of the genetics of cannabis use while Chapter 4 and 5 utilize genetic statistical analyses to test associations within the OUD population. More specifically, Chapter 2 of this thesis is a protocol which outlines a detailed study design and search strategy for the systematic review to follow, ensuring methodological transparency and a peer-review process prior to conducting the systematic review. This protocol is published in *Systematic Reviews*. Chapter 3 applies the search strategy outlined in the published protocol and summaries relevant genome-wide association results to highlight potential SNPs of interest and identify gaps within the literature. This systematic review is published in *BMC Medical Genomics.* Chapter 4 employs a GWAS and PRS to identify novel genetic variants associated with cannabis use and quantifying the variance explained by genetic variability. In addition, Chapter 4 explores genetic differences by sex. Finally, Chapter 5 applies the methods utilized in Chapter 4, including the GWAS, PRS and sex analyses, to investigate differences in MMT outcomes.

Due to the overlap between each manuscript, including but not limited to the population of interest, methodology and the genetic nature of the studies, the chapter-specific backgrounds, rationale, and methods might contain overlapping information and concepts. Despite these similarities, each manuscript discussed in this thesis is unique and serves a specific purpose.

2 CHAPTER 2

2.1 COPYRIGHT STATEMENT

Hillmer, A., Chawar, C., Sanger, S., D'Elia, A., Butt, M., Kapoor, R., Kapczinski, F., Pare, G., Thabane, L., & Samaan, Z. (2020). Genetic determinants of cannabis use: a systematic review protocol. *Systematic reviews*, 9(1), 1-6.

2.2 Genetic determinants of cannabis use: a systematic review protocol.

Authors & Affiliations
Alannah Hillmer[1]; Caroul Chawar[2]; Stephanie Sanger[3]; Alessia D'Elia[4], Mehreen Butt[5]; Raveena Kapoor[6]; Flavio Kapczinski[7], Guillaume Pare[8], Lehana Thabane[9]; *Zainab Samaan[10]

1. Neuroscience Graduate Program; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7 hillmea@mcmaster.ca
2. Neuroscience Graduate Program; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; chawarc@mcmaster.ca
3. Health Science Library; McMaster University; 1280 Main St. W., Hamilton, ON, L8S 4L8; sangers@mcmaster.ca
4. Neuroscience Graduate Program; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; deliaa@mcmaster.ca
5. Integrated Science Program; McMaster University; 1280 Main St. W., Hamilton, ON, L8S 4L8; buttm6@mcmaster.ca
6. Health Sciences Program; McMaster University; 1280 Main St. W., Hamilton, ON, L8S 4L8; kapoor8@mcmaster.ca
7. Department of Psychiatry and Behavioural Neuroscience; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; kapczinf@mcmaster.ca
8. Population Health Research Institute; McMaster University; 1280 Main St. W., Hamilton, ON, L8S 4L8; pareg@mcmaster.ca
9. Department of Health Research Method, Evidence & Impact; 1280 Main St. W., Hamilton, ON, L8S 4L8; thabanl@mcmaster.ca
10. Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; samaanz@mcmaster.ca

*Corresponding Author:
Dr. Zainab Samaan
100 West 5th St., Hamilton, ON L8N3K7, Canada
Telephone: 905-522-1155 ext. 35448
Fax: 905-381-5629

2.2.1 Abstract

Background: With the legalization of Cannabis in Canada there is an increase trend in use. Cannabis has been known to have several health implications, one of which is the development of Cannabis Use Disorder (CUD). CUD more common in males than females, as well as in certain ethnic groups such as Native Americans. Additionally, both environmental and genetic risk factors have been found for cannabis use. The objective of this systematic review will be to summarize the genetic variants associated with cannabis use which have reached borderline genome-wide significance.

Methods: This systematic review will incorporate articles that have performed a Genome-Wide Association Study (GWAS) investigating cannabis use. MEDLINE, Web of Science, EMBASE, GWAS Catalog, GWAS Central, and NIH Database of Genotype and Phenotype will be searched using a comprehensive search strategy. The Quality of Genetic Association Studies (Q-Genie) tool will be utilized to assess the quality of the included studies. All screening and data extraction will occur independently by two authors. If feasible, a random-effects meta-analysis will be conducted on pooled odds ratios of single nucleotide polymorphisms reaching genome-wide significance.

Discussion: This systematic review will synthesise available GWAS on cannabis use. Results from this review will inform and direct further investigation of genetic variants associated with cannabis use.

Systematic Review Registration PROSPERO CRD42020176016.

Keywords: Systematic Review, Cannabis, Genetics, Genome-wide

2.2.2 Background

On October 17th, 2018 the Cannabis Act came into effect in Canada allowing for the legal growth of cannabis plants as well as the recreational possession and consumption of cannabis for those who are 18 years or older(1). In response to the Cannabis Act, Statistics Canada has introduced a National Cannabis Survey which has been conducted every three months since February 2018. The NCS showed that nearly 17 percent of Canadians aged 15 years and older reported using cannabis within a 3-month period between mid-August to mid-September of 2019, a rate that was consistent with the rate of the year prior, when cannabis was an illicit substance. However, in the fourth quarter of 2019 cannabis use was increased when compared to the fourth quarter of 2018. Additionally, regardless of the year of study, cannabis consumption rates continue to be higher among males than females(2).

Cannabis Use Disorder (CUD) is defined as a problematic pattern of cannabis use leading to clinically significant impairment or distress. In 2013 the Diagnostic and statistical manual reported that CUD is prevalent in 3.4% of youth aged 12- to 17- years old and 1.5% of adults age 18 years or older. Trends of CUD also differ among sex and ethnicities. Rates of CUD is higher in males compared to females and rates of CUD is higher in Native American and Alaska Natives compared to other ethnic groups(3). Results from a meta-analysis on twin studies estimated the heritability for cannabis use initiation to be 40-48% and 51-59% for problematic cannabis use, suggesting a genetic component to cannabis use and CUD(4). A Genome-Wide Association Study (GWAS) combined five cohorts identifying several genes and Single Nucleotide Polymorphisms (SNPs) associated with cannabis use and dependence(5). A cluster of correlated SNPs in a novel region of chromosome 10 were identified at genome-wide significant levels in participants of European descent(5). However, of three meta-analyses conducted on cannabis use in the literature, only one study identified a significate association(6–8). One region on chromosome 16 was significantly associated with age of first cannabis use, with the strongest association for the intronic variant rs1574587(7).

Interestingly, one study investigated the genetic and environmental risk factors for cannabis availability reported variation in cannabis initiation and symptoms of cannabis use disorder. Cannabis availability and initiation had a correlation of 0.48 and cannabis availability and symptoms of cannabis use disorder had a correlation of 0.23. Additionally, much of the variation associated with problematic use can be explained by shared environmental risk in cannabis availability leading to initiation and the genetic non-shared environmental risks for cannabis initiation(9). These findings are of specific interest to Canada and other countries with legalization of cannabis is already in effect or being considered, as cannabis is increasingly more available since the legalization.

With cannabis availability increasing, and known heritability of CUD, it is important to understand the genetic risk factors associated with cannabis use. While meta-analyses of GWASs provide regions of interest, no known systematic review exists that summarises identified genes and/or SNPs that have reached genome-wide significance for cannabis use. It is important to provide a summary of the literature which includes recent GWASs in the context of cannabis legalization. Further, understanding the genetic basis

of cannabis use will assist health care workers in making science-informed decisions regarding recreational and prescription cannabis use.

2.2.2.1 Objectives

The main goal of this systematic review is to identify genetic variants from Genome-Wide Association Studies (GWASs) associated with cannabis use. Though genetic variants most commonly reported by GWASs are SNPs, this review will be inclusive of any other genetic markers reported in GWASs. We will summarize the results of GWASs which meet our inclusion criteria, and if possible, we will meta-analyze genetic variants that are reported in more than one primary study.

Primary objectives of this systematic review include:

1. Identify genetic variants associated with current cannabis use. Current cannabis use is defined by either self-report or positive urine drug screens within one month of the study being conducted.
2. Identify genetic variants associated with lifetime cannabis use. Lifetime cannabis use is defined by any self-reported or positive urine drug screens of cannabis use within one's lifetime.
3. Identify genetic variants associated with CUD. CUD is defined by any diagnostic and classification systems used to diagnosis CUD or questionnaires validated to assess CUD.

Secondary objectives of this systematic review include:

1. Identify genetic variants associated with the adverse outcomes of cannabis use including psychiatric (cognitive impairment, psychotic symptoms, depression, anxiety, suicidal behaviour), and non-psychiatric (chronic bronchitis, lung infections, chronic cough, increased risk of motor vehicle accidents), and any other reported adverse outcomes(10–12).
2. When feasible, perform subgroup summaries by sex or ethnic differences.

2.2.3 Methods

This protocol is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) statement(13) (see PRISMA-P checklist in Additional file 1). This protocol was registered within the International Prospective Register of Systematic Reviews (PROSPERO) (registration number: CRD42020176016).

2.2.3.1 Eligibility Criteria

GWAS studies presenting original data on associations between cannabis use and genetic polymorphisms using any study design (i.e. case-control, cohort, etc.) will be included in this systematic review. All other types of studies will be excluded. Studies in

any setting will be included and no restriction will be placed on age, sex, ethnic background or language. Additionally, articles that do not present sufficient data to calculate the odds ratio (OR) with a 95% confidence interval will be excluded from quantitative analyses if data cannot be obtained after contacting the studies' authors and the calculations cannot be made with the available published information. However, we will include these studies in the qualitative description of the review findings.

We will include studies investigating Cannabis Use Disorder as defined by the Diagnostic and Statistical Manual -5 (DSM-5), or other diagnostic and classification systems such as the International Statistical Classification of Diseases and Related Health Problems -10 (ICD-10) or specific diagnostic scales designed to screen and diagnose dependence or use disorder of cannabis will be included as well as any studies measuring any use of cannabis. We define cannabis use based on the included studies' definitions and accept the following definition: Current cannabis use is defined as either self-report or positive urine drug screens within one month of the study being conducted; and lifetime cannabis use is defined as any self-reported or positive urine drug screens of cannabis use within one's lifetime (14). Clinical diagnoses and questionnaires validated to assess CUD will also be accepted. All studies not investigating current cannabis use, lifetime cannabis use or CUD will be excluded. In the case of polymorphisms reported in duplicate publications from the same study population, the article most recent will be included.

2.2.3.2 Information Sources and Search Strategy

A Health Science Librarian was consulted to develop a comprehensive search strategy. No language restriction will be placed on the search strategy, though studies will be limited to human studies. MEDLINE, Web of Science, EMBASE, GWAS Catalog, GWAS Central, and NIH Database of Genotype and Phenotype will be searched using the agreed-upon strategy, modified for each database. The search strategy will include all terms relevant to cannabis and genome-wide association studies. Databases will be searched from inception onwards. Sources of grey literature including dissertations and theses, clinical guidelines and reports from regulatory agencies will be searched. Reference lists of relevant systematic reviews and all included studies will be checked to identify additional articles.

Search strategy:

Draft search strategies for multiple electronic databases are provided in Additional file 2.

2.2.3.3 Study Records

2.2.3.3.1 Data management

All of the references will be managed and organized through Zotero(15). Covidence will be used for the management of this systematic review at the title and abstract, full text and data extraction stages(16). Prior to the formal screening process, a

calibration will take place to pilot and refine the screening process. Training will be given to all team members on using Covidence prior to starting the review.

2.2.3.3.2 Selection Process

Two independent reviewers will screen titles and abstracts for inclusion criteria. Full-text review will also be completed independently by two reviewers. Disagreements between reviewers will be resolved by consensus or including a third reviewer. We will record the reason for excluding studies at the full-text review stage.

2.2.3.3.3 Data Collection Process

Data extraction will take place independently and in duplicate for each eligible study. Standardized full-text data extraction forms will be constructed. The data extraction form will be pilot tested by two independent reviewers to determine the feasibility of this review and ensure all details are captured. In the event of missing data, we will contact study authors to obtain missing information where possible. All contact with the authors will be documented.

2.2.3.3.4 Data Items

We will extract the following information: author, year of study, country, cohort population used, number of participants (separated by those included in the cannabis use group and non-cannabis use group), control population, the ethnicity of participants, mean age, sex ratio, the measure of Cannabis Use Disorder or cannabis use or definition of cannabis use, inclusion and exclusion criteria, how cannabis use was reported (i.e. self-report, drug urine screens), frequency of cannabis use and finally any genetic variants which reached the significance threshold set of $p \leq 10^{-7}$. Genome-wide significance is generally considered any SNP with a p-value less than 5×10^{-8}, however, SNPs reaching borderline significance, $p < 10^{-7}$, will also be extracted as borderline significance has been found to be generally replicable(17).

2.2.3.4 Outcomes and Prioritization

The main aim of the systematic review will be to assess variants reaching the given threshold associated with cannabis use outcomes from the primary studies included in this review.

The primary outcomes are as follows:

1. Current cannabis use is defined as either self-reported cannabis use or positive cannabis urine drug screens within one-month of the study being conducted.
2. Lifetime cannabis use, defined as self-reported ever used cannabis during the individual's lifetime.
3. CUD, defined by a diagnosis from the DSM-5 or other diagnostic and classification system such as the ICD-10 or specific diagnostic scales designed to screen and diagnose dependence or use disorder of cannabis.

For each of the outcomes above we will collect information on each outcome as reported in the primary studies meeting the eligibility criteria, including dichotomous use of cannabis, percent positive urine screens, questionnaires, diagnostic classification and any other form of data collection.
The secondary outcomes are as follows:

1. Adverse outcomes of cannabis use including psychiatric and non-psychiatric outcomes and any other reported adverse outcomes. We will collect data as reported in the primary studies included such as comorbid diagnosis, additional medication condition or other adverse outcomes reported in the studies.
2. We will collection information from the included primary studies on sex and ethnic groups within the study. We will provide a qualitative summary and, if feasible, conduct a subgroup meta-analysis of genetic variants within specific ethnic groups.

2.2.3.5 Risk of Bias in Individual Studies

Quality assessment will be completed in duplicate for each study included. The Quality of Genetic Association Studies (Q-Genie) tool [Version 1.1] will be used. Disagreements of quality assessments will be resolved through discussion(18). If a consensus is not reached through discussion, a third author will be consulted to resolve the disagreement.

2.2.3.6 Data Synthesis

Studies included in this systematic review will undergo qualitative synthesis. Summary tables will be used which will include the sample size, size of cannabis group and non-cannabis group, sex distribution, mean age, study design, ethnic population and outcome (current cannabis use, lifetime cannabis use or CUD). A separate table will be used to display any variants reaching borderline genome-wide significance, the corresponding study it was reported in, the corresponding chromosome and position, minor allele, gene/locus, population size, outcome associated, measure, measure of association value, measure of variability, ethnicity, and p-value reported.

Heterogeneity between the studies will be assessed through the I^2 statistic with a 95% confidence interval. We will also report summary tables including the study design, population, and cannabis use measure/definition to describe heterogeneity qualitatively. If appropriate, a random-effects meta-analysis will be conducted on pooled odds ratios for the main outcome previously mentioned. If appropriate, the a random-effects meta-analysis will be conducted on pooled odds ratio for the secondary outcomes previously mentioned as well as a subgroup analyses of the participants sex and ethnicities. Subgroup analyses by participant sex accounts for any differences in cannabis use between sex's which has been previously reported in the literature (19–21). Additionally, due to genetic differences between ethnicities, genetic associations may be more predominant in certain ethnic groups than others, as such a subgroup analysis will be conducted, if feasible (22). Studies excluded from the quantitative analysis will be listed and an exclusion reason will be given.

If quantitative methods of analysis are not feasible for both the primary or secondary outcomes due to either low heterogeneity found by the I^2 statistic or qualitative synthesis or no two study reports the same genetic variant, only qualitative synthesis results will be reported. We will not conduct a meta-analysis of individual participant data.

2.2.3.7 Meta-bias

To help mitigate publication bias conference abstracts will included, manual searches of references lists will be conducted and Cochrane Clinical Trail Protocols Registry and ClinicalTrails.gov databases will be searched for relevant clinical trial protocols. Additionally, the GWAS catalog will be manual searched for borderline significant variants associated with current cannabis use, lifetime cannabis use, or CUD to ensure all variants are captured within this review. Authors of conference abstracts will be contacted to determine the stage of the research project and all correspondence will be documented. If the published work was not captured by the search strategy, and deemed eligible by two independent reviewers, it will be included. Two independent reviewers will search the references lists of all included studies. Any identified references, deemed eligible by two independent reviewers, will be included.

2.2.3.8 Confidence in Cumulative Evidence

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) will be used to assess the strength of evidence. GRADE scores according to the risk of bias, publication bias, consistency, directness and precision. A score of high-, moderate, low-, or very low-quality evidence will be assigned and summarized in a table(23).

2.2.3.9 Presenting and Reporting of Results

The full review will follow the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines with special consideration to the Human Genome Epidemiology Network (HuGENet) guidelines(24). Although HuGENet reviews typically focus on a single gene, we will present information on each genetic variant-phenotype association reported which will include the study details, population, findings and source of data.

2.2.4 Discussion

A lack of consistent evidence exists in the current literature for genetic variants associated with cannabis use. In addition, this is the first known systematic review to synthesize the available evidence on genetic variants associated with cannabis use. The proposed systematic review aims to identify all genetic variants that have reached borderline genome-wide significance associated with cannabis use and CUD. The proposed systematic review will provide an overview of the current literature on the

genetics of cannabis, aiding in the genetic understanding of cannabis use. Understanding the genetic contribution to cannabis use and its effects such as cannabis use disorder has the potential to aid medical practitioners in making decisions related to cannabis use for medical reasons and the associated potential risks. Additionally, variants reaching borderline genome-wide significance will be examined in the context of their known or biologically plausible relevance to further our understanding.

Anticipated limitations of this review existed at both the study and review level. Limitations at the study level that may include a lack of reporting quality control steps, reporting of variants within linkage disequilibrium, small sample size, and a lack of reporting variants that failed to reach genome-wide significance ($p<5x10^{-8}$) but may have reached borderline significance levels ($p<10^{-7}$). At the review level, limitations exist in the expected high heterogeneity, differing outcomes for cannabis use reported in the literature and the exclusion of meta-analysis and candidate gene studies.

On completion of the systematic review, we will publish in a peer-review academic journal to reach both clinical and academic experts in the field. This systematic review will then inform and direct the further investigation of genetic variants associated with cannabis through candidate gene studies.

2.2.5 List of Abbreviations

CUD: Cannabis Use Disorder
DSM-5: Diagnostic and Statistical Manual 5th edition
GRADE: Grading of Recommendations Assessment, Development and Evaluation
GWAS: Genome-wide Association Study
GWASs: Genome-wide Association Studies
HuGeNET: The Human Genome Epidemiology Network
ICD-10: International Statistical Classification of Diseases and Related Health Problems - 10
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
Q-Genie: The quality of genetic association studies

2.2.6 Declarations

2.2.6.1 Ethics Approval and Consent to Participate

Not applicable.

2.2.6.2 Consent for Publication

Not applicable.

2.2.6.3 Availability of Data and Materials

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

2.2.6.4 Competing Interests

The authors declare that they have no competing interests.

2.2.6.5 Funding

This is study is supported in part by CIHR (grant number PJT-156306). The funding source has no role in the study design, analysis, reporting or the decision to publish the study.

2.2.6.6 Authors' Contributions

ZS is the guarantor. AH and CC drafted the manuscript. AH, CC and ZS contributed to the development of the selection criteria, the risk of bias assessment strategy and data extraction criteria. SS provided expertise in developing the search strategy. All authors read, provided feedback and approved the final manuscript.

2.2.6.7 Amendments

If amendments to this protocol are made, they will be documented and communicated to the journal. A data of amendment, description, and rationale will accompany each amendment.

2.2.6.8 Additional Files

Additional File 1 is provided in .pdf format, is titled “PRISMA-P 2015 Checklist” and contains PRISMA-P checklist.

2.2.7 References

1. The Government of Canada. Cannabis Legalization and Regulation. 2018.
2. Statistics Canada. National Cannabis Survey, third quarter 2019. 2019.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Washington, DC; 2013.
4. Verweij KJH, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG, et al. Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. Addiction. 2010;105(3):417–30.
5. Agrawal A, Chou YL, Carey CE, Baranger DAA, Zhang B, Sherva R, et al. Genome-wide association study identifies a novel locus for cannabis dependence. Mol Psychiatry. 2018;23(5):1293–302.
6. Verweij K, Vinkhuyzen A, Benyamin B, Lynskey M, Quaye L, Agrawal A, et al. The genetic etiology of cannabis use initiation: a meta-analysis of genome-wide association studies and a SNP-based heritability estimation. Addict Biol. 2013;18(5):846–50.
7. Minică CC, Verweij KJH, Most PJ, Mbarek H, Bernard M, Eijk KR, et al. Genome-wide association meta-analysis of age at first cannabis use. Addiction [Internet]. 2018 Nov;113(11):2073–86. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=132211589&site=ehost-live
8. Stringer S, Minică CC, Verweij KJH, Mbarek H, Bernard M, Derringer J, et al. Genome-wide association study of lifetime cannabis use based on a large meta-analytic sample of 32 330 subjects from the International Cannabis Consortium. Transl Psychiatry. 2016;6(3):e769.
9. Gillespie NA, Neale MC, Kendler KS. Pathways to cannabis abuse: a multi-stage model from cannabis availability, cannabis initiation and progression to abuse. Addiction. 2009;104(3):430–8.
10. The Government of Canada. Cannabis Health Effects. 2019.
11. Hall W. What has research over the past two decades revealed about the adverse health effects of recreational cannabis use? Addiction. 2015;110(1):19–35.
12. Gobbi G, Atkin T, Zytynski T, Wang S, Askari S, Boruff J, et al. Association of cannabis use in adolescence and risk of depression, anxiety, and suicidality in young adulthood: a systematic review and meta-analysis. JAMA psychiatry. 2019;76(4):426–34.
13. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. Bmj. 2015;349.
14. National Academies of Sciences Engineering and Medicine. The health effects of cannabis and cannabinoids: The current state of evidence and recommendations for research. National Academies Press; 2017.
15. Center for History New Media. Zotero: The Next-Generation Research Tool [Internet]. George Mason University; 2009. Available from: http://www.zotero.org/

16. Veritas Health Innovation. Covidence systematic review software [Internet]. Melbourne, Australia; Available from: https://www.covidence.org/
17. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. Genet Epidemiol Off Publ Int Genet Epidemiol Soc. 2008;32(3):227–34.
18. Sohani ZN, Meyre D, de Souza RJ, Joseph PG, Gandhi M, Dennis BB, et al. Assessing the quality of published genetic association studies in meta-analyses: The quality of genetic studies (Q-Genie) tool. BMC Genet. 2015;16.
19. Cooper ZD, Haney M. Investigation of sex-dependent effects of cannabis in daily cannabis smokers. Drug Alcohol Depend. 2014;136:85–91.
20. Lev-Ran S, Imtiaz S, Taylor BJ, Shield KD, Rehm J, Le Foll B. Gender differences in health-related quality of life among cannabis users: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Drug Alcohol Depend. 2012;123(1–3):190–200.
21. Van Gastel WA, MacCabe JH, Schubart CD, Van Otterdijk E, Kahn RS, Boks MPM. Cannabis use is a better indicator of poor mental health in women than in men: a cross-sectional study in young adults from the general population. Community Ment Health J. 2014;50(7):823–30.
22. Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, et al. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. Int J Methods Psychiatr Res. 2018;27(2):e1608.
23. Guyatt GH, Oxman AD, Schünemann H, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol. 2011;64(4):380–2.
24. Little J, Higgins JPT, Bray M, Ioannidis J, Khoury M, Manolio T, et al. The HuGENet^{TM} HuGE review handbook, version 1.0. Ottawa, Ontario, Canada HuGENet Canada Coord Cent. 2006;

3 CHAPTER 3

3.1 COPYRIGHT STATEMENT

Hillmer, A., Chawar, C., Sanger, S., D'Elia, A., Butt, M., Kapoor, R., Kapczinski, F., Thabane, L., & Samaan, Z. (2021). Genetic basis of cannabis use: a systematic review. *BMC medical genomics*, 14(1), 1-21.

3.2 Genetic Basis of Cannabis Use: A Systematic Review

Authors & Affiliations
Alannah Hillmer[1]; Caroul Chawar[2]; Stephanie Sanger[3]; Alessia D'Elia[4], Mehreen Butt[5]; Raveena Kapoor[6]; Flavio Kapczinski[7], Lehana Thabane[8]; *Zainab Samaan[9]

1. Neuroscience Graduate Program; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7 hillmea@mcmaster.ca
2. Neuroscience Graduate Program; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; chawarc@mcmaster.ca
3. Health Science Library; McMaster University; 1280 Main St. W., Hamilton, ON, L8S 4L8; sangers@mcmaster.ca
4. Neuroscience Graduate Program; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; deliaa@mcmaster.ca
5. Integrated Science Program; McMaster University; 1280 Main St. W., Hamilton, ON, L8S 4L8; mehreen.saleem.butt@gmail.com
6. Michael G. DeGroote School of Medicine; McMaster University; 1280 Main St. W., Hamilton, ON, L8S 4L8; raveena.kapoor@medportal.ca
7. Department of Psychiatry and Behavioural Neuroscience; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; kapczinf@mcmaster.ca
8. Department of Health Research Method, Evidence & Impact; 1280 Main St. W., Hamilton, ON, L8S 4L8; thabanl@mcmaster.ca
9. Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; samaanz@mcmaster.ca

*Corresponding author:
Dr. Zainab Samaan
100 West 5th St., Hamilton, ON L8N 3K7, Canada
Telephone: 905-522-1155 ext. 35448
Fax: 905-381-5629

1.2.1 Abstract

Background: With the increase in cannabis use rates, cannabis use disorder is being reported as one of the most common drug use disorders globally. Cannabis use has several known physical, psychological, and social adverse events, such as altered judgement, poor educational outcomes, and respiratory symptoms. The propensity for taking cannabis and the development of a cannabis use disorder may be genetically influenced for some individuals. Heritability estimates suggest a genetic basis for cannabis use, and several genome-wide association studies (GWASs) have identified possible regions of association, albeit with inconsistent findings. This systematic review aims to summarize the findings from GWASs investigating cannabis use and cannabis use disorder.

Methods: This systematic review incorporates articles that have performed a GWAS investigating cannabis use or cannabis use disorder. MEDLINE, Web of Science, EMBASE, CINAHL, GWAS Catalog, GWAS Central, and NIH Database of Genotype and Phenotype were searched using a comprehensive search strategy. All studies were screened in duplicate, and the quality of evidence was assessed using the quality of genetic association studies (Q-Genie) tool. All studies underwent qualitative synthesis; however, quantitative analysis was not feasible.

Results: Our search identified 5984 articles. Six studies met our eligibility criteria and were included in this review. All six studies reported results that met our significance threshold of $p \leq 1.0 \times 10^{-7}$. In total 96 genetic variants were identified. While meta-analysis was not possible, this review identified the following genes, *ANKFN1*, *INTS7*, *PI4K2B*, *CSMD1*, *CST7*, *ACSS1*, and *SCN9A*, to be associated with cannabis use. These regions were previously reported in different mental health conditions, however not in relation to cannabis use.

Conclusion: This systematic review summarized GWAS findings within the field of cannabis research. While a meta-analysis was not possible, the summary of findings serves to inform future candidate gene studies and replication efforts.

Systematic Review Registration PROSPERO CRD42020176016.

Keywords: Systematic Review, Cannabis, Genetics, Genome-wide Association Study

3.2.1 Introduction

3.2.1.1 Rationale

Over the past two decades cannabis use and dependence are estimated to have increased, with cannabis use disorder (CUD) reported as one of the most common drug use disorders globally(1). In Canada, it has been reported that nearly 17 percent of Canadians aged 15 years and older reported using cannabis between October and December of 2019, an increase from 14 percent between January to March of 2018. Additionally, cannabis consumption rates are higher among males than females(2). Concerningly, cannabis has been associated with substantial adverse effects. Like other drugs, cannabis can result in cravings, dependence, and drug-seeking behaviour(3,4). During intoxication, cannabis can interfere with memory, motor coordination, altered judgement, and at higher doses, paranoia or psychosis(3). Further, repeated use of cannabis can have long lasting effects, including altered brain development, poor education outcome, cognitive impairment, diminished life satisfaction and achievement, poor professional and social achievements, symptoms of chronic bronchitis and increased risk of chronic psychotic disorders(3,5).

Heritability estimates for cannabis use initiation varied from 30-48%, and from 51-59% for problematic cannabis use, suggesting a genetic component exists(6). Genome-wide association study (GWAS) meta-analyses have identified possible regions of association on chromosome 3 for lifetime cannabis use (*CADM2*), chromosome 10 for CUD (rs77300175), and chromosome 16 for age of first cannabis use (*ATP2C2*)(7–9). Moreover, candidate gene studies have detected some significant associations with cannabis use on the *CNR1*, *GABRA2*, *FAAH*, and *ABCB1* genes, but as with genome-wide association studies (GWASs), replication of these associations has been inconsistent(10).

GWASs provide a 'hypothesis-free' method of identifying novel variant-trait associations, leading to the discovery of novel biological mechanisms and diverse clinical applications(11). As such, in this systematic review, we will summarize GWAS findings relevant to cannabis use or CUD outcomes and discuss future directions.

3.2.1.2 Objectives

The main goal of this systematic review is to identify genetic variants from GWASs associated with cannabis use.
Primary objectives of this systematic review include the following:

1. Identify genetic variants associated with current cannabis use. Current cannabis use is defined by either self-report or positive urine drug screens within 1 month of the study being conducted.
2. Identify genetic variants associated with lifetime cannabis use. Lifetime cannabis use is defined by any self-reported or positive urine drug screens of cannabis use within one's lifetime.
3. Identify genetic variants associated with CUD. CUD is defined by any diagnostic and classification systems used to diagnose CUD or questionnaires validated to assess CUD.

Secondary objectives of this systematic review include the following:

1. Identify genetic variants associated with the adverse outcomes of cannabis use, including psychiatric (cognitive impairment, psychotic symptoms, depression, anxiety, suicidal behavior) and non-psychiatric (chronic bronchitis, lung infections, chronic cough, increased risk of motor vehicle accidents)(12–14).
2. When feasible, perform subgroup summaries by sex or ethnic differences.

3.2.2 Methods

This systematic review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement(15) (see PRISMA checklist in Additional file 1). The Human Genome Epidemiology Network (HuGENet) guideline was used to supplement the PRISMA guideline. While this review does not conform with the HuGENet guideline expectations of reporting on candidate gene study findings, the HuGENet is used to uphold the standard of reporting research specific to genetic association studies(16).

3.2.2.1 Protocol and registration

The protocol for this systematic review has been registered within the International Prospective Register of Systematic Reviews (PROSPERO) (registration number: CRD42020176016)(17). The full protocol has been published in the journal of Systematic Reviews(18).

3.2.2.2 Eligibility criteria

This review investigates GWASs presenting original data on associations between cannabis use and genetic polymorphisms using any study design (i.e. case-control, cohort, etc.). We include studies investigating CUD as well as any studies measuring any use of cannabis. Studies that investigated CUD as defined by any version of the Diagnostic and Statistical Manual (DSM) or other diagnostic and classification systems such as the International Statical Classification of Diseases and Related Health Problems-10 (ICD-10) were included. We define cannabis use based on the included studies' definitions and accept the following definitions: current cannabis use is defined as either self-report or positive urine drug screens within one month of the study being conducted, and lifetime cannabis use is defined as any self-reported or positive urine drug screens of cannabis use within one's lifetime(19). All other studies that did not perform a GWAS and investigate cannabis use or CUD were excluded. No restrictions were placed on the study setting or participant's age, sex, ethnic background or language. Further details on the inclusion criteria can be found in the study protocol(18).

3.2.2.3 Information sources and search strategy

A Health Science Librarian was consulted to develop a comprehensive search strategy. OVID MEDLINE 1946-Present, Web of Science 1976-Present, OVID EMBASE 1974-Present, EBSCOHost CINHAL 1981-Present, GWAS Catalog, GWAS Central, and NIH Database of Genotype and Phenotype databases were searched using the established strategy, modified for each database. All databases were searched from inception to February 2nd, 2021. The search strategy included all terms relevant to genome-wide association studies and cannabis. The search strategies for each electronic database are provided in Table 1.

3.2.2.4 Study selection and data collection process

Calibration was completed prior to the formal screening process. Title and abstract screening, full-text screening and data extraction phases were completed in duplicate through Covidence(20). Conflict resolution at the title and abstract and full-text stages was performed by a senior reviewer (AH or CC), blind to the reviewer's vote. Disagreements at the data extraction stage was resolved by the consensuses of the two reviewers. The reason for study exclusion was recorded at the full-text stage.

3.2.2.5 Data items

Data extracted included baseline participant characteristics, the measure of cannabis used, relevant and significant measured outcomes, statistical measures, and reported study limitations and conflicts. For this review, the threshold of significance of genetic variants reaching $p \leq 10\text{-}7$ was set, as some GWAS results with this significance level have been shown to be replicable within the literature(21).

3.2.2.6 Risk of bias within studies and data analysis

Quality assessment was completed in duplicate for each included study using the Quality of Genetic Association Studies (Q-Genie) tool [Version 1.1](22). Disagreements of quality assessment was resolved through discussion between the two reviewers, and the first author reviewed and confirmed all quality assessments.

3.2.2.7 Summary measures and synthesis of results

A random-effects meta-analysis through pooled odds ratios was planned to quantitatively assess the data. However, these measures were not appropriate as data extracted from each study were unique and could not be combined. For the aforementioned reasons, a heterogeneity test, and a subgroup meta-analyses could not be completed.

3.2.2.8 Risk of bias across studies

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) was used to assess the strength of evidence, with specific consideration of

prognostic factors(23,24). GRADE scores assess outcomes according to the risk of bias, publication bias, consistency, directness, and precision(23).

3.2.3 Results

3.2.3.1 Study selection

The search strategy, along with hand-searching, yielded 5984 studies. After removing duplicates through the Zotero reference manger and Covidence, 4344 studies were unique and screened for eligibility at the title and abstract phase(20,25). Of the 69 studies eligible for full-text screening, 6 studies were included in this review and underwent data extraction and quality assessment.

Studies frequently failed to meet the eligibility criteria for inclusion for the following reasons (i) conducted a GWAS meta-analysis, (ii) conducted a candidate gene study or (iii) were investigating a factor associated with cannabis use (i.e. aggression) rather than cannabis use itself.

Please see the PRISMA flow diagram in Figure 1.

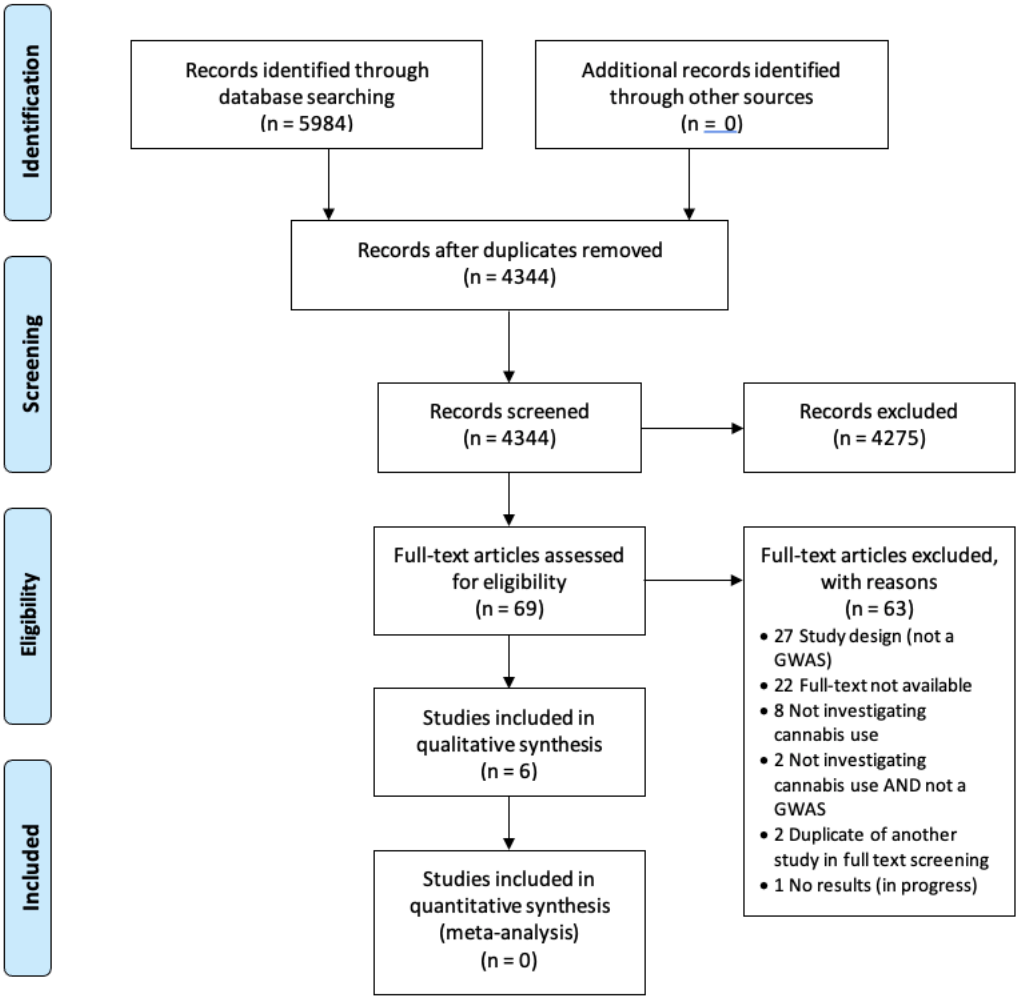


Figure 1. PRISMA Flow Diagram of Study Inclusion.

3.2.3.2 Study characteristics

Individual study characteristics are reported in Table 2. Two studies were case-control, two were cohort, one was case-cohort, and another was case-cohort and cohort. Interestingly, the first GWAS in the field of cannabis use was published in 2011 and the most recent conducted in 2019(26,27). All studies used data from large study datasets. Three studies utilized the Study of Addiction: Genetics and Environment (SAGE)(4,26,28). The International Cannabis Consortium (ICC), UKBiobank, and 23andMe were utilized in one study which performed three independent GWAS on the aforementioned datasets(9). Another study combined the Yale-Penn and the International Consortium on the Genetics of Heroin Dependence (ICGHD) to perform a single GWAS(4). Finally, one study utilized the Integrative Psychiatric Research (iPSYCH)(27)

and another the Netherlands twin registry(29). Studies varied in size from 3,053-51,372 participants. Of the studies which reported participants' sex and age, three studies had a population comprised of mostly female participants(9,26,28,29), while only one reported majority male(4). The mean age of study participants varied from mid-thirties to mid-fifties. Three studies reported on participants of European or African American ethnicities(4,26,28) and three studies reported a European only ethnicity(9,27,29). Reported outcomes of interest included lifetime cannabis use(8,9), CUD as defined by either the DSM-IV(26) or ICD-10(27), CUD criteria count(4,28) or age of onset of cannabis use(29).

3.2.3.3 Risk of bias within studies

The Q-Genie tool [version 1.1] was completed in duplicate and used to assess study quality. Studies were assessed on a scale of 1 to 7 for 11 items. An overall score greater than or equal to 45 for studies with a control group and studies with an overall score greater than 40 without a control group were considered good quality according to the Q-Genie tool(22). All studies were considered to be good quality except for one study, Minica et al., which was deemed moderate quality. It should be noted that Minica et al. did not discuss any potential sources of bias or limitations within their study. Additionally, the study was conducted using the Netherlands twin registry and while individuals with a genetic relatedness larger than 0.025 were excluded for some analyses, heritability was not accounted for in all analyses and may therefore introduce bias(29). Three studies reported potential conflicts of interest due to involvement with industry funding(4,26,28), two studies report conflict in a patent involved in identifying SNPs associated with addiction(26,28) and one study reports authors are employees of deCODE genetics(27). Please see Table 3 for the studies Q-genie scores of the included studies.

3.2.3.4 Results of individual studies

All six studies included in this systematic review reported outcomes that reached the significance threshold set a priori (Table 4).

Agrawal et al. (2011) identified two SNPs associated with DSM-IV cannabis dependence within the *ANKFN1* gene (chromosome 17). European and African American participants were selected from the SAGE study which was aimed to primarily study DSM-IV alcohol dependence. Case status was defined as a lifetime history of DSM-IV cannabis dependence, with controls defined as using cannabis at least once in their lifetime but not meeting criteria for DSM-IV cannabis dependence(26).

Agrawal et al. (2014) identified a SNP reaching borderline significance threshold on chromosome 3 associated with CUD factor scores in African Americans, however, no associated gene was identified. Participants were European and African Americans selected from the SAGE study. DSM factor scores were developed from 12 DSM-IV and DSM-5 criteria for CUD(28).

Demontis et al. identified 26 SNPs associated with CUD on chromosome 8, with no associated gene identified. However, only 5 SNPs were discussed and identified in the paper, and thus only 5 SNPs are reported in this review. Participants were selected from

the iPSYCH cohort and were of European ancestry. The iPSYCH cohort was established to study six major psychiatric disorders, however, identified participants meeting ICD-10 CUD(27).

Minica et al. reported 3 SNPs associated with cannabis initiation and 24 SNPs associated with the age of onset of cannabis use. Identified SNPs were found on chromosomes 5, 9, 18 and 19, with one SNP associated with cannabis initiation was found on the Zinc finger protein, *ZNF181*. All participants were of European descent and were selected from the Netherlands Twin Registry. Cannabis initiation was defined as ever/never having used cannabis while age of onset was determined by asking participants an open-ended question(29).

Pasman et al. conducted three independent GWASs in three separate cohorts, all of which included European participants: ICC, UKBiobank, and 23andMe. While results from 23andMe were unable to be shared due to privacy policies, the lead author kindly provided SNPs reaching borderline significance threshold with lifetime cannabis use for GWAS conducted in the ICC and UKBiobank cohort. One SNP in the ICC cohort and 18 SNPs in the UKBiobank were associated with lifetime cannabis use, with no genes specified in either. Lifetime cannabis use was defined as any cannabis use during lifetime(9).

Sherva et al. identified 42 SNPs associated with DSM-IV cannabis dependence criteria count across 27 different genes/regions including *INTS7, SNORA26, RPS20P10, PI4K2B, CSMD1, PSMB7, HABP2, MEFV, CST7, APMAP, ACSS1, snoU13, TPST2, SCN9A, CTA-445C9.15, CTA-445C9.14-CTA-4, SCN9A-SCN7A, ARL2BPP5-RP11-541P9.3, RP11-755E23.3-CCDC67, SNORD11-RNU6-1014P, RP5-860P4.2-CST7, RNU6-1257P, APMAP-ACSS1, C9.15, RPS20P10-CYP26B1, PI4K2B-ZCCHC4*, and *CST7-APMAP*. European and African American participants were selected from the Yale-Penn Study, the SAGE study and the ICGHD cohorts(4).

While no SNPs were reported within the same region not allowing further quantitative analysis, several phenotypic similarities exist across studies. Interestingly, two studies found that educational attainment was negatively associated with CUD(26,27) and a third found positive genetic correlations with educational attainment(9). Two studies found that cannabis dependence was significantly related to alcohol, nicotine, and cocaine dependence(4,26) with a third reporting a positive genetic correlation between lifetime cannabis use and smoking and alcohol use and dependence(9).

3.2.3.5 Risk of bias across studies

Outcomes assessed for GRADE include lifetime cannabis use, diagnosis of CUD, criterion count for CUD and age of onset of cannabis use. All outcomes included two studies except for age of onset of cannabis use which only included one study. The full GRADE table can be found in Table 5. All outcomes were rates as important, and no outcome was rated as having a "very serious" concern pertaining to any certainty criteria. Only the outcomes of diagnosis of CUD and criterion count for CUD had a serious rating, both of which were in the category of indirectness. Both of these outcomes were downgraded due to the use of different diagnostic criteria. More specifically, for the

outcome of diagnosis of CUD Agrawal et al. (2011) utilized the DSM-IV and Demontis et al. (2019) utilized the ICD-10 and for the outcome of criterion count of CUD Sherva et al. (2016) utilized the DSM-IV criteria and Agrawal et al. (2014) utilized a combination of DSM-IV and DSM-5 criteria.

3.2.4 Discussion

3.2.4.1 Summary of evidence

In this review we identified 96 genetic variants to be associated with different measures of cannabis. Of these genetic variants, 18 reached the genome-wide significance threshold of $p \leq 5 \times 10^{-8}$, all of which are available in Table 4. As no genetic variants included in this review were reported in more than one study, meta-analyses were not possible. However, of the genetic variants identified in this review, several are located on genes in which previous studies have reported associations with mental health, namely *ANKFN1*, *INTS7*, *PI4K2B*, *CSMD1*, *CST7*, *ACSS1*, and *SCN9A*.

With cannabis being a legal substance, research on the benefits and harms of cannabis has been on the rise. However, a limited number of GWASs have been conducted on cannabis use to determine any genetic associations. This systematic review was able to qualitatively summarize findings from GWASs reporting borderline genome-wide significance to aid in identifying SNPs that may be replicable in future studies. We have identified six eligible studies that reported independent GWAS results, one of which primarily focused on a GWAS meta-analysis. Of the included studies, only participants from European or African American ethnicities were included, suggesting a need for genetic studies being conducted in more diverse ethnic populations. All six studies reported at least one borderline significant SNP; however, no two studies identified the same SNP. SNPs were found to be associated with CUD, cannabis initiation, age of onset of cannabis use, DSM-IV cannabis dependence criteria count, or lifetime cannabis use on various gene regions. According to assessment using the Q-genie tool and GRADE tool, no study or outcome was deemed to be of poor quality. Additionally, with GWAS requiring a sample size of thousands of participants for adequate power, all studies met this threshold(30).

While the majority of genes identified in the included studies had either no known function or biological plausibility, and none had any additional associations with cannabis use, as mentioned above, several did have associations with mental health conditions and are discussed briefly, namely *ANKFN1*, *INTS7*, *PI4K2B*, *CSMD1*, *CST7*, *ACSS1*, and *SCN9A*. *ANKFN1* is a protein coding gene which has been associated with smoking cessation and nicotine dependence(31). *INTS7* is a component of the integrator complex, which is involved in the small nuclear RNA U1 and U2 transcriptions(32) and has been associated with bipolar temperament(31,33). *PI4K2B* contributes to the overall PI4-kinase activity of the cell(32) and is associated with attention deficit hyperactivity disorder (ADHD), logical memory and abnormality of neuronal migration(33). *CSMD1* has been associated with behavioural disinhibition, schizophrenia, cognitive tests, chronic bronchitis, and bipolar disorder(31,33). *CST7* is associated with alcohol consumption and myocardial infarction(31,33). *ACSS1* catalyzes the synthesis of acetyl-CoA and has been

associated with performance on standardized cognitive tests and bitter alcoholic beverage consumption(31–33). *SCN9A* medicates the voltage-dependent sodium ion permeability of excitable membranes and plays a role in pain mechanisms, especially in the development of inflammatory pain(31). As it is known that cannabis can have a negative impact on learning, memory and chronic bronchitis, known relation to mental illness and suggested role in pain management, these regions may have implications in cannabis use despite having no clear known biological relevance(3,19).

Additionally, it is also important to highlight that genes identified in this review associated with cannabis use or CUD have also been associated with other neuropsychiatric disorders namely nicotine dependence, ADHD, bipolar disorder, schizophrenia, and alcohol consumption suggesting that the genetic risk for the development of these disorders may not be independent. Previously genetic associations have been found amongst schizophrenia, bipolar disorder, ADHD, depression, and autism spectrum disorder, with a high genetic correlation between schizophrenia and bipolar disorder and a moderate correlation between ADHD and depression, ADHD and autism spectrum disorder, and ADHD and depression(34). A recent GWAS meta-analysis added to the evidence on shared genetic associations amongst neuropsychiatric disorders by identifying that an increased risk of cannabis use disorder is genetically correlated with increased liability for smoking initiation, alcohol use, nicotine dependence, and psychiatric disorders (e.g. ADHD, schizophrenia, major depression)(35). These genetic correlations among neuropsychiatric disorders, including cannabis, could reflect genuine pleiotropy or could indicate these psychiatric disorders, including CUD, are not completely independent(34,35). As such, it is important to discuss the biological and individual factors that influence the development of neuropsychiatric disorders.

Neuropsychiatric disorders are influenced by a range of factors, including genetics, personality/mood characteristics, psychological status, behaviour, neurocognitive functioning, and demographic characteristics(36,37). To begin, non-specific to CUD, the prenatal environment, including prenatal nutrition, maternal stress, and maternal substance abuse, can impact brain development and therefore the behavioural outcome of children. Potential mechanisms through which the prenatal environment can impact brain development occurs on multiple levels including genetic selection, epigenetic modification, mediation of brain-immune communications, abnormal metabolism pathways, synthetic mediation of hormones and the hypothalamic-pituitary-adrenal axis, and mediation of the microbiota-gut brain axis(37). Furthermore, nutritional deficiency during critical stages of pregnancy has been linked to emotional and behavioural problems in children including decreased attention, decreased IQ, language delay, and neurodevelopment and related neuropsychiatric disorders(37–39). More specifically, prenatal malnutrition has been linked to an increased risk of schizophrenia during the 1944-1945 Dutch Hunger Winter and the 1959-1961 Chinese famine. Additionally, a “U” relationship between serum 12(OH)D concentration and emotion, behaviour and attention has been found(38,40). Interestingly, the hippocampus, which plays an important role in learning and memory, has been suggested to be sensitive to the exposure of prenatal nutrition deficiency(39,41). The hippocampus has also been proven to be crucial in the pathophysiology of many neuropsychiatric disorders, in which the

changes result from alerted brain development(41). Maternal stress has been associated with poor offspring outcomes including cognition, health and educational attainment, however methodological challenges exist leading to potential misattribution of socially mediated (i.e. postnatal parenting) mechanisms to biological ones (i.e. alterations to developing fetal brain)(42,43). Finally, prenatal exposure to alcohol and other substances has been increasingly common and the consequence of the exposure differs depending on the substance used. Alcohol, tobacco, cannabis and opioids are among the most frequent used substances during pregnancy and offspring outcomes may include birth defects, developmental disability, fetal alcohol syndrome, childhood obesity, decreased birth weight, poor inhibitory control and other organ deficits(44). Thus, many neuropsychiatric disorders appear to result from interactions among genetic background, the prenatal environment and postnatal lifestyle choices(45,46). Given the known association between deficits within the prenatal environment and other neuropsychiatric disorders it is plausible to suggest that the prenatal environment and subsequential gene expression may play a role in future cannabis use and/or CUD.

As previously mentioned, a variety of factors contribute to the complex etiology of neuropsychiatric disorders such as epigenetic modification. Epigenetic modifications that can regulate gene expression include DNA methylation, nucleosomal structure and positioning, post-translational modification of nucleosome histones, histone replacement and small RNA molecules that influence protein production(47). The most studied form of epigenetic modification is DNA methylation, which can be influenced by a range of factors including genetic factors, disease, environmental exposures, and lifestyle. DNA methylation changes can be either persistent or reversible once the exposure is no longer present, adding value for biomarker development(48). How cannabis, THC and other exogenous cannabis receptor modulators alter epigenetic mechanisms have been previously reviewed(47). Relatively little is known about the molecular pathways influenced by cannabis, however, one study identified 13 proteins, 3 metabolites and 2 lipids significantly associated with a metabolite of THC and another found acute effects of cannabis or THC on the central nervous system and heart rate(49,50). In addition to DNA methylation, post-transcriptional chemical medication of RNA is rapidly emerging as a key role in regulating gene expression, known as epitranscriptomics(51). Of growing interest within this felid is N4-acetylcytidine (ac4C), a key role in the transcriptional translation process. ac4C has been implicated in the occurrence of various disease such as inflammation, metabolic diseases, autoimmune diseases, and cancer(52). While the role ac4C may play in neuropsychiatric disorders remains unknown, it is important to consider the role epitransciptomics plays in the gene expression with normal development.

Current knowledge on cannabis has demonstrated that cannabis can induce structural changes to brain regions including the hippocampus, amygdala, cerebellum, prefrontal cortex and striatum as well as grey matter volume(53–55). Potential pre-existing neurobiological factors may exist in cannabis use as well as gene x drug interactions. For instance, in young teens, reduced orbitofrontal cortex volume has been found to predict initiation of cannabis use in later adolescence. The G allele of rs2023239 of *CNR1* is linked with higher cortical CBR1 and is associated with smaller hippocampal volume in chronic cannabis users, but not healthy controls and findings that suggest only

individuals with a high genetic risk of schizophrenia experience a negative impact on cortical maturation during early adolescence thus suggestive of gene x drug interactions(56–58). In addition, functional MRI evidence suggest specific brain activity signatures with cannabis use such as increased functional connectivity associated with the default node network and insula networks and hippocampal and parahippocampal atrophy have been associated with chronic cannabis use(59,60). However, neuroimaging studies of cannabis users have yielded inconsistent findings and may reflect individual differences that preceded cannabis use. The inconsistent findings in the literature highlight the need for large longitudinal studies utilizing before-and-after cannabis use neuroimaging(61). Taken together, it is plausible that structural differences in brain regions could be influenced by genetic differences between individuals, explaining the mixed evidence within neuroimaging. Further research is required to determine the complex interactions amongst individual genetic predispositions, prenatal environment, and postnatal environment contributing to individual cannabis use behaviour and/or the development of CUD. Understanding the genetic predispositions is one piece of the puzzle in understanding the complex development of cannabis use and CUD.

Finally, it is important to consider the shared genetic basis of other substance use disorders. Heritability estimates across substance use disorders vary, with heritability lowest for hallucinogens (0.39) and highest for cocaine use (0.72)(62,63). Additionally, substance use disorders are the result of gene x environment interactions, with partial risk inborn and another part determined by environmental experiences(62). Previous reviews have summarized the literature on GWASs for various substance use disorders including alcohol use disorder, nicotine use disorder, CUD, OUD, and cocaine use disorder. However, genetic studies within specific substance use disorders have had varying success in replicating previously identified associations, limiting evidence for shared genetic basis across substance use disorders(63,64). The complexity of substance use disorder make genetic prediction efforts difficult, and while currently only alcohol use disorder have been genetically correlated with CUD, continued advancements in molecular genetic studies and substance use disorder at larges further our understanding of the biological pathways underlying substance use disorders(9,63,65). For instance, *CNR1* and *CNR2*, components of the endocannabinoid system, are major targets of investigation for their impact in neuropsychiatry and addiction phenotypes suggested shared genetic risk factors(66,67). In regards to neuropsychiatric disorders, Mendelian randomization studies have found mixed evidence on the causal effect of cannabis initiation and schizophrenia, finding weak evidence that cannabis initiation increases schizophrenia risk and strong evidence that schizophrenia liability increases the odds of cannabis initiation, and causal evidence of ADHD on cannabis initiation(68–72). Through continued advances, it is hoped that the underlying genetic basis for CUD, or a shared genetic basis for all substance use disorders, will be identified to provide preventative measures and treatment for substance use disorders in the future.

3.2.4.2 Limitations

While this systematic review was rigorous and involved a peer-reviewed protocol, it is not without limitations. First, our inclusion criteria limited our review to only GWASs, meaning any GWAS meta-analyses and candidate gene studies were excluded. GWAS meta-analyses and candidate gene studies are often more powered due to their larger sample sizes and minimal genetic variants tested, respectfully(11). However, including only GWASs was decided *a priori* to capture novel genetic variants associated with cannabis use and avoid the inclusion of multiple studies which could use the same genetic dataset. Second, it is important to note that this review is susceptible to publication bias, as studies that do not achieve genome-wide significance may be less likely to be published, and thus, not included in this review. Unpublished GWAS findings may exist with SNPs reaching the borderline significance threshold. While we cannot eliminate publication bias entirely, we searched abstracts, GWAS catalogs, and databases for any near significant findings that were not published. If a relevant abstract was identified, without the full study published, the first author was contacted to determine whether the full GWAS had been published or was going to be submitted to a journal. Finally, if a study met our inclusion criteria but did not report any SNPs that fell below the genome-wide significance threshold, study authors were contacted to confirm if any SNPs had reached the borderline significant threshold set for this review. Third, due to the heterogeneity of the reported findings, it was not possible to conduct a meta-analysis or sex and ethnicity subgroup analyses. Although we could not conduct a meta-analysis, we qualitatively summarized the studies and reported a comprehensive list of all SNPs reaching the significance threshold for this study.

3.2.5 Conclusions

This systematic review was able to summarize GWAS findings within the field of cannabis use. The results can inform future candidate gene studies and GWASs of possible replicable SNPs that require further investigation. We were able to identify all GWASs conducted on cannabis use, highlighting the need for further research as no two GWASs reported the same SNP or gene associated with cannabis use. Further, included GWASs had limited ethnic diversity, with only European or African American participants. Recommendations are made for future research to replicate reported associations and include diverse ethnic populations to test whether SNPs associated with cannabis use reported are generalizable across study populations and if associations differ by ethnicity.

3.2.6 List of Abbreviations

GWAS: Genome-wide Association Study
CUD: Cannabis Use Disorder
GWASs: Genome-wide Association Studies
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
HuGeNET: The Human Genome Epidemiology Network
PROSPERO: International Prospective Register of Systematic Reviews
DSM: Diagnostic and Statistical Manual
ICD-10: International Statistical Classification of Diseases and Related Health Problems - 10
Q-Genie: The quality of genetic association studies
SAGE: Study of Addiction: Genetics and Environment
ICC: International Cannabis Consortium
ICGHD: International Consortium on the Genetics of Heroin Dependence
iPSYCH: Integrative Psychiatric Research
ADHD: Attention-Deficit Hyperactivity Disorder

3.2.7 Declarations

3.2.7.1 Ethics approval and consent to participate

Not applicable.

3.2.7.2 Consent for publication

Not applicable.

3.2.7.3 Availability of data and materials

All data generated or analysed during this study are included in this published article. Primary articles included in this systematic review can be found at the following links:
https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1369-1600.2010.00255.x
https://www.sciencedirect.com/science/article/abs/pii/S0376871613004766
https://www.nature.com/articles/s41593-019-0416-1
https://link.springer.com/article/10.1007/s10519-015-9723-9
https://www.nature.com/articles/s41593-018-0206-1?fbclid=IwAR20cNM7ZDBKTGp7QNtm521tkEECrKyDTFLQ5pQV9emvsIbrsM_TMeAUTSg
https://jamanetwork.com/journals/jamapsychiatry/article-abstract/2504223

3.2.7.4 Competing interests

The authors declare they have no competing interests.

3.2.7.5 Funding

This systematic review was partially funded by CIHR (Grant number PJT-156306). The funding agency had no role in the study design, interpretation of data, or dissemination of results.

3.2.7.6 Authors contributions

ZS is the guarantor. AH, CC and ZS conceptualized the systematic review. AH implemented the design of the review and search strategy with the aid of CC and SS. AH, CC, AD, MB and RK screened studies, extracted data and assess the quality of the studies. AH prepared the first draft, and the final manuscript was reviewed and revised by CC, AD, MB, RK, FK, LT and ZS. All authors read and approved the final manuscript.

3.2.7.7 Acknowledgements

Not applicable.

3.2.8 References

1. Degenhardt L, Charlson F, Ferrari A, Santomauro D, Erskine H, Mantilla-Herrara A, et al. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Psychiatry. 2018;5(12):987–1012.
2. Statistics Canada. National Cannabis Survey, forth quarter 2019. 2019.
3. Volkow ND, Baler RD, Compton WM, Weiss SRB. Adverse Health Effects of Marijuana Use. N Engl J Med. 2014;370(23):2219–27.
4. Sherva R, Wang Q, Kranzler H, Zhao H, Koesterer R, Herman A, et al. Genome-wide association study of cannabis dependence severity, novel risk variants, and shared genetic risks. JAMA Psychiatry. 2016;73(5):472–80.
5. Bahji A, Stephenson C. International perspectives on the implications of cannabis legalization: a systematic review & thematic analysis. Int J Environ Res Public Health. 2019;16(17):3095.
6. Verweij KJH, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG, et al. Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. Addiction. 2010;105(3):417–30.
7. Agrawal A, Chou YL, Carey CE, Baranger DAA, Zhang B, Sherva R, et al. Genome-wide association study identifies a novel locus for cannabis dependence. Mol Psychiatry. 2018;23(5):1293–302.
8. Minică CC, Verweij KJH, Most PJ, Mbarek H, Bernard M, Eijk KR, et al. Genome-wide association meta-analysis of age at first cannabis use. Addiction [Internet]. 2018 Nov;113(11):2073–86. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=132211589&site=ehost-live
9. Pasman JA, Verweij KJH, Gerring Z, Stringer S, Sanchez-Roige S, Treur JL, et al. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. Nat Neurosci [Internet]. 2018;21(9):1161–70. Available from: http://dx.doi.org/10.1038/s41593-018-0206-1
10. Stringer S, Minică CC, Verweij KJH, Mbarek H, Bernard M, Derringer J, et al. Genome-wide association study of lifetime cannabis use based on a large meta-analytic sample of 32 330 subjects from the International Cannabis Consortium. Transl Psychiatry. 2016;6(3):e769.
11. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019;20(8):467–84.
12. The Government of Canada. Cannabis Health Effects. 2019.
13. Hall W. What has research over the past two decades revealed about the adverse health effects of recreational cannabis use? Addiction. 2015;110(1):19–35.
14. Gobbi G, Atkin T, Zytynski T, Wang S, Askari S, Boruff J, et al. Association of cannabis use in adolescence and risk of depression, anxiety, and suicidality in young adulthood: a systematic review and meta-analysis. JAMA psychiatry. 2019;76(4):426–34.

15. Moher D, Liberati A, Tetzlaff J, Altman D, Group P, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med. 2009;6(7):e1000097.
16. Little J, Higgins JPT, Bray M, Ioannidis J, Khoury M, Manolio T, et al. The HuGENet™ HuGE review handbook, version 1.0. Ottawa, Ontario, Canada HuGENet Canada Coord Cent. 2006;
17. Hillmer A, Chawar C, D'Elia A, Kapoor R, Butt M, Samaan Z. A systematic review and meta-anlaysis of genome-wide association studies and cannabis use. PROSPERO: International prospective register of systematic reivews. 2020.
18. Hillmer A, Chawar C, Sanger S, D'Elia A, Butt M, Kapoor R, et al. Genetic determinants of cannabis use: a systematic review protocol. Syst Rev. 2020;9(1):1–6.
19. National Academies of Sciences Engineering and Medicine. The health effects of cannabis and cannabinoids: The current state of evidence and recommendations for research [Internet]. Washington, DC: National Academies Press; 2017. Available from: https://www.ncbi.nlm.nih.gov/books/NBK423845/ doi: 10.17226/24625
20. Veritas Health Innovation. Covidence systematic review software [Internet]. Melbourne, Australia; Available from: https://www.covidence.org/
21. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. Genet Epidemiol Off Publ Int Genet Epidemiol Soc. 2008;32(3):227–34.
22. Sohani ZN, Meyre D, de Souza RJ, Joseph PG, Gandhi M, Dennis BB, et al. Assessing the quality of published genetic association studies in meta-analyses: The quality of genetic studies (Q-Genie) tool. BMC Genet. 2015;16.
23. Guyatt GH, Oxman AD, Schünemann H, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol. 2011;64(4):380–2.
24. Foroutan F, Guyatt G, Zuk V, Vandvik PO, Alba AC, Mustafa R, et al. GRADE Guidelines 28: Use of GRADE for the assessment of evidence about prognostic factors: rating certainty in identification of groups of patients with different absolute risks. J Clin Epidemiol. 2020;121:62–70.
25. Center for History New Media. Zotero: The Next-Generation Research Tool [Internet]. George Mason University; 2009. Available from: http://www.zotero.org/
26. Agrawal A, Lynskey MT, Hinrichs A, Grucza R, Saccone SF, Krueger R, et al. A genome-wide association study of DSM-IV cannabis dependence. Addict Biol. 2011;16(3):514–8.
27. Demontis D, Rajagopal VM, Thorgeirsson TE, Als TD, Grove J, Leppälä K, et al. Genome-wide association study implicates CHRNA2 in cannabis use disorder. Nat Neurosci. 2019;22(7):1066.
28. Agrawal A, Lynskey MT, Bucholz KK, Kapoor M, Almasy L, Dick DM, et al. DSM-5 cannabis use disorder: A phenotypic and genomic perspective. Drug Alcohol Depend [Internet]. 2014;134(1):362–9. Available from: http://dx.doi.org/10.1016/j.drugalcdep.2013.11.008

29. Minică CC, Dolan C V, Hottenga J-J, Pool R, Fedko IO, Mbarek H, et al. Heritability, SNP-and gene-based analyses of cannabis use initiation and age at onset. Behav Genet. 2015;45(5):503–13.
30. Ziyatdinov A, Kim J, Prokopenko D, Privé F, Laporte F, Loh P-R, et al. Estimating the effective sample size in association studies of quantitative traits. bioRxiv. 2019;
31. Bethesda (MD): National Library of Medicine (US) National Centre for Biotechnology Infromation. dbGene [Internet]. 2004. Available from: https://www.ncbi.nlm.nih.gov/gene/
32. UniProt Consortium. UniProt: a worldwide hub of protien knowledge. Nucleic Acids Res. 2019;47(D1):D506–15.
33. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. Ensembl 2020. Nucleic Acids Res. 2020;48(D1):D682–8.
34. Wu Y, Cao H, Baranova A, Huang H, Li S, Cai L, et al. Multi-trait analysis for genome-wide association study of five psychiatric disorders. Transl Psychiatry. 2020;10(1):1–11.
35. Johnson EC, Demontis D, Thorgeirsson TE, Walters RK, Polimanti R, Hatoum AS, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. The Lancet Psychiatry. 2020;7(12):1032–45.
36. Taylor M, Cousijn J, Filbey F. Determining Risks for Cannabis Use Disorder in the Face of Changing Legal Policies. Curr Addict reports [Internet]. 2019/10/24. 2019 Dec;6(4):466–77. Available from: https://pubmed.ncbi.nlm.nih.gov/33312839
37. Yan X, Zhao X, Li J, He L, Xu M. Effects of early-life malnutrition on neurodevelopment and neuropsychiatric disorders and the potential mechanisms. Prog Neuro-Psychopharmacology Biol Psychiatry. 2018;83:64–75.
38. Wang X, Jiao X, Xu M, Wang B, Li J, Yang F, et al. Effects of circulating vitamin D concentrations on emotion, behavior and attention: A cross-sectional study in preschool children with follow-up behavior experiments in juvenile mice. J Affect Disord. 2020;275:290–8.
39. Chen J, Zhao X, Cui L, He G, Wang X, Wang F, et al. Genetic regulatory subnetworks and key regulating genes in rat hippocampus perturbed by prenatal malnutrition: implications for major brain disorders. Aging (Albany NY). 2020;12(9):8434.
40. Xu M-Q, Sun W-S, Liu B-X, Feng G-Y, Yu L, Yang L, et al. Prenatal malnutrition and adult schizophrenia: further evidence from the 1959-1961 Chinese famine. Schizophr Bull. 2009;35(3):568–76.
41. Oberlander VC, Xu X, Chini M, Hanganu-Opatz IL. Developmental dysfunction of prefrontal–hippocampal networks in mouse models of mental illness. Eur J Neurosci. 2019;50(6):3072–84.
42. Aizer A, Stroud L, Buka S. Maternal stress and child outcomes: Evidence from siblings. J Hum Resour. 2016;51(3):523–55.
43. DiPietro JA. Maternal stress in pregnancy: considerations for fetal development. J Adolesc Heal. 2012;51(2):S3–8.
44. Chang G. Maternal Substance Use: Consequences, Identification, and Interventions. Alcohol Res [Internet]. 2020 Jun 30;40(2):6. Available from:

https://pubmed.ncbi.nlm.nih.gov/32612898
45. Newnham JP, Moss TJM, Nitsos I, Sloboda DM, Challis JRG. Nutrition and the early origins of adult disease. Asia Pac J Clin Nutr. 2002;11:S537–42.
46. Taylor PD, Poston L. Developmental programming of obesity in mammals. Exp Physiol. 2007;92(2):287–98.
47. Szutorisz H, Hurd YL. Epigenetic Effects of Cannabis Exposure. Biol Psychiatry [Internet]. 2015/11/03. 2016 Apr 1;79(7):586–94. Available from: https://pubmed.ncbi.nlm.nih.gov/26546076
48. Markunas CA, Hancock DB, Xu Z, Quach BC, Fang F, Sandler DP, et al. Epigenome-wide analysis uncovers a blood-based DNA methylation biomarker of lifetime cannabis use. Am J Med Genet Part B Neuropsychiatr Genet. 2021;186(3):173–82.
49. Hinckley JD, Saba L, Raymond K, Bartels K, Klawitter J, Christians U, et al. An Approach to Biomarker Discovery of Cannabis Use Utilizing Proteomic, Metabolomic, and Lipidomic Analyses. Cannabis Cannabinoid Res. 2020;
50. Zuurman L, Ippel AE, Moin E, Van Gerven JMA. Biomarkers for the effects of cannabis and THC in healthy volunteers. Br J Clin Pharmacol. 2009;67(1):5–21.
51. Karthiya R, Wasil SM, Khandelia P. Emerging role of N4-acetylcytidine modification of RNA in gene regulation and cellular functions. Mol Biol Rep. 2020;1–11.
52. Jin G, Xu M, Zou M, Duan S. The processing, gene regulation, biological functions, and clinical relevance of N4-acetylcytidine on RNA: a systematic review. Mol Ther Acids. 2020;20:13–24.
53. Battistella G, Fornari E, Annoni J-M, Chtioui H, Dao K, Fabritius M, et al. Long-term effects of cannabis on brain structure. Neuropsychopharmacology. 2014;39(9):2041–8.
54. Scheffler F, Du Plessis S, Asmal L, Kilian S, Phahladira L, Luckhoff HK, et al. Cannabis use and hippocampal subfield volumes in males with a first episode of a schizophrenia spectrum disorder and healthy controls. Schizophr Res. 2021;231:13–21.
55. Batalla A, Bhattacharyya S, Yuecel M, Fusar-Poli P, Crippa JA, Nogue S, et al. Structural and functional imaging studies in chronic cannabis users: a systematic review of adolescent and adult findings. PLoS One. 2013;8(2):e55821.
56. Schacht JP, Hutchison KE, Filbey FM. Associations between cannabinoid receptor-1 (CNR1) variation and hippocampus and amygdala volumes in heavy cannabis users. Neuropsychopharmacology. 2012;37(11):2368–76.
57. French L, Gray C, Leonard G, Perron M, Pike GB, Richer L, et al. Early cannabis use, polygenic risk score for schizophrenia and brain maturation in adolescence. JAMA psychiatry. 2015;72(10):1002–11.
58. Cheetham A, Allen NB, Whittle S, Simmons JG, Yücel M, Lubman DI. Orbitofrontal volumes in early adolescence predict initiation of cannabis use: a 4-year longitudinal and prospective study. Biol Psychiatry. 2012;71(8):684–92.
59. Bloomfield MAP, Hindocha C, Green SF, Wall MB, Lees R, Petrilli K, et al. The neuropsychopharmacology of cannabis: A review of human imaging studies.

Pharmacol Ther. 2019;195:132–61.
60. Ferland J-MN, Hurd YL. Deconstructing the neurobiology of cannabis use disorder. Nat Neurosci. 2020;23(5):600–10.
61. Volkow ND, Swanson JM, Evins AE, DeLisi LE, Meier MH, Gonzalez R, et al. Effects of cannabis use on human behavior, including cognition, motivation, and psychosis: a review. JAMA psychiatry. 2016;73(3):292–7.
62. Ducci F, Goldman D. The genetic basis of addictive disorders. Psychiatr Clin. 2012;35(2):495–519.
63. Deak JD, Johnson EC. Genetics of substance use disorders: a review. Psychol Med. 2021;1–12.
64. Prom-Wormley EC, Ebejer J, Dick DM, Bowers MS. The genetic epidemiology of substance use disorder: a review. Drug Alcohol Depend. 2017;180:241–59.
65. Zhou H, Sealock JM, Sanchez-Roige S, Clarke T-K, Levey DF, Cheng Z, et al. Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. Nat Neurosci. 2020;23(7):809–18.
66. Onaivi ES, Ishiguro H, Sgro S, Leonard CM. Cannabinoid receptor gene variations in drug addiction and neuropsychiatric disorders. J Drug Alcohol Res. 2013;2(1):1–11.
67. Marco EM, García-Gutiérrez MS, Bermúdez-Silva F-J, Moreira F, Guimarães F, Manzanares J, et al. Endocannabinoid system and psychiatry: in search of a neurobiological basis for detrimental and potential therapeutic effects. Front Behav Neurosci. 2011;5:63.
68. Gage SH, Jones HJ, Burgess S, Bowden J, Davey Smith G, Zammit S, et al. Assessing causality in associations between cannabis use and schizophrenia risk: a two-sample Mendelian randomization study. Psychol Med [Internet]. 2017 Apr;47(5):971–80. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=121624529&site=ehost-live
69. Vaucher J, Keating BJ, Lasserre AM, Gan W, Lyall DM, Ward J, et al. Cannabis use and risk of schizophrenia: a Mendelian randomization study. Mol Psychiatry. 2018;23(5):1287–92.
70. Johnson EC, Hatoum AS, Deak JD, Polimanti R, Murray RM, Edenberg HJ, et al. The relationship between cannabis and schizophrenia: a genetically informed perspective. Addiction. 2021;
71. Treur JL, Demontis D, Smith GD, Sallis H, Richardson TG, Wiers RW, et al. Investigating causality between liability to ADHD and substance use, and liability to substance use and ADHD risk, using Mendelian randomization. Addict Biol. 2021;26(1):e12849.
72. Artigas MS, Sánchez-Mora C, Rovira P, Richarte V, Garcia-Martínez I, Pagerols M, et al. Attention-deficit/hyperactivity disorder and lifetime cannabis use: genetic overlap and causality. Mol Psychiatry. 2020;25(10):2493–503.

3.2.9 Tables

Table 1. Search strategy

OVID MEDLINE	1. Genome-Wide Association Study/ 2. Genotyping Techniques/ 3. Genome, Human/ 4. Genetic Variation/ 5. genetics/ or exp human genetics/ 6. (human* adj2 (genotyp* or genome* or genetic*)).ti,ab,kw,kf. 7. (GWS or GWAS or GWA).mp. 8. genome wide.ti,ab,kw,kf. 9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 10. exp Cannabis/ 11. ((cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) adj2 (overdose* or use* or using or misuse* or abus* or dependence* or addict*)).ti,ab,kw,kf. 12. 10 or 11 13. 9 and 12 *14. Limit 13 to humans*
Web of Science	1. TS=(genome-wide association study or genome-wide association or GWAS or GWA or genome wide) 2. TS=(human NEAR/2 genome) 3. TS=((cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) NEAR/2 (overdose* or use* or using or misuse* or abus* or dependence* or addict*)) 4. TS=(cannabis* or marijuana* or marihuana*) 5. #1 OR #2 6. #3 OR #4

	7. #5 and #6
OVID EMBASE	1. Genome-Wide Association Study/ 2. Genotyping Techniques/ 3. Genome, Human/ 4. Genetic Variation/ 5. genetics/ or exp human genetics/ 6. (human* adj2 (genotyp* or genome* or genetic*)).ti,ab,kw. 7. (GWS or GWAS or GWA).mp. 8. genome wide.ti,ab,kw. 9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 10. exp Cannabis/ 11. ((cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) adj2 (overdose* or use* or using or misuse* or abus* or dependence* or addict*)).ti,ab,kw. 12. 10 or 11 13. 9 and 12 *14. Limit 13 to human*
EBSCOHost CINAHL	1. genome-wide association study or genome-wide association or GWAS or GWA or genome wide or genome 2. cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) 3. overdose* or use* or using or misuse* or abus* or dependence* or addict* 4. S2 and S3 5. S1 and S4 6. *Limit to Human*
GWAS Catalog	Terms Searched: - Cannabis - Cannabis dependence

	- Marihuana - Marijuana - Cannabinoids - Hash - Kush - Weed - Pot - THC - CBD
GWAS Central	Terms Searched: - Cannabis - Cannabis dependence - Marijuana - Marihuana - Cannabinoids - Hash - Kush - Weed - Pot - THC - CBD
NIH Database of Genotypes and Phenotypes	Terms Searched: - Cannabis - Cannabis dependence - Marijuana - THC - Marihuana - Cannabinoids

	- Hash - Kush - Weed - Pot - CBD

Table 2. Individual Study Characteristics

First Author Last Name, Year	Title of Publication	Study Design	Cohort used	Sample Size	% Male	Mean age	Ethnicity	Outcome of interest
Agrawal, 2011	A Genome-wide Association Study of DSM-IV Cannabis Dependence	Case-Control	SAGE	3054	NR	39.00	2019 European-Americans, 1035 African Americans	Life-time history of DSM-IV cannabis dependence, modified to included cannabis withdrawal
Agrawal, 2014	DSM-5 Cannabis Use Disorder: A Phenotypic and Genomic Perspective	Case-control	SAGE	3053	49%	38.10	2018 European-Americans, 1035 African Americans	DSM-5 cannabis use disorder factor scores
Demontis, 2019	Genome-wide association study implicates CHRNA2 in cannabis use disorder	Case-cohort	iPSYCH	CUD: 2387 Control: 48985 Total: 51372	NR	CUD: 24.77 Control: 22.67	European	International Statistical Classification of Disease and Related Health Problems, 10th revision (ICD-10) diagnosis reflecting problematic

								and persistent use of cannabis
Minica, 2015	Heritability, SNP- and Gene-Based Analyses of Cannabis Use Initiation and Age at Onset	Cohort	Netherlands twin registry	6744	39.1%	39.09	European	Self-reported use of cannabis ever in lifetime and self-reported age of onset of cannabis use
Pasman, 2018	GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal effect of schizophrenia liability	Cohort	ICC study	35297	44.3%	35.7	European	Any cannabis use within lifetime
			UK Biobank	126785	43.7%	55		
			23andMe	22683	NR	NR		
Sherva, 2016	Genome-wide Association Study of Cannabis Dependence Severity, Novel Risk Variants, and Shared Genetic Risks	Cohort/Case-cohort	Yale-Penn, SAGE, ICGHD	14754	53.4%	39.24	8754 European-American, 6000 African American	Criterion count for DSM-IV Cannabis Dependence

Table 3. Q-genie scores

First Author Last Name, Year	Reported conflicts of interest	Reported study limitations	Q-Genie Score	Quality Assessment
Agrawal, 2011	Drs. LJ Bierut, J. Rice, A. Goate and S Saccone are listed as inventors on the patent "Markers for Addiction" (US 20070258898): covering the use of certain SNPs in determining the diagnosis, prognosis, and treatment of addiction. Dr. Bierut has acted as a consultant for Pfizer, Inc. in 2008. All other authors report no competing interests.	• SAGE study was ascertained for alcohol dependence led to a high level of comorbidity in the cannabis dependent cases and exposed controls • Use of controls with other forms of substance dependence but not cannabis dependence protected against signals that may have been less specific • Power computations revealed that minor allele frequencies ranging from 15-40% association signals with odds ration exceeding 1.45 were able to be detected	53	Good Quality
Agrawal, 2014	Laura J. Bierut is listed as an inventor on Issued U.S. Patent 8,080,371,"Markers for Addiction" covering the use of certain SNPs in determining the diagnosis, prognosis, and treatment of addiction.	• The sample was ascertained from three family studies of substance use disorders for the express purpose of identifying genetic variants for alcoholism, nicotine and cocaine dependence and related pathology • Differences in DSM-IV and DSM-5 criteria leading to different assessments of withdrawal and diagnosis of CUD across study populations	55	Good quality

Demontis, 2019	T. Werge has been a lecturer and advisor to H. Lundbeck A/S. T.E. Thorgeirsson, D.F. Gudbjartsson, G.W. Reginsson, H. Stefansson and K. Stefansson are employees of deCODE genetics/Amgen.	• None reported	57	Good quality
Minica, 2015	None	• No statistically significant GWAS findings that pass the threshold $p<1.0x10^{-8}$	49	Moderate quality
Pasman, 2018	P.F., S.L.E. and members of the 23andMe Research Team are employees of 23andMe Inc. J.A.R.-Q. was on the speakers' bureau and/or acted as consultant for Eli Lilly, Janssen- Cilag, Novartis, Shire, Lundbeck, Almirall, BRAINGAZE, Sincrolab and Rubió in the last 5 years. He also received travel awards (air tickets and hotel) for taking part in psychiatric meetings from Janssen-Cilag, Rubió, Shire and Eli Lilly. The Department of Psychiatry chaired by him received unrestricted educational and research support from the following pharmaceutical companies in the last 5 years: Eli Lilly, Lundbeck, Janssen- Cilag, Actelion, Shire, Ferrer and Rubió.	• Lifetime cannabis was analyzed as a single dichotomous measure combining experimental and regular users in a single group • The power of some analyses may have been limited	52	Good quality
Sherva, 2016	Dr Kranzler reports being a	• One of the significant SNPs identified	52	Good quality

	consultant or an advisory board member for Alkermes, Indivior, Lundbeck, and Otsuka (unrelated to the present study) and being a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which is supported by AbbVie, Ethypharm, Lilly, Lundbeck, and Pfizer. No other disclosures were reported.	(rs143244591 on chromosome 3) has little supportive evidence for association from other SNPs in the region • None of the GWAS SNPs identified in the full GWAS analysis are rare • Lack of evidence of associations in both the European American and African American participants • The Yale-Penn samples who underwent genotyping on the HumanOmni1-Quad and Human Core Exome chips showed more consistent results than the corresponding SAGE population • The cohorts used have higher rates of polysubstance dependence than the general population and may not be generalizable to individuals who only use cannabis		

Table 4. SNPs reaching borderline significance threshold

First Author Last Name, Year	Outcome associated with SNP	Build	SNP IDs	Chr:Pos	Alleles	Minor Allele	Gene or Locus	MAF	N	Measure of Association type	Measure of Association value	Measure of Variability	Measure of Variability value	p-value	Ethnicity
Agrawal, 2011	DSM-IV cannabis dependence	hg18	rs1019238	17			ANKFN1	EA=0.4, AA=0.1	3054	OR	1.453	95% CI	1.254 - 1.682	6.1E-07	AA & EA
			rs1431318	17			ANKFN1	EA=0.44, AA=0.25	3054	OR	0.708	95% CI	0.616 - 0.812	9.14E-07	AA & EA
Agrawal, 2014	Cannabis use disorder factor scores from 12 DSM IV + 5 criteria	hg18	rs4364205	3		T		0.41	1035	β	-0.19	95% CI	-0.26 - (-0.12)	1.3E-07	AA
Demontis, 2019	Cannabis use disorder	hg19	rs56372821	8	A/G	A		0.163	51372	OR	0.728			**9.31E-12**	EA
			rs4732724	8	C/G	C		0.324	51372	OR	0.82			**1.34E-08**	EA

			rs6558008	8	A/C	A		0.755	51372	OR	1.237			**2.45E-08**	EA
			rs73229090	8	A/C	A		0.11	51372	OR	0.702			**7.41E-10**	EA
			rs937220	8	A/G	A		0.259	51372	OR	0.8258			2.46E-07	EA
Minica, 2015	Cannabis initiation	GR Ch37	rs35487050	19:35221228	C/T		ZNF181		6744	β	0.81	SE	0.16	1.68E-07	E
			rs35917943	19:35147183	C/A			<0.05	6744	β	0.77	SE	0.15	1.62E-07	E
			rs35760174	19:35221582	C/G				6744	β	0.76	SE	0.15	7.04E-07	E
	Age of onset of cannabis use		rs14232406 0	5:95425757	G/A			<0.05	5148	β	0.68	SE	0.11	7.66E-08	E
			rs78505392	5:95422966	C/G				5148	β	0.58	SE	0.1	2.16E-07	E
			rs12003072	9:86771161	A/C				5148	β	0.52	SE	0.09	3.04E-07	E
			rs77097806	5:95456735	A/G				5148	β	0.56	SE	0.1	3.54E-07	E
			rs6879646	5:95450187	A/G				5148	β	0.57	SE	0.1	3.61E-07	E
			rs4613744	5:95451494	C/T				5148	β	0.55	SE	0.1	5.07E-07	E
			rs60218730	5:95492765	G/T				5148	β	0.59	SE	0.11	5.98E-07	E

			rs74305417	9:86779774	C/G				5148	β	0.52	SE	0.09	6.2E-07	E
			rs14298106 9	18:58826022	G/A				5148	β	0.47	SE	0.09	7.25E-07	E
			rs12386084	18:58827145	C/G				5148	β	0.47	SE	0.09	7.25E-07	E
			rs11791893 6	18:58828323	G/A				5148	β	0.47	SE	0.09	7.25E-07	E
			rs2160801	18:58829024	T/A				5148	β	0.47	SE	0.09	7.25E-07	E
			rs14542417 3	18:58829597	T/C				5148	β	0.47	SE	0.09	7.25E-07	E
			rs11753840 9	18:58830942	G/C				5148	β	0.47	SE	0.09	7.25E-07	E
			rs17817245	18:58832135	A/G				5148	β	0.47	SE	0.09	7.25E-07	E
			rs14020680 9	18:58833215	A/G				5148	β	0.47	SE	0.09	7.25E-07	E
			rs11769271 2	18:58834506	T/G				5148	β	0.47	SE	0.09	7.25E-07	E

			rs17817423	18:58835462	C/T				5148	β	0.47	SE	0.09	7.25E-07	E
			rs9916935	18:58835931	T/C				5148	β	0.47	SE	0.09	7.25E-07	E
			rs192013604	18:58838324	T/C				5148	β	0.47	SE	0.09	7.25E-07	E
			rs11747164 0	18:58838402	A/G				5148	β	0.47	SE	0.09	7.25E-07	E
			rs78456402	9:86781900	C/A				5148	β	0.5	SE	0.09	9.09E-07	E
			rs11998981	9:86783107	T/C				5148	β	0.5	SE	0.09	9.09E-07	E
			rs79236058	5:95478830	G/A				5148	β	0.57	SE	0.1	9.59E-07	E
Pasman, 2018	Lifetime cannabis use	hg19	rs7609594	3:85482595	G/A	A	CADM2	0.38	126785	β	0.068	SE	0.01	**5.86E-11**	E
			rs4099556	4:37067936	G/A	A	MIR4801	0.176	126785	β	-0.079	SE	0.013	**3.56E-09**	E
			rs78680891	3:51028954	C/G	C		0.002299	126574	β	-0.52473	SE	0.106473	8.30E-07	E
			rs2071704	4:3240159	C/T	C		0.419929	125953	β	-0.04974	SE	0.010162	9.83E-07	E

			rs76565656	7:148048453	G/A	G		0.122539	124918	β	-0.07974	SE	0.015258	1.73E-07	E
			rs4308708	8:76702058	A/C	A		0.041699	126193	β	-0.13628	SE	0.025107	5.70E-08	E
			rs10883796	10:10465531 5	G/A	G		0.296441	126444	β	-0.05638	SE	0.010951	2.62E-07	E
			rs11186071	10:91969337	G/A	G		0.004168	126711	β	-0.90907	SE	0.18524	9.22E-07	E
			rs11214441	11:11284671 3	T/A	T		0.394257	126761	β	-0.05354	SE	0.010263	1.82E-07	E
			rs5442	12:6954864	G/A	G		0.066921	126785	β	-0.09622	SE	0.019181	5.27E-07	E
			rs75448266	12:6798632	A/C	A		0.084049	125672	β	-0.0944	SE	0.017746	1.04E-07	E
			rs17083392	13:69026559	A/G	A		0.004286	126745	β	-1.60924	SE	0.3177	4.08E-07	E
			rs2019135	13:69034323	T/C	T		0.004267	126772	β	-1.76061	SE	0.323766	5.39E-08	E
			rs2019150	13:69034188	C/A	C		0.004262	126771	β	-1.76061	SE	0.323766	5.39E-08	E

			rs677755	13:69041438	T/C	T		0.00486	126760	β	1.426	SE	0.290601	9.24E-07	E
			rs74632168	13:69048954	C/T	C		0.004056	126772	β	-1.72062	SE	0.319014	6.91E-08	E
			rs2650494	16:28318440	A/G	A		0.414312	124286	β	-0.05259	SE	0.010223	2.68E-07	E
			rs325363	18:38331295	G/A	G		0.229965	126624	β	0.059212	SE	0.011689	4.07E-07	E
			rs4492854	11:11298353 4	C/T	C		0.4363	30366	β	0.1113	SE	0.021251	1.63E-07	E
Sherva, 2016	DSM-IV cannabis dependence criteria count	GR Ch37	-	1:212183114:I			INTS7		1250					2.05E-07	AA
			rs14148222 8				INTS7		1250					**8.84E-10**	AA
			rs77349458				SNORA26, INTS7		1250					1.57E-07	AA
			rs77448142				RPS20P10, RPS20P10-CYP26B1		1250					2.13E-07	AA

			-	2:167239818:D			SCN9A-SCN7A		4750					7.43E-07	AA
			rs313542				PI4K2B		2640					**7.12E-10**	EA
			rs7689780				PI4K2B,PI4K2B-ZCCHC4		2640					4.63E-07	EA
			rs73323306				ARL2BPP5-RP11-541P9.3		4750					**7.61E-09**	AA
			rs7832545				CSMD1		2640					**3.61E-10**	EA
														3.02E-07*	
			rs77378271				CSMD1		2640					**5.30E-10**	EA
														6.69E-07*	
			rs75721860				CSMD1		2640					**2.75E-09**	EA
			rs7853028				PSMB7		1250					**8.94E-09**	AA
			rs4643011				HABP2		1250					7.21E-07	AA

			rs11647404 2				RP11-755E2 3.3-CCD C67		4750					3.81E-07	AA
			rs18916703 8				SNORD11 - RNU6-1014P		1250					**2.64E-09**	AA
			rs14216241 5				SNORD11 - RNU6-1014P		1250					**1.88E-08**	AA
			rs11704781 0				SNORD11 - RNU6-1014P		1250					5.68E-08	AA
			rs11466037				MEFV		1250					8.08E-07	AA
			rs18356857 8				MEFV		1250					5.62E-07	AA
							RP5-860P4		1250					2.13E-07	AA

			rs11426999 2				.2-CST7						1.63E-07*	
			rs11462052 9				RP5-860P4 .2-CST7		1250				2.15E-07 1.61E-07	AA
			rs11662908 4				RP5-860P4 .2-CST7		1250				1.70E-07 1.27E-07*	AA
			rs11538451 2				RP5-860P4 .2-CST7		1250				8.91E-08 7.19E-08*	AA
			rs14764166 2				CST7, RP5-860P4 .2-CST7		1250				8.87E-08 7.13E-08*	AA
			rs19178314 4				CST7, RP5-860P4 .2-CST7		1250				8.81E-08 7.07E-08*	AA
			rs14680633 8				CST7, RP5-860P4 .2-CST7		1250				8.80E-08 7.03E-08*	AA

			rs11482872				CST7,APMAP,CST7-APMAP		1250					9.01E-08 6.92E-08*	AA
			rs11412800 2				APMAP		1250					8.06E-08 6.33E-08*	AA
			rs11534271 1				RNU6-1257P,APMAP		1250					2.54E-07 1.15E-07*	AA
			-	20:24955422:I			RNU6-1257P,APMAP		1250					2.54E-07 1.15E-07*	AA
			rs11576444 0				APMAP		1250					2.33E-07 9.03E-08*	AA
			rs11606833 5				APMAP-ACSS1		1250					1.66E-07 5.62E-08*	AA

			rs11463714 2				ACSS1,APMAP-ACSS1		1250					9.15E-08	AA
			rs11407190 1				ACSS1		1250					1.06E-07	AA
			rs11323274 2				ACSS1		1250					8.82E-08	AA
			rs11419992 8				ACSS1		1250					8.42E-08	AA
			rs11483636 4				ACSS1		1250					8.23E-08	AA
			rs11666936 8				ACSS1		1250					5.08E-08	AA
			rs14537993 4				ACSS1		1250					**3.96E-08**	AA
			rs14302022 5	2:16721471 4	G/A	G	snoU13,SCN9A	0.95	4750					5.85E-07	AA
			-	chr22:26917069:I			C9.15,TPST2,CTA-445C		2640					**3.15E-10**	EA

							9.14-CTA-4								
			rs4149485				TPST2,CTA-445C9.15		2640					**2.24E-10**	EA
*Adjusted for the DSM-IV criteria counts for alcohol, cocaine, and opioid dependence AA=African American, EA=European American, E=European Bold type indicates variants for which the p value reached genome-wide significance ($p \leq 5 \times 10^{-8}$)															

Table 5. GRADE Assessment

№ of studies	Certainty assessment						Effect			Certainty	Importance
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	№ of events	№ of individuals	Rate (95% CI)		
Cannabis use in lifetime											
2	observational studies	not serious	not serious	not serious	not serious					-	IMPORTANT
Diagnosis of cannabis use disorder											
2	observational studies	not serious	not serious	serious [a]	not serious					-	IMPORTANT
Criterion count of cannabis use disorder											
2	observational studies	not serious	not serious	serious [a]	not serious		-			-	IMPORTANT
Age of onset of cannabis use											
1	observational studies	not serious	not serious	not serious	not serious		-			-	IMPORTANT

Explanations
a. Different diagnostic classification

4 CHAPTER 4

4.1 Genetics of cannabis use in opioid use disorder: A Genome-Wide Association Study and Polygenic Risk Score study

Authors & Affiliations
Alannah Hillmer[1], Caroul Chawar[2], Amel Lamri[3], Jacqueline Hudson[4], Flavio Kapczinski[5], Luciano Minuzzi[6], David C. Marsh[7], Lehana Thabane[8], Zainab Samaan[9]* & Andrew D. Paterson[10]*

1. Neuroscience Graduate Program; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; hillmea@mcmaster.ca
2. MEND Research Lab; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; caroulchawar@gmail.com
3. Population Health Research Institute, Hamilton, ON, Canada; Department of Health Research Method, Evidence, and Impact, McMaster University, Hamilton, ON, Canada; lamria@mcmaster.ca
4. Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; jhudson@mcmaster.ca
5. Professor Emeritus, McMaster University, Hamilton, ON, Canada; Faculty of Medical Sciences, Universidade Federal do Rio Grande do Sol, Porto Alegre, Brazil; flavio.kapczinski@gmail.com
6. Department of Psychiatry and Behavioural Neuroscience; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; minuzzi@mcmaster.ca
7. Canadian Addiction Treatment Centres; Northern Ontario School of Medicine; 935 Ramsey Lake Rd., Sudbury, ON, P3E 2C6; dmarsh@nosm.ca
8. Department of Health Research Method, Evidence & Impact; 1280 Main St. W., Hamilton, ON, L8S 4L8; thabanl@mcmaster.ca
9. Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; samaanz@mcmaster.ca
10. Program in Genetics and Genome Biology, The Hospital for Sick Children; Divisions of Biostatistics and Epidemiology, Dalla Lana School of Public Health; University of Toronto; 686 Bay Street, Toronto, Ontario Canada M5G 0A4; andrew.paterson@sickkids.ca

*Corresponding Authors:
Dr. Zainab Samaan (responsible for phenotype and study methods)
100 West 5th St., Hamilton, ON L8N3K7, Canada
Telephone: 905-522-1155 ext. 35448
Fax: 905-381-5629

Dr. Andrew Paterson (responsible for genotype analysis and methods)
686 Bay Street, Toronto, ON M5G0A4, Canada
Telephone: 416-813-6994
Fax: 416-813-2150

4.1.1 Abstract

Background: Trends in cannabis use rates have increased in the Canadian population from 14.0% to 16.7% between 2018-2019. Further, individuals living with an Opioid Use Disorder (OUD) have increased rates of cannabis use in comparison to the general population. Research on the short- and long-term impacts of cannabis use in OUD patients, including its use as a harm reduction strategy, has been inconclusive. A genetic component may contribute to cannabis dependence; however, replication of findings has been inconsistent.
Objectives: This study aims to explore novel genetic variants associated with cannabis use through Genome-wide Association Study (GWAS) methods and to investigate genetic variants previously associated with cannabis use in a novel sample through a Polygenic Risk Score (PRS).
Methods: Participants were recruited from the GENetics of Opioid Addiction (GENOA) and the Pharmacogenetics of Opioid Substitution Treatment (POST) studies. The study outcomes of interest are: regular cannabis use (yes/no) (n= 2616) and heaviness of cannabis use (product of number of days used within a 30-day period and typical amount in grams) (n= 1293). Logistic and linear regressions were preformed, respectively, to test the association between each outcome, regular cannabis use and heaviness of cannabis use. GWAS summary statistics from a recent large GWAS meta-analysis investigating cannabis use disorder were used to conduct a polygenic risk score.
Results: Results from the GWASs indicated no genome-wide significant associations, however, rs1813412 on chromosome 17 for regular cannabis use and rs62378502 on chromosome 5 for heaviness of cannabis use were approaching genome-wide significance. Both SNPs approaching genome-wide significance were found to be significant ($p<0.05$) within both males and females when analyzed separately, however sex did not modify the association. The PRS identified statistically significant results for the regular cannabis PRS at the p-value 0.05. PRS scores for both regular cannabis use and heaviness of use R^2 were less than 2.50E-03.
Conclusion: This study provides promising results in understanding the genetic contribution to cannabis use in individuals living with OUD. However, findings are limited to a European ancestry sample.

Keywords: genetics, cannabis, opioid, polygenic risk score, genome-wide analysis

4.1.2 Introduction

Background/Rationale

In Canada, trends in cannabis use have been increasing, where 14.0% of Canadians 15 years or older reported cannabis in the first quarter of 2018 compared to 16.7% in the last quarter of 2019 (1). Globally, over the past two decades cannabis use and dependence are estimated to have increased and Cannabis Use Disorder (CUD) is reported as one of the most common drug use disorders (2). Concerningly, cannabis use is associated with adverse events such as impaired ability to concentrate and react quickly, impaired memory, motor coordination and judgement, increased anxiety, fear or panic, as well as possible paranoid feelings and psychosis (3). Repeating use of cannabis can have long lasting effects including, but not limited to, social outcomes such as poor education outcomes, diminished life satisfaction and achievement, poor professional and social achievements, as well as physical outcomes such as altered brain development, cognitive impairment, symptoms of chronic bronchitis and increased risk of chronic psychologic disorders (3–5).

It has been suggested that a genetic component may exist to cannabis dependence, with heritability estimates varied from 30-48% for cannabis use initiation and from 51-59% for problematic cannabis use (6). Genes of interest include those that encode for the cannabinoid receptors CB1 (*CNR1)* and CB2 *(CNR2)*, however only limited evidence exists for *CNR1* role in cannabis dependence and *CNR2* has never been examined in relation to any substance-related behaviour other than alcohol related use, to our knowledge (7). Additional candidate genes of interest include *FAAH*, *MGLL*, *TRPV1*, *GPR55*, *GABRA2*, *DRD2*, *ANKK1*, and *ABCB1*, however, replication of these associations has been inconsistent (7,8). A recent systematic review identifying genetic variants from genome-wide association studies (GWASs) associated with cannabis found gene regions of interest in *INTS7*, *PI4K2B*, *CSMD1*, *CST7*, *ACSS1*, and *SCN9A* (9). Further, to date, five GWAS meta-analyses have been conducted identifying *CADM2*, *ATP2C2*, and *FOXP2* as gene regions of interest associated with cannabis use, and regions of interest on chromosome 10 (rs77300175) and chromosome 8 (rs4732724) associated with CUD (8,10–13). Several metabolites of cannabis have been identified and *FAAH* and *CYP2C9* are genes of interest in the metabolism of cannabis (14–17).

Individuals living with an Opioid Use Disorder (OUD) have higher rates of substance use than the general population(18). Cannabis use in particular is highly prevalent in individuals with OUD, at rates more than 50%, higher than that of the general Canadian population (16.7%). Despite the high prevalence of cannabis use in the OUD population, the short and long-term impacts of cannabis use in OUD are inconclusive, with some studies showing harm reduction as a substitute, some showing no association, and others identifying worse outcomes for patients receiving Medication-Assisted Treatment (MAT) (19–31). Being that cannabis is a widely used substance, with policies supporting the legalization and with a growing viewpoint with the potential substitution of cannabis for other drugs, it is important to not only understand the genetic factors associated with cannabis use but also how cannabis use effects vulnerable populations such as the OUD population.

We aim to examine if novel genetic variants are associated with cannabis use within the OUD population. Further, we aim to determine if the known genetic variants associated with cannabis use in other populations are associated with cannabis use within OUD patients receiving MAT through a Polygenic Risk Score (PRS).

4.1.2.1 Objectives

The study aims to identify novel genetic variants and test the association between genetic variants previously associated with cannabis use in a novel sample. The objectives of the study are to:

1. Conduct a GWAS to identify novel genetic variants
2. Investigate sex differences of any novel genetic variants approaching or reaching genome-wide significance
3. Conduct a PRS based on a large previous GWAS meta-analysis
4. Conduct an exploratory PRS using an outcome to approximate cannabis dependence

4.1.3 Methods

This study is reported in accordance to the Strengthening the Reporting of Genetic Association studies (STREGA) guidelines, an extension of Strengthening the Reporting of Observational studies in Epidemiology (STROBE) statement (32). An accompanying STREGA checklist can be found in Supplementary File 1.

4.1.3.1 Study Design and Setting

Data were collected as part of the Genetics of Opioid Addiction (GENOA) and Pharmacogenetics of Opioid Substitution Treatment Response (POST) programs, prospective cohort studies conducted in collaboration with the Canadian Addiction Treatment Centre (CATC) and McMaster University designed to identify factors associated with opioid use and treatment outcomes including genetic risk factors in patients diagnosed with OUD and receiving treatment. Participants (n_{GENOA}=1536, n_{POST}=2544) were recruited from 76 CATC sites across Ontario, Canada, from 2013 to 2016 and 2018 to present for GENOA and POST respectively. At study recruitment, participants completed an extensive interview with a trained researcher in which they completed the Maudsley Addition Profile (MAP) among other measures, were asked about their substance use and any related behaviour in the past 30 days and provided a DNA sample (33). Participants were followed for a 12-month period through their electronic medical record documenting weekly or biweekly routine urine drug screening and participants within the GENOA study completed follow up interview at 12 months. In addition, when participants were recruited at study entry, their drug urine screens for 3-months prior to study enrollment were collected from medical records. Therefore, in total, 15-months of urine drug screens were obtained from participants electronic medical records. Participants were given a small coffee shop gift card after every face-to-face interview in appreciation for participating.

The GENOA and POST studies were reviewed and approved by the Hamilton Integrated Research Ethics Board (GENOA: 11-056, POST: 4556).

4.1.3.2 Participants

Patients were eligible to participate in the GENOA study if they were 18 years or older and met the criteria for Diagnostic and Statistical Manual – fourth edition (DSM-IV) opioid addiction requiring treatment (later replaced in DSM-5 as Opioid Use Disorder (OUD)) (34). Patients were excluded if they did not speak English or refused to provide a blood sample (for DNA). Patients were eligible to participate in the POST study if they were 16 years or older and met the criteria for Diagnostic and Statistical Manual – fifth edition (DSM-5) Opioid Use Disorder (OUD). Patients were excluded if they did not speak English or refused to provide a saliva sample (for DNA). In addition to meeting eligibility for the GENOA or POST study criteria, further inclusion criteria for this study included completing the cannabis related questions on the MAP or self-reported current cannabis use and a viable DNA sample.

4.1.3.3 Outcomes and Quantitative variables

Outcomes measured in the study include the following:

1. Regular cannabis use, defined as self-reported cannabis use at least twice in a 30-day period.
2. Heaviness of cannabis use, defined as the product of number of days of cannabis use and the typical dose within a 30-day period (35).
3. Cannabis dependence, defined as the score obtained from the Marijuana Cravings Questionnaire – Short Form (MCQ-SF) (36).

Covariates for the measures of regular cannabis use, heaviness of cannabis and cannabis dependence that were accounted for in the statistical models included: sex, age in years, and principal components accounting for differences due to population stratification.

4.1.3.4 Data Sources/Measurement

Cannabis use was self-reported by participants at baseline in which they indicated how many days in the past 30 days they used cannabis, the amount used in a typical day and how it was taken (ie. orally, smoked) as per the MAP. Other substances were self-reported as with cannabis through use of the MAP and also by urine drug screens. Urine drug screens were done on average once a week. Participant's urine sample was tested based on reported or suspected drug use, analyzed and reported as the number of positive screens for the drug detected in the test using the FaStep Assay (Trimedic Supply Network Ltd, Concord, Ontario, Canada) (37). Unfortunately, urine drug testing for cannabis was not consistent in all the clinical sites, however previously we reported the validity of self-reported cannabis use versus objective urine drug screen (35). Sensitivity and specificity were determined using participants who had data for both urinalysis and MAP (n=349). The sensitivity was 79.9% (95% CI 72.7, 85.8) and specificity was 80.0%

(95% CI 73.6, 85.4). Sensitivity and specificity values did not significantly differ between men and women, and there were no significant differences between false negatives and false positives. Thus, self-reported cannabis use was deemed as an appropriate measure of cannabis use (35).

Regular cannabis was reported as a binary construct. Participants were reported as regularly using cannabis if they had self-reported cannabis use two or more days on the MAP. As previously stated, heaviness of cannabis use was determined by the number of days of cannabis use as well as the typical dose used within a 30-day period for self-reported cannabis users. To determine typical dose, we followed a protocol previously used in the literature and converted all amounts to grams of cannabis (35). In addition, participants who reported one or two "puffs of a joint" were given the equivalent grams as reported in the literature, assuming 10 puffs is equal to 0.5 grams of cannabis (38). The log of heaviness of cannabis use was used in all statistical analyses to approach a normal distribution.

Post-hoc, we conducted an exploratory analysis using the MCQ-SF to approximate cannabis dependence. Cannabis dependence was reported as a continuous variable with a range of 12-84 based on scores obtained by the MCQ-SF. The MCQ-SF was administered to POST research participants who had indicated past month cannabis use. The administration of the MCQ-SF has been previously described (39). Briefly, the MCQ-SF is a 12-item Questionnaire, scored on a 7-point Likert scale, which assess four components of cannabis cravings (compulsivity, emotional benefit, expectancy of positive outcomes through use and purposefulness of cannabis) (36). The MCQ-SF score was summed for all participants, and the log of the total score was used for the statistical analysis to approach a normal distribution.

4.1.3.5 Quality Control Checks

Blood samples collected as part of the GENOA study and saliva samples collected as part of the POST study at study recruitment and were genotyped by Génome Québec Innovation Centre using GenomeStudio and the Illumina Global Screening Array – 24 v1.0 (40–42). Further details on the DNA collection can be found in Supplementary File 2. Variant quality control procedures were applied using PLINK v1.90 (43). Additionally, R version 3.3.3 was used for quality control checks (44).

Samples were excluded if they have a low variant call rate (<99%), inconsistencies between self-reported vs. genetically determined sex or ancestry or if they exhibit excess heterozygosity suggestive of sample contamination (exchange of DNA between two or more samples). Samples were also excluded if they were a duplicate and a heritability analysis was performed to exclude first-degree and second-degree relatives. Variants were excluded if they have a low call rate across samples (<99%), or if the minor allele was less than 0.05.

Due to the predominantly European ancestry, only data from samples of European decent (n=2958) were selected for imputation. Imputation was completed by TOPMed Imputation Server (version R2) using the software Eagle v2.4 and Minimac4 for phasing and Single Nucleotide Polymorphism (SNP) imputation, respectively, with the reference

panel of TOPMed r2 (45–48). Post-imputation filtering excluded SNPs with Rsq quality metrics of less than 0.3 and minor allele frequencies lower than 0.05. A detailed description of the steps taken is reported and available in Supplementary File 2.

4.1.3.6 Bias

While measures were taken to identify and mitigate potential bias, possible sources of bias inevitably remained. Although sensitivity and specificity analyses were conducted on self-reported cannabis use behaviour, cannabis was an illegal substance at the earlier time of the study (cannabis was legalized in Canada as of October 2018), potentially leading to social desirability and recall biases. Participants might have provided incorrect information to be viewed more favourable or could have had difficulty recalling the duration and amount of cannabis used within 30 days. Further, the findings might have been affected by volunteer bias, not only for participating in a research study but more specifically in a genetic study, where the sample may not be representative of the entire OUD population receiving treatment. Additionally, the study results may not be generalizable outside of the European ancestry and may not be replicable in other populations. With respect to measuring cannabis use, bias exists as cannabis includes different components, mainly detla-9-tetrahydrocannabinol (THC) and cannabidiol (CBD), amounts (referring to the amount of THC and CBD included in various strains of cannabis as well as physically measuring of cannabis) and methods of consumption which are imprecisely estimated as no standard currently exists (38,49). Finally, due to the observational nature of this study, it is not possible to control for all variables or for extraneous confounding variables.

4.1.3.7 Sample Size

Of the total 4990 participants enrolled in the GENOA and POST studies, 4621 genetic samples were available. As only participants from European ancestry were included, 2625 participants of European ancestry passed genetic quality control steps and were used for this study. A detailed chart outlining the steps conducted to reach the final sample size is reported in the Supplementary File 2.

4.1.3.8 Statistical methods

Descriptive statistics were conducted on the total samples to describe the demographic and clinical characteristics, with continuous variables expressed as means with standard deviations, while categorical variables are expressed as counts.

Separate regression analyses were performed to test the association between genetic variants and the outcomes of interest. Logistic regressions were conducted to measure the association of regular cannabis use and linear regressions were used for heaviness of cannabis use and cannabis dependence. All aforementioned covariates were adjusted for by using additive models for genotype coding. Further, identical regression analyses were conducted separately for male and female subsets as well as an interaction model for SNPs approaching or meeting genome-wide significance.

The statistical software PRSice-2 was used to conduct the PRS to investigate the polygenic risk of cannabis use (50). Discovery GWAS results were obtained from a large meta-GWAS conducted on CUD which included summary statistics from the Psychiatric Genomics Consortium Substance Use Disorders group, the iPSYCH sample and deCODE sample (13). The UCSC*liftOver* tool was used to convert the GWAS summary statistics from GRCh build 37 to build 38 in order to match the target data (51). Discovery GWAS results were pruned using PRSice-2 using an R^2 threshold of 0.5 and within 250kb. Subset of SNPs were selected at decreasingly liberal P-value thresholds (1, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.001, 0.0001, $1x10^{-5}$ and $5x10^{-8}$). PRS were calculated separately for regular cannabis use, heaviness of cannabis use and cannabis dependence, with all aforementioned covariates adjusted for.

All statistical analyses were performed on PLINK v1.09, R studio 3.3.3 and PRSice-2 (43,44,50,52).

4.1.4 Results

4.1.4.1 Participants

Of the 2625 participants whose samples passed the genetic quality control checks, 9 participants with genetic data had missing or incomplete self-reported cannabis use data as collected on the MAP and were excluded from the current study, therefore 2616 participants were included. For the outcome of heaviness of cannabis use, participants were required to self-report at least one day of cannabis use in the past 30 days and report the average number of grams used. Of the 2616 participants, 1321 participants completed the MAP and reported at least 1 day of cannabis use. However, 28 participants did not report the average number of grams of cannabis or were not able to convert their reported usage in grams (e.g. reported the use of cannabis oil), therefore 1293 participants were included in the outcome of heaviness of use. Finally, for the outcome cannabis dependence, of the 1321 participants who reported at least 1 day of cannabis use, 485 individuals did not complete the MCQ-SF and therefore 836 individuals were included in the cannabis dependence analysis. The total number of participants included for each outcome were 2616 for regular cannabis use, 1293 for heaviness of cannabis use and 836 for cannabis dependence. A flow chart is available in Supplementary File 2, Figure S4.

4.1.4.2 Descriptive Data

For the current study, 965 participants from the GENOA study and 1,651 from the POST study were included. Slightly more participants were male than female (57.45%) with a mean age of 38.23. Majority of participants had never been married, are unemployed and completed some level of high school. Nearly half of the sample reported regular cannabis use (47.94%), of those who reported cannabis use a mean of 21.30 days out of 30 with and an average of 1.36 grams per day was reported. Cannabis users reported an average score of 37.17 on the MCQ-SF. Finally, 6,377,206 SNPs passed the quality control steps and were included in the GWAS, 381,569 SNPs were included in the

regular cannabis use PRS, 381,358 SNPs in the heaviness of use PRS and 380,998 SNPs in the cannabis dependence PRS. Participants demographics can be found in Table 1.

Table 1. Sample demographics

	Total	Male	Female
N (%)	2616	1503 (57.45)	1113 (42.55)
GENOA	965	560 (58.03)	405 (41.97)
POST	1651	943 (57.12)	708 (42.88)
Age in years, Mean (SD)	38.23 (11.08)	39.89 (11.17)	38.35 (10.90)
Marital status, N (%)			
Never Married	1251 (47.83)	761 (50.63)	490 (44.03)
Common law	539 (20.60)	275 (18.30)	264 (23.72)
Currently Married	258 (9.86)	156 (10.38)	102 (9.16)
Separated	241 (9.21)	118 (7.85)	123 (11.05)
Divorced	251 (9.59)	157 (10.45)	94 (8.45)
Widowed	76 (2.91)	36 (2.40)	40 (3.59)
Currently employed, N (%)	931 (35.59)	631 (41.98)	300 (26.95)
Education, N[a] (%)			
Less than grade 9	574 (21.97)	326 (21.69)	248 (22.34)
Grade 9-12	1212 (46.38)	769 (51.16)	443 (39.91)
Trade School	87 (3.33)	59 (4.56)	28 (2.52)
College/University/Graduate School	722 (27.63)	336 (22.36)	386 (34.77)
Regular use of cannabis, N (%)	1254 (47.94)	782 (52.03)	472 (42.41)
Day's cannabis used in last 30, Mean (SD)[b]	21.30 (11.59)	21.50 (11.56)	21.05 (11.65)
Average cannabis dose in g/day, Mean (SD)[b]	1.36 (2.27)	1.41 (1.69)	1.27 (2.98)
Average MCQ-SF total score, Mean (SD)[c]	37.17 (16.18)	37.16 (15.87)	37.18 (16.67)
[a]Data available for n_{Total}= 2,613, n_{Male}= 1503, n_{Female}= 1,110 [b]Data available for n_{Total}= 1,293, n_{Male}= 802, n_{Female}= 491 [c]Data available for n_{Total}= 836, n_{Male}= 507, n_{Female}= 329			

4.1.4.3 Main results

We identified two SNPs in our GWASs approaching genome-wide significance; for regular cannabis use rs1813412 on chromosome 17 ($p=2.05x10^{-7}$) and for heaviness of use rs62378502 on chromosome 5 ($p=5.56x10^{-7}$). For both regular and heaviness of cannabis use, using the top SNP as a covariate in the analysis did not result in any further

significant results, thus no other strong signals within the region exist. Results from the top SNPs of each respective GWAS can be found in Table 2.

Table 2. Top SNPs from each GWAS

Outcome	Chr	SNP	BP	A1	OR/BETA	P
Regular Cannabis Use	17	rs1813412	22193901	G	1.352	2.05×10^{-7}
Heaviness of Cannabis Use	5	rs62378502	168815119	A	0.189	$5.56\text{x}10^{-7}$

Chr=chromosome, SNP=single nucleotide polymorphism, BP=base pair, A1=reference allele, OR=odds ratio, BETA= beta coefficient

Results from the sex-stratified association analyses between the SNPs approaching GWAS significance for regular cannabis use and heaviness of cannabis use are reported in Table 3 and interpreted with the significance threshold of $p<0.5$. The G allele of rs1813412 was significantly associated with an increased odds of regular cannabis use in both males [odds ratio (OR)=1.31, 95% confidence interval (CI)=1.13, 1.52, p=4.32E-04] and females [OR= 1.47, 95% CI=1.23, 1.76, p=2.33E-05], however the sex by SNP interaction was not significant. The C allele of rs62378502 was significantly associated with heaviness of cannabis use in both males [Beta=0.19, standard error (SE)=0.05, p=6.59e-05] and females [Beta=0.19, SE=0.07, p=3.44E-03], however, as with regular cannabis use, the sex by SNP interaction was not significant.

Table 3. SNPs and associated outcomes stratified by sex

Outcome	SNP	N	Reference Allele	OR/BETA	95 % CI/SE	P
Regular cannabis use	**rs1813412**		G			
	Males	1497		1.31	1.13, 1.52	4.32E-4
	Females	1109		1.47	1.23, 1.76	2.33E-5
	Interaction	2616		1.15	0.92, 1.46	0.23
Heaviness of cannabis use	**rs62378502**		C			
	Males	799		0.19	0.05	6.59E-5
	Females	490		0.19	0.07	3.44E-3
	Interaction	1293		0.01	0.08	0.92

OR=odds ratio, BETA= beta coefficient, 95% CI = 95% confidence interval levels (lower, upper), SE=standard error

Results from the PRS include the best-fit model and the PRS model fit across p-value thresholds observed in the previously reported GWAS for regular cannabis use and heaviness of cannabis use are reported in Table 4 and Table 5, respectively. Two PRS reached significance in the regular cannabis use outcome, and none reached significance in heaviness of use. Using the p-value threshold of 0.0001 and 0.001 for the base data, a significant PRS was found for regular cannabis use with an R^2 value of $2.38E^{-3}$ and $2.28E^{-3}$ respectively. A bar plot depicting the model of fit for the regular cannabis use and heaviness of cannabis use can be seen in Figure 5 and 6, respectively.

Table 4. PRS best model fit for each outcome

Outcome	Threshold	PRS R^2	Full R^2	Null R^2	Coefficient	SE	P	Number of SNPs
Regular cannabis use	0.0001	2.50E-3	4.82E-2	4.58E-2	38.58	17.56	**0.028**	204
Heaviness of cannabis use	1.0 E-5	9.19E-4	1.51E-2	1.42E-2	-3.71	3.41	0.277	45

PRS R^2=Variance explained by the PRS, Full R^2=Variance explained by the full model, Null R^2=Variance explained by the covariates, SE=standard error, Number of SNPs=number of SNPs included in the model

Table 6. PRS model fit across thresholds for each outcome

Outcome	Threshold	R^2	Coefficient	SE	P	Number of SNPs
Regular cannabis use	$5.0E^{-8}$	5.11E-04	1.63	1.60	0.308	2
	$1.0E^{-5}$	3.82E-04	7.05	8.00	0.379	45
	0.0001	2.38E-03	38.58	17.56	**0.028**	204
	0.001	2.28E-03	104.55	48.69	**0.032**	1280
	0.05	4.20E-04	277.29	300.29	0.356	36660
	0.1	7.68E-04	520.81	417.11	0.212	66179
	0.2	2.67E-04	436.65	592.98	0.462	118682
	0.3	7.22E-05	285.56	745.90	0.702	166435
	0.4	1.79E-06	53.68	889.52	0.952	209778
	0.5	3.00E-05	252.46	1023.25	0.805	249212
	1	2.54E-05	344.23	1514.88	0.820	381569
Heaviness of cannabis use	$5.0E^{-8}$	4.14E-04	0.52	0.70	0.462	2
	$1.0E^{-5}$	9.06E-04	-3.71	3.41	0.277	45
	0.0001	7.30E-04	-7.30	7.48	0.329	203
	0.001	6.35E-07	-0.60	20.80	0.977	1280

	0.05	8.16E-05	42.08	128.96	0.744	36662
	0.1	5.51E-05	48.25	180.01	0.789	66241
	0.2	5.42E-05	67.87	255.24	0.790	118664
	0.3	3.96E-05	73.07	321.35	0.820	166396
	0.4	4.59E-05	-94.13	384.80	0.807	209735
	0.5	3.44E-05	-93.98	443.35	0.832	249091
	1	4.64E-05	-161.91	658.24	0.806	381358

R^2=Variance explained by the PRS, SE=standard error, Number of SNPs=number of SNPs included in the model

4.1.4.4 Exploratory Analysis

Post-hoc we decided to conduct an exploratory PRS with a new outcome using the MCQ-SF to approximate cannabis dependence (36). The results from the exploratory PRS can be found below in Table 6 and Table 7, however no PRS was found to be significantly associated with cannabis dependence.

Table 7. PRS best model fit for cannabis dependence

Outcome	Threshold	PRS R^2	Full R^2	Null R^2	Coefficient	SE	P	Number of SNPs
MCQ-SF	$5.0E^{-8}$	1.37E-03	1.56E-02	1.42E-02	0.28	0.27	0.287	2

PRS R^2=Variance explained by the PRS, Full R^2=Variance explained by the full model, Null R^2=Variance explained by the covariates, SE=standard error, Number of SNPs=number of SNPs included in the model

Table 8. PRS model fit across thresholds for cannabis dependence

Outcome	Threshold	R^2	Coefficient	SE	P	Number of SNPs
MCQ-SF	$5.0E^{-8}$	1.35E-03	0.29	0.266	0.287	2
	$1.0E^{-5}$	4.01E-05	0.24	1.32	0.854	45
	0.0001	2.89E-04	1.39	2.82	0.622	203
	0.001	1.94E-04	3.16	7.83	0.687	1285
	0.05	2.15E-04	20.51	48.26	0.671	36679
	0.1	8.22E-05	17.62	67.01	0.793	66169
	0.2	1.30E-07	-1.00	95.76	0.992	118625
	0.3	4.23E-05	22.65	120.10	0.850	166251
	0.4	9.06E-05	-39.87	144.42	0.783	209574
	0.5	1.31E-04	-55.21	166.40	0.740	248834
	1	8.07E-05	-64.41	247.17	0.794	380998

R^2=Variance explained by the PRS, SE=standard error, Number of SNPs=number of SNPs included in the model

4.1.5 Discussion

4.1.5.1 Key results

For the GWASs, rs1813412 on chromosome 17 for regular cannabis use and rs62378502 on chromosome 5 for heaviness of cannabis use were approaching genome-wide significance. Further, both rs1813412 and rs62378502 were found to be significantly associated with the outcome of regular cannabis use and heaviness of cannabis use, respectively, in males and in females separately, however sex did not modify the association between the SNP and the outcomes of the study. For the PRS, statistically significant results were observed for the regular cannabis PRS at the p-value threshold of 0.0001 and 0.001. The exploratory PRS did not result in any significant results for cannabis dependence or increased variance explained by genetic contribution of cannabis dependence.

4.1.5.2 Interpretation

This GWAS did not replicate any previous known genetic associations of cannabis use from the literature. Additionally, rs1813412 on chromosome 17 and rs62378502 on chromosome 5 have no know associations with other traits or pathways within the literature. It is important to note that cannabis use and CUD have been reported to be modestly genetically correlated and show divergent genetic relationships with educational attainment, BMI and age at birth of first child (13). Thus, it is plausible that the lack of replication is due to majority of known GWASs reported in the literature which investigated CUD, a narrow phenotype compared to cannabis use without the associated features of a disorder (9). Additionally, within our exploratory PRS using the MCQ-SF to approximate cannabis dependence, we did not find an increased amount of genetic variance explained, which is in contrast with our predictions, therefore suggesting that within our study, even when approximating cannabis dependence, potential confounding variables exist leading to a lack of replication in our GWAS's.

While pervious literature on the genetics of cannabis use have identified several SNPs and genes of interest, growing evidence in genetic studies suggest that the majority of genetic variants have small effects which collectively contribute to the risk of certain diseases (9,53–55). PRS provide a method to aggregate the effects of variants across the genome by generating a weighted sum of the number of risk alleles an individual carries (54,56). We found a statistically significant PRS for regular cannabis use, suggesting this score can be applied to assess individual genetic risk of cannabis use. However, the variance explained by genetic variants was minor, less than 2.50E-03, therefore suggesting that the majority of variance contributing to regular cannabis use and heaviness of cannabis use is due to other, non-genetic, factors (i.e. environmental). From a clinical stance, as genetics are not as modifiable as factors such as environment, there is a greater possibility for changing the modifying risk factors in those who currently use cannabis. Additionally, it is possible that so little of the variance is explained by the PRS is due to the high degree of polysubstance use within this population, suggesting that

genetic factors may be contributing to addiction more broadly rather than to a single substance, such as cannabis.

It is important to recognize that this sample includes individuals currently undergoing treatment for an OUD. Individuals who live with an OUD are likely to have used multiple substances and it is relatively common for individuals receiving addiction treatment to engage in continued substance use and polysubstance use while on treatment (18,57). Further, in our population it has been previously found that cannabis use is associated with continued substance use while on treatment (58). However, the evidence surrounding the association between cannabis and illicit opioid use is mixed, with one study reporting cannabis users being less likely to be using heroin and another reporting cannabis use being significantly associated with illicit opioid use in women (35,58). Therefore, it is possible that the lack of significant findings is due to a shared genetic contribution among various substances, commonalities of addictive substances such as being dopamine-agonists or the high degree of polysubstance use within this population (59).

Finally, this study was unique in stratifying analyses by sex. Previous studies on cannabis use suggest that neurodevelopmental, pharmacological, metabolic, behavioural, and hormonal differences all contribute to sex differences in cannabis use (60,61). Compared to men, women show a greater sensitivity to the subjective effects of cannabis, have faster trajectories to CUD and worse mental health outcomes, such as increased anxiety and a higher risk of early onset psychosis (62–64). Given that animal and clinical studies suggest sex-specific differences, it is important to investigate how sex-specific genetic differences are associated with cannabis use outcomes (64). The current study did not find sex-specific associations for both top SNPs identified in the GWASs, rs1813412 and rs62378502. Nonetheless, investigating sex-specific differences at the genetic level can aid in understanding sex-specific differences at the neurodevelopment, pharmacological, metabolic, and hormonal levels. Although no sex-specific differences were found, it is important to continue to explore sex analyses within the field of genetics to understand sex-specific differences within addiction for clinical implications, ensuring that if underlying genetic differences exist within patterns of substance use, sex-specific treatment decisions are being developed to provide personalized care.

4.1.5.3 Limitations and generalizability

The current study has several limitations. This study utilized self-report data on cannabis use introducing potential reporting bias. Additionally, missingness within the data in respect to the measure of heaviness of cannabis use, resulting in a smaller sample size for this outcome. Second, due to limited genetic data on individuals of various ancestral backgrounds, the current findings are limited to those of European ancestry. Third, as no other studies to our knowledge have reported on results by sex, they are not comparable to other study findings. Fourth, as our study populations did not screen for CUD we used an approximation of cannabis dependence. Additionally, the GWASs for regular cannabis use and heaviness of cannabis use were both under-powered. Thus, the results of both GWASs and further investigation of sex-specific differences for the top

SNPs should be interpreted with caution. Finally, it is important to note that participation bias in genetic studies exist. While our study collected genetic data at study recruitment mitigating the loss to follow, a bias may exist wherein potential genetic differences in willingness to participate in genetic studies is overrepresented (65).

Despite these limitations, the current study provides additional evidence for the genetic liability of cannabis use within a high risk population with polysubstance addiction. While further research is required on the genetic susceptibility to cannabis use within the OUD population, the current study identifies the need for investigation of the contribution of sex to cannabis use.

4.1.6 Other Information

4.1.6.1 Funding

This study was supported by CIHR (PJT-156306), which played no role in the study design, analysis, reporting or publication.

4.1.6.2 Authors Contributions

ZS is the guarantor. AH, ZS and AP conceptualized the genetic study. AH implemented the design, quality control steps, and methodology with the aid of CC, AL, ZS and AP. AP provided genetic statistical guidance and data clean up methods. DM and JH provided support and collaboration with the CATC clinics. AH prepared the first draft. FK, LM, LT, ZS and AP revised the study design, data analysis, and data interpretation. All authors reviewed and approved the final manuscript.

4.1.7 List of Abbreviations

Canadian Addition Treatment Centre (CATC)
Cannabidiol (CBD)
Cannabis Use Disorder (CUD)
Confidence Interval (CI)
Detla-9-tetrahydrocannabinol (THC)
Diagnostic and Statistical Manual – fourth edition (DSM-IV)
Diagnostic and Statistical Manual – fifth edition (DSM-5)
Genetics of Opioid Addiction (GENOA)
Genome-Wide Association Study (GWAS)
Marijuana Craving Questionnaire-Short form (MCQ-SF)
Maudsley Addition Profile (MAP)
Medication-Assisted Treatment (MAT)
Odds Ratio (OR)
Opioid Use Disorder (OUD)
Pharmacogenetics of Opioid Substitution Treatment Response (POST)
Polygenic Risk Score (PRS)
Single Nucleotide Polymorphism (SNP)
Standard Error (SE)

Strengthening the Reporting of Genetic Association studies (STREGA)
Strengthening the Reporting of Observational studies in Epidemiology (STROBE)

4.1.8 References

1. The Government of Canada. Cannabis Stats Hub. 2019.
2. Degenhardt L, Charlson F, Ferrari A, Santomauro D, Erskine H, Mantilla-Herrara A, et al. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Psychiatry. 2018;5(12):987–1012.
3. Volkow ND, Baler RD, Compton WM, Weiss SRB. Adverse Health Effects of Marijuana Use. N Engl J Med. 2014;370(23):2219–27.
4. The Government of Canada. Cannabis Health Effects. 2019.
5. Bahji A, Stephenson C. International perspectives on the implications of cannabis legalization: a systematic review & thematic analysis. Int J Environ Res Public Health. 2019;16(17):3095.
6. Verweij KJH, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG, et al. Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. Addiction. 2010;105(3):417–30.
7. Agrawal A, Lynskey MT. Candidate Genes for Cannabis Use Disorders: Findings, Challenges and Directions. Addiction. 2009;104(4):518–32.
8. Stringer S, Minică CC, Verweij KJH, Mbarek H, Bernard M, Derringer J, et al. Genome-wide association study of lifetime cannabis use based on a large meta-analytic sample of 32 330 subjects from the International Cannabis Consortium. Transl Psychiatry. 2016;6(3):e769.
9. Hillmer A, Chawar C, Sanger S, D'Elia A, Butt M, Kapoor R, et al. Genetic basis of cannabis use: a systematic review. BMC Med Genomics. 2021;14(1):1–21.
10. Agrawal A, Chou YL, Carey CE, Baranger DAA, Zhang B, Sherva R, et al. Genome-wide association study identifies a novel locus for cannabis dependence. Mol Psychiatry. 2018;23(5):1293–302.
11. Minică CC, Verweij KJH, Most PJ, Mbarek H, Bernard M, Eijk KR, et al. Genome-wide association meta-analysis of age at first cannabis use. Addiction [Internet]. 2018 Nov;113(11):2073–86. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=132211589&site=ehost-live
12. Pasman JA, Verweij KJH, Gerring Z, Stringer S, Sanchez-Roige S, Treur JL, et al. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. Nat Neurosci [Internet]. 2018;21(9):1161–70. Available from: http://dx.doi.org/10.1038/s41593-018-0206-1
13. Johnson EC, Demontis D, Thorgeirsson TE, Walters RK, Polimanti R, Hatoum AS, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. The Lancet Psychiatry. 2020;7(12):1032–45.
14. Chayasirisobhon S. Mechanisms of action and pharmacokinetics of cannabis. Perm J. 2021;25.
15. Conzelmann A, Reif A, Jacob C, Weyers P, Lesch K-P, Lutz B, et al. A polymorphism in the gene of the endocannabinoid-degrading enzyme FAAH

(FAAH C385A) is associated with emotional–motivational reactivity. Psychopharmacology (Berl). 2012;224(4):573–9.
16. Ahn K, McKinney MK, Cravatt BF. Enzymatic pathways that regulate endocannabinoid signaling in the nervous system. Chem Rev. 2008;108(5):1687–707.
17. Gasse A, Vennemann M, Köhler H, Schürenkamp J. Toxicogenetic analysis of Δ9-THC-metabolizing enzymes. Int J Legal Med. 2020;134(6):2095–103.
18. Compton WM, Valentino RJ, DuPont RL. Polysubstance use in the US opioid crisis. Mol Psychiatry. 2021;26(1):41–50.
19. Reiman A. Cannabis as a substitute for alcohol and other drugs. Harm Reduct J. 2009;6(1):35.
20. Lucas P, Walsh Z, Crosby K, Callaway R, Belle-Isle L, Kay R, et al. Substituting cannabis for prescription drugs, alcohol and other substances among medical cannabis patients: The impact of contextual factors. Drug Alcohol Rev. 2016;35(3):326–33.
21. Ashrafioun L, Bohnert KM, Jannausch M, Ilgen MA. Characteristics of substance use disorder treatment patients using medical cannabis for pain. Addict Behav. 2015;42:185–8.
22. Wasserman DA, Weinstein MG, Havassy BE, Hall SM. Factors associated with lapses to heroin use during methadone maintenance. Drug Alcohol Depend. 1998;52(3):183–92.
23. Proctor S, Copeland A, Kopak A, Hoffmann N, Herschman P, Polukhina N. Outcome predictors for patients receiving methadone maintenance treatment: findings from a retrospective multi-site study. J Substamce Use. 2016;21(6):601–13.
24. Lucas P, Baron EP, Jikomes N. Medical cannabis patterns of use and substitution for opioids & other pharmaceutical drugs, alcohol, tobacco, and illicit substances; results from a cross-sectional survey of authorized patients. Harm Reduct J. 2019;16(1):9.
25. Lau N, Sales P, Averill S, Murphy F, Sato S, Murphy S. A safer alternative: Cannabis substitution as harm reduction. Drug Alcohol Rev. 2015;34(6):654–9.
26. Epstein DH, Preston KL. Does cannabis use predict poor outcome for heroin-dependent patients on maintenance treatment? Past findings and more evidence against. Addiction. 2003;98(3):269–79.
27. Peirce, J Mp, N Mr, J Mk, KKrasnansky, JStabile, et al. Correlates of stimulant treatment outcome across treatment modalities. Am J Drug Alcohol Abuse [Internet]. 2009;35(1 CC-Drugs and Alcohol):48-53. Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01621014/full
28. Cichewicz DL. Synergistic interactions between cannabinoid and opioid analgesics. Life Sci. 2004;74(11):1317–24.
29. Saxon AJ, Calsyn DA, Greenberg D, Blaes P, Haver VM, Stanton V. Urine screening for marijuana among methadone-maintained patients. Am J Addict. 1993;2(3):207–11.
30. Saxon AJ, Wells EA, Fleming C, Jackson TR, Calsyn DA. Pre-treatment

characteristics, program philosophy and level of ancillary services as predictors of methadone maintenance treatment outcome. Addiction. 1996;91(8):1197–210.
31. Nirenberg TD, Cellucci T, Liepman MR, Swift RM, Sirota AD. Cannabis versus other illicit drug use among methadone maintenance patients. Psychol Addict Behav. 1996;10(4):222.
32. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):p.e296.
33. Marsden J, Gossop M, Stewart D, Best D, Farrell M, Lehmann P, et al. The Maudsley Addiction Profile (MAP): a brief instrument for assessing treatment outcome. Addiction. 2002;93(12).
34. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychaitric Publishing; 2013.
35. Zielinski L, Bhatt M, Sanger N, Plater C, Worster A, Varenbut M, et al. Association between cannabis use and methadone maintenance treatment outcomes: an investigation into sex differences. Biol Sex Differ. 2017;8(1):8.
36. Heishman SJ, Evans RJ, Singleton EG, Levin KH, Copersino ML, Gorelick DA. Reliability and validity of a short form of the Marijuana Craving Questionnaire. Drug Alcohol Depend. 2009;102(1–3):35–40.
37. Trimedic Supply Network Ltd. FaStep Assay [Internet]. Available from: https://www.trimedic-inc.com/wp-content/uploads/2018/04/Fastep-Package-Insert.pdf
38. Zeisser C, Thompson K, Stockwell T, Duff C, Chow C, Vallance K, et al. A 'standard joint'? The role of quantity in predicting cannabis-related problems. Addict Res Theory. 2012;20(1):82–92.
39. Rosic T, Kapoor R, Panesar B, Naji L, Chai DB, Sanger N, et al. The association between cannabis use and outcome in pharmacological treatment for opioid use disorder. Harm Reduct J. 2021;18(1):1–9.
40. Centre d'expertise et de services Génome Québec [Internet]. Available from: https://cesgq.com/
41. illumina. GenomeStudio [Internet]. 2022. Available from: https://support.illumina.com/array/array_software/genomestudio/downloads.html
42. illumina. Infinium Global Screening Array - 24 v1.0 [Internet]. 2017. Available from: https://grcf.jhmi.edu/wp-content/uploads/2017/12/infinium-commercial-gsa-data-sheet-370-2016-016.pdf
43. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
44. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
45. Loh P-R, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48(11):1443.

46. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284–7.
47. TOPMed. TOPMed Imputation Server - Reference Panels [Internet]. Available from: https://topmedimpute.readthedocs.io/en/latest/reference-panels.html
48. Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature. 2021;590(7845):290–9.
49. Freeman AM, Petrilli K, Lees R, Hindocha C, Mokrysz C, Curran HV, et al. How does cannabidiol (CBD) influence the acute effects of delta-9-tetrahydrocannabinol (THC) in humans? A systematic review. Neurosci Biobehav Rev. 2019;107:696–712.
50. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience. 2019;8(7):giz082.
51. Kuhn RM, Haussler D, Kent WJ. The UCSC genome browser and associated tools. Brief Bioinform. 2013;14(2):144–61.
52. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4(1):s13742-015.
53. Johnson EC, Tillman R, Aliev F, Meyers JL, Salvatore JE, Anokhin AP, et al. Exploring the relationship between polygenic risk for cannabis use, peer cannabis use and the longitudinal course of cannabis involvement. Addiction [Internet]. 2019 Apr;114(4):687–97. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=135199548&site=ehost-live
54. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. 2020;15(9):2759–72.
55. Khera A V, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet. 2018;50(9):1219–24.
56. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. Genome Med. 2020;12(1):1–11.
57. Arnold TD, Lin L, Cotton BP, Bryson WC, Polenick CA. Gender Differences in Patterns and Correlates of Continued Substance Use among Patients in Methadone Maintenance Treatment. Subst Use Misuse. 2021;56(4):529–38.
58. Shams I, Sanger N, Bhatt M, Rosic T, Luo C, Shahid H, et al. The association between health conditions and cannabis use in patients with opioid use disorder receiving methadone maintenance treatment. BJPsych open. 2019;5(6).
59. Popescu A, Marian M, Drăgoi AM, Costea R-V. Understanding the genetics and neurobiological pathways behind addiction. Exp Ther Med. 2021;21(5):1–10.
60. Crane NA, Schuster RM, Fusar-Poli P, Gonzalez R. Effects of cannabis on neurocognitive functioning: recent advances, neurodevelopmental influences, and sex differences. Neuropsychol Rev [Internet]. 2012/11/06. 2013 Jun;23(2):117–37. Available from: https://pubmed.ncbi.nlm.nih.gov/23129391

61. Greaves L, Hemsing N. Sex and gender interactions on the use and impact of recreational cannabis. Int J Environ Res Public Health. 2020;17(2):509.
62. Cooper ZD, Haney M. Investigation of sex-dependent effects of cannabis in daily cannabis smokers. Drug Alcohol Depend. 2014;136:85–91.
63. Van Gastel WA, MacCabe JH, Schubart CD, Van Otterdijk E, Kahn RS, Boks MPM. Cannabis use is a better indicator of poor mental health in women than in men: a cross-sectional study in young adults from the general population. Community Ment Health J. 2014;50(7):823–30.
64. Nia AB, Mann C, Kaur H, Ranganathan M. Cannabis Use: Neurobiological, Behavioral, and Sex/Gender Considerations. Curr Behav Neurosci reports [Internet]. 2018/11/10. 2018 Dec;5(4):271–80. Available from: https://pubmed.ncbi.nlm.nih.gov/31548918
65. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2018;47(1):226–35.

4.1.9 Figures

Figure 1. Regular cannabis use Manhattan Plot

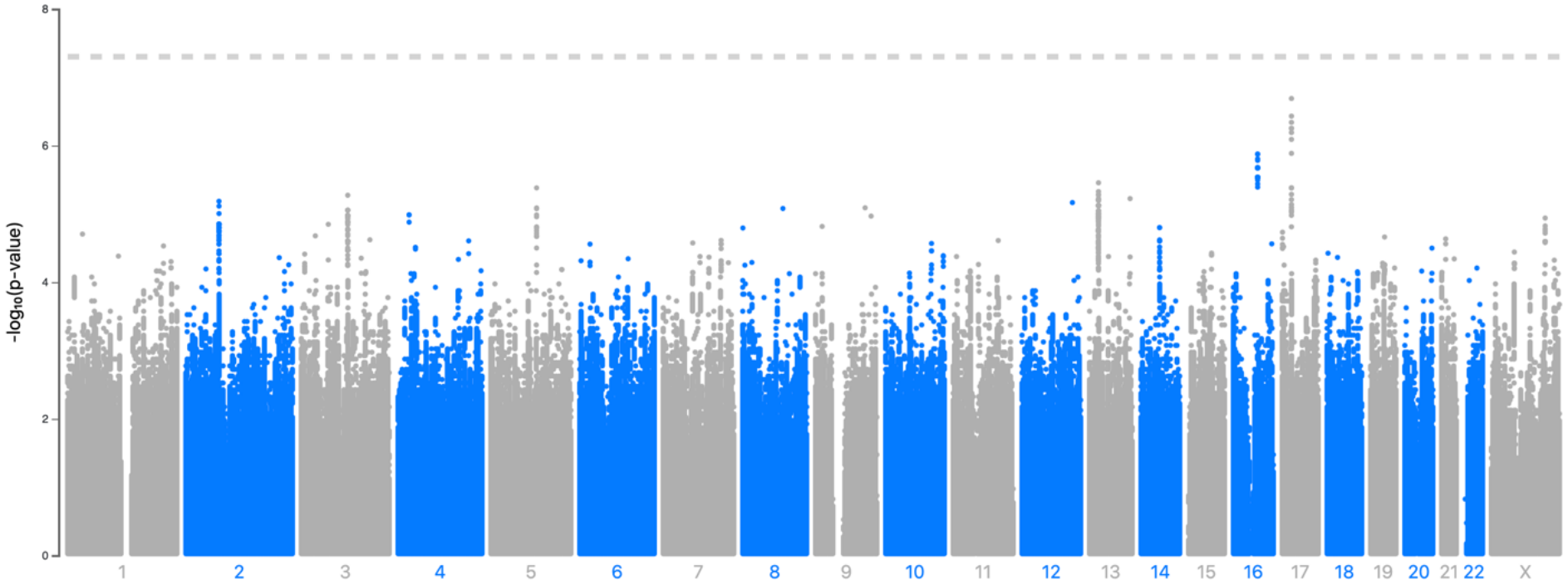
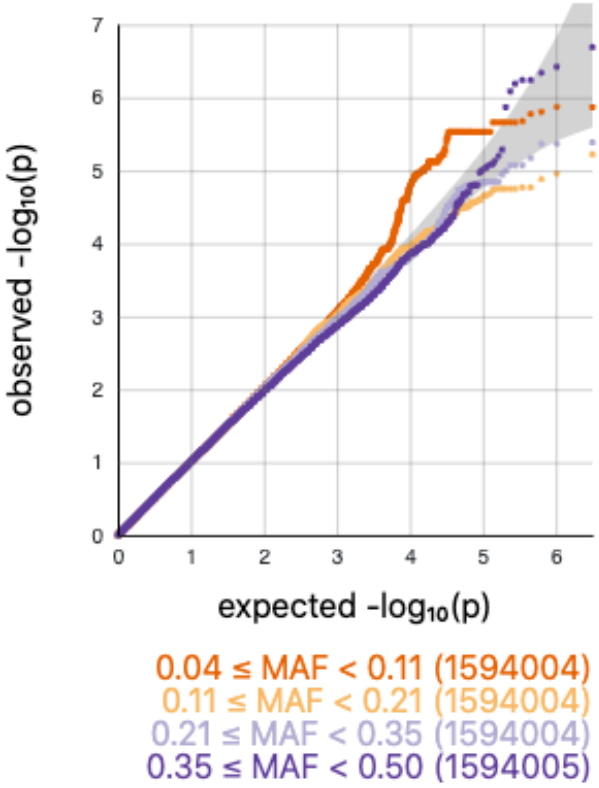


Figure 2. Regular cannabis use QQ Plot



GC lambda 0.5: 1.011
GC lambda 0.1: 1.006
GC lambda 0.01: 0.997
GC lambda 0.001: 0.988

Figure 3. Heaviness of use Manhattan Plot

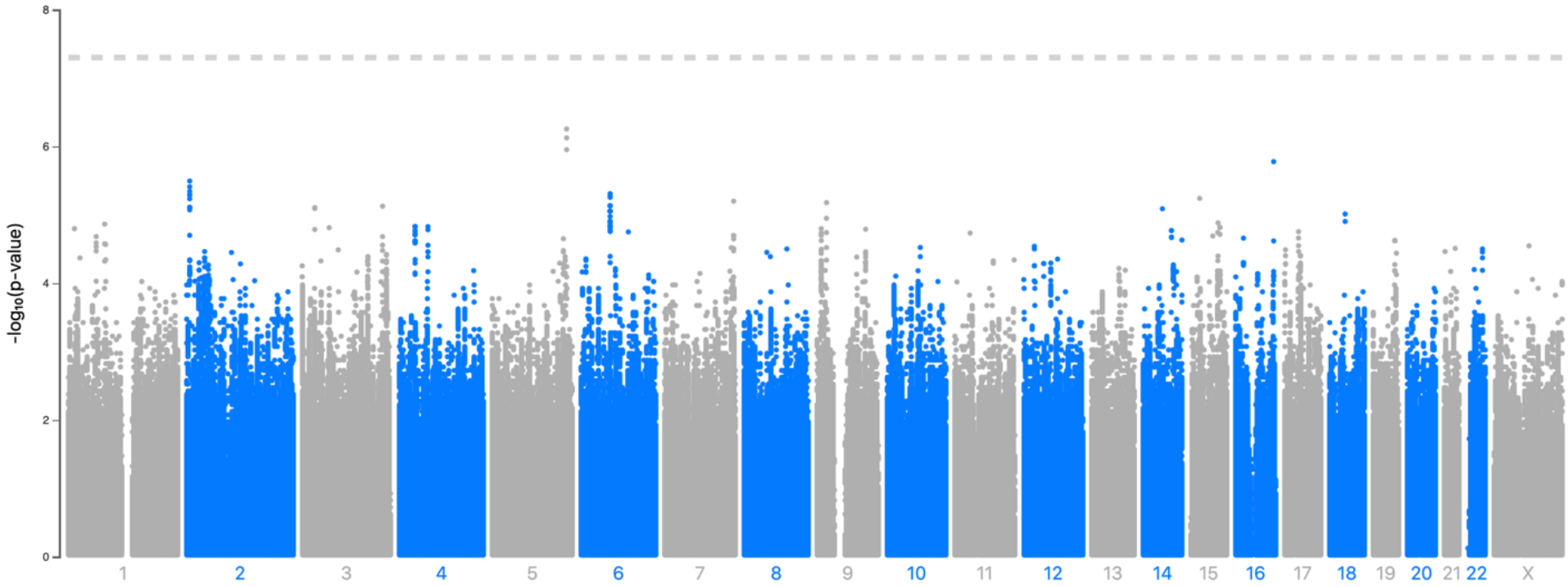
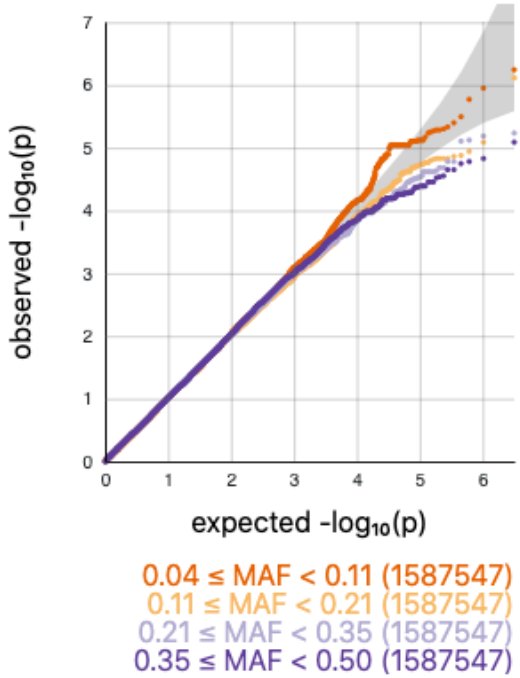


Figure 4. Heaviness of cannabis use QQ Plot



GC lambda 0.5: 1.006
GC lambda 0.1: 1.012
GC lambda 0.01: 1.012
GC lambda 0.001: 1.000

Figure 5. Regular cannabis PRS model fit

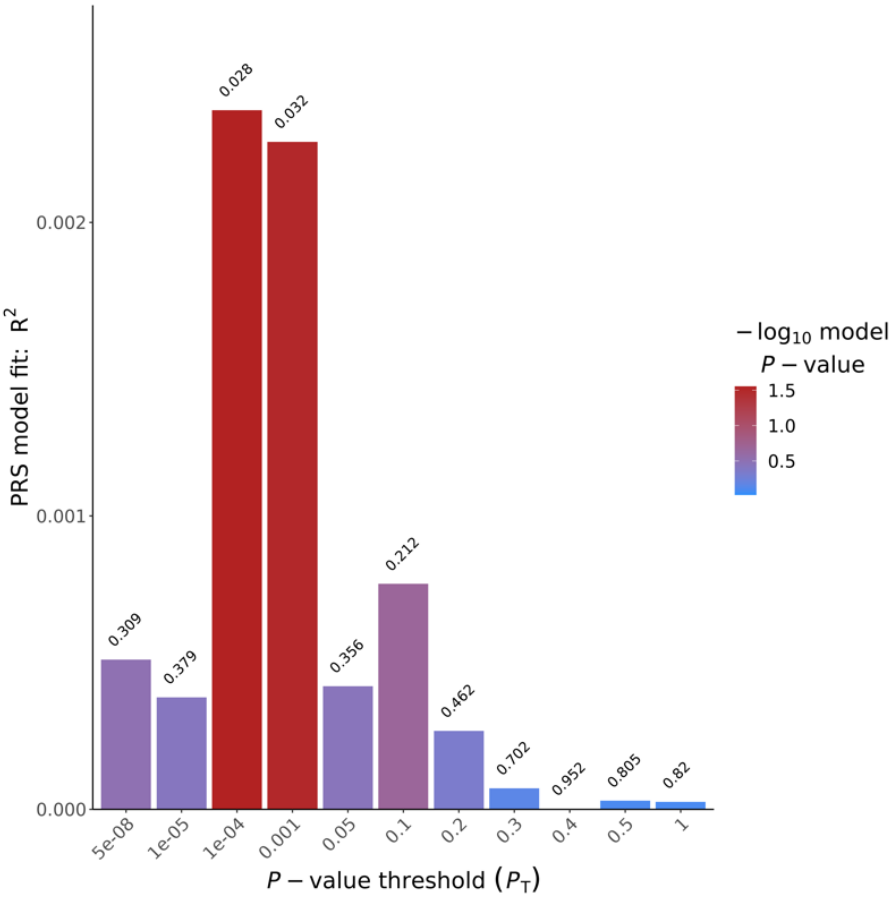


Figure 6. Heaviness of use PRS model fit

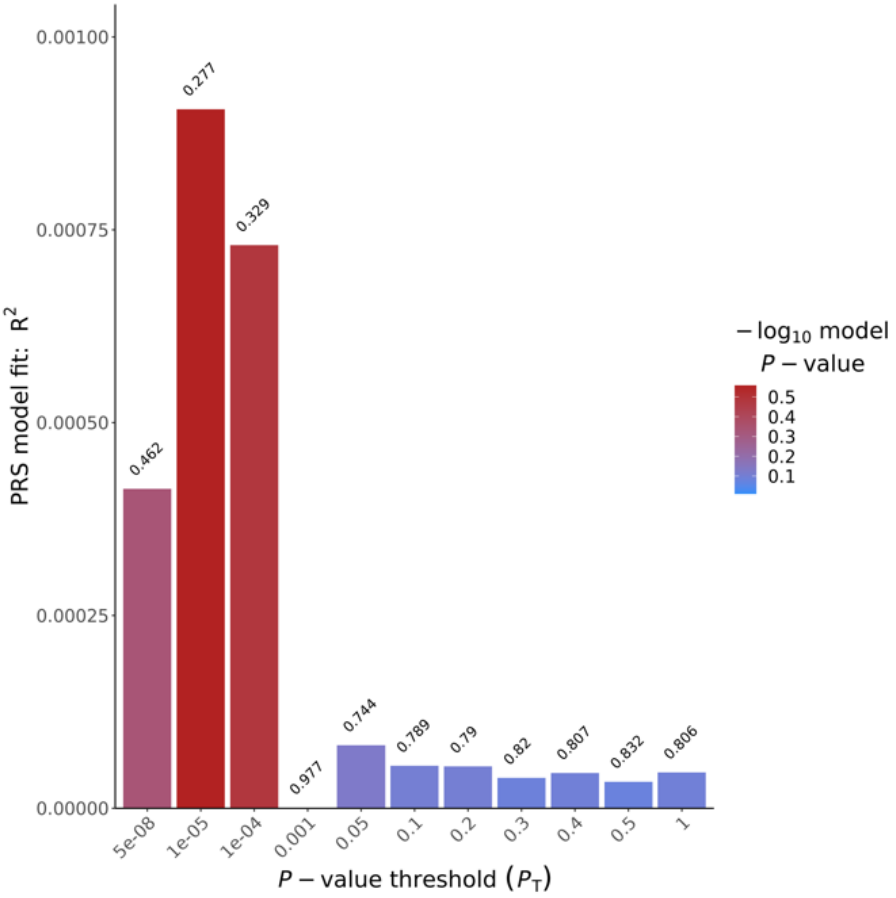
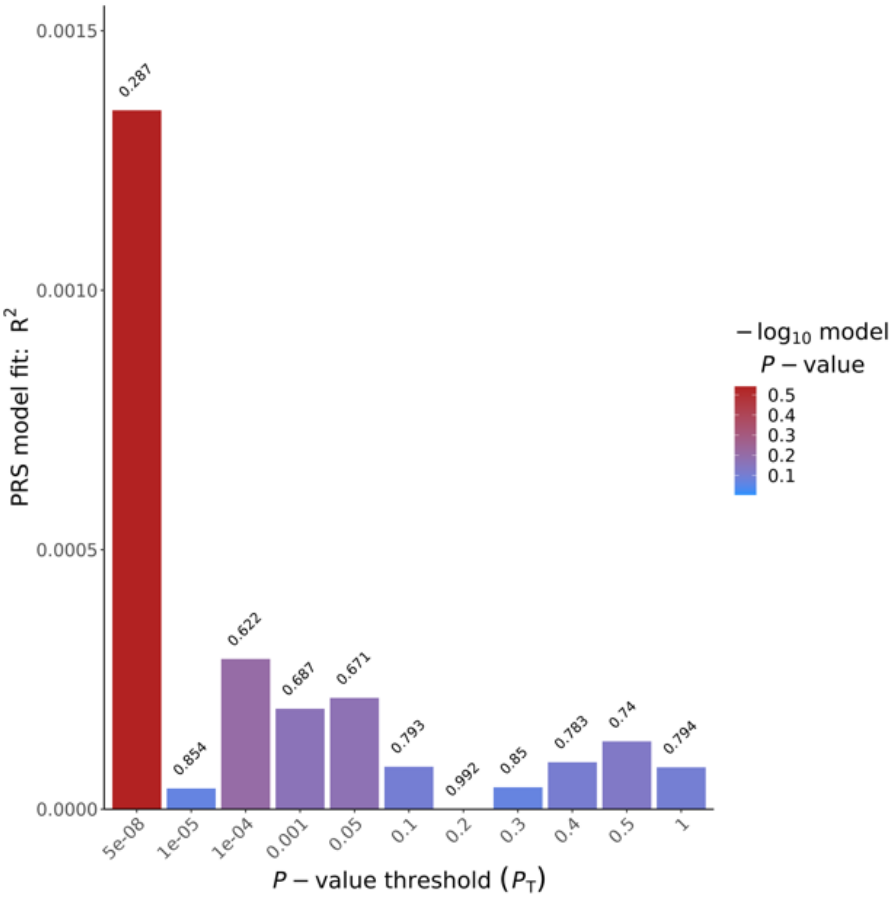


Figure 7. Cannabis dependence PRS model of fit



5 CHAPTER 5

5.1 Genetic contribution to opioid use disorder treatment outcomes: A genome-wide association study and polygenic risk score study

Authors & Affiliations
Alannah Hillmer[1], Caroul Chawar[2], Amel Lamri[3], Jacqueline Hudson[4], Flavio Kapczinski[5], Luciano Minuzzi[6], David C. Marsh[7], Lehana Thabane[8], Zainab Samaan[9]* & Andrew D. Paterson[10]*

1. Neuroscience Graduate Program; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; hillmea@mcmaster.ca
2. MEND Research Lab; Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; caroulchawar@gmail.com
3. Population Health Research Institute, Hamilton, ON, Canada; Department of Health Research Method, Evidence, and Impact, McMaster University, Hamilton, ON, Canada; lamria@mcmaster.ca
4. Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; jhudson@mcmaster.ca
5. Professor Emeritus, McMaster University, Hamilton, ON, Canada; Faculty of Medical Sciences, Universidade Federal do Rio Grande do Sol, Porto Alegre, Brazil; flavio.kapczinski@gmail.com
6. Department of Psychiatry and Behavioural Neuroscience; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; minuzzi@mcmaster.ca
7. Canadian Addiction Treatment Centres; Northern Ontario School of Medicine; 935 Ramsey Lake Rd., Sudbury, ON, P3E 2C6; dmarsh@nosm.ca
8. Department of Health Research Method, Evidence & Impact; 1280 Main St. W., Hamilton, ON, L8S 4L8; thabanl@mcmaster.ca
9. Department of Psychiatry and Behavioural Neurosciences; McMaster University; 100 West 5th St., Hamilton, ON, L8N 3K7; samaanz@mcmaster.ca
10. Program in Genetics and Genome Biology, The Hospital for Sick Children; Divisions of Biostatistics and Epidemiology, Dalla Lana School of Public Health; University of Toronto; 686 Bay Street, Toronto, Ontario Canada M5G 0A4; andrew.paterson@sickkids.ca

*Corresponding Authors:
Dr. Zainab Samaan (responsible for phenotype and study methods)
100 West 5th St., Hamilton, ON L8N3K7, Canada
Telephone: 905-522-1155 ext. 35448
Fax: 905-381-5629

Dr. Andrew Paterson (responsible for genotype analysis and methods)
686 Bay Street, Toronto, ON M5G0A4, Canada
Telephone: 416-813-6994
Fax: 416-813-2150

5.1.1 Abstract

Background: Opioid Use Disorder continues to be a health concern with a high rate of opioid related deaths occurring worldwide as well as in Canada. Medication Assisted Treatment (MAT) has been shown to reduce opioid withdrawal and cravings, and reducing opioid use, however variability exists in individual's treatment outcomes. Pharmacogenetic studies investigating methadone (a common MAT) dose has identified several candidate genes, namely *OPRM1*, *OPRD1*, *ABCB1*, and *CYP2B6,* however only 5 genome-wide associations studies have ever been reported on MAT, showing inconsistent findings.

Objectives: This study aims to explore genetic variants associated with MAT outcomes through genome-wide association study (GWAS) methods and test the association between genetic variants previously associated with opioid use outcomes in a novel sample through a polygenic risk score (PRS).

Methods: Participants were recruited from the GENetics of Opioid Addiction (GENOA) and the Pharmacogenetics of Opioid Substitution Treatment (POST) studies. The study outcomes of interest are: continued opioid use while on treatment (yes/no), relapse (yes/no), methadone dose (daily dose reported in milligrams) and opioid overdose (yes/no). Logistic regressions were performed for binary outcomes and a linear regression preformed for the continuous outcome, to test the association between genetic variants and each outcome. GWAS summary statistics from a recent large GWAS meta-analysis investigating methadone dose were used to conduct PRS.

Results: All GWASs conducted in this study did not identify any SNPs reaching genome-wide significance. SNPs approaching genome-wide significance included rs5868616 on chromosome 5, rs10912116 on chromosome 1, rs6670338 on chromosome 1 and rs12777585 on chromosome 10 for continued opioid use, relapse, methadone dose and opioid overdose respectively. The PRS identified statistically significant results ($p<0.05$) for the outcome of methadone dose (R^2=3.45E-3) only.

Conclusion: While this study did not identify any SNPs reaching genome-wide significance, it provides an association between a PRS and methadone dose. The PRS may have a potential as a tool to help identify a suitable methadone dose in this population and indirectly influence treatment outcome.

Keywords: genetics, opioid, methadone, polygenic risk score, genome-wide

5.1.2 Introduction

Opioid Use Disorder (OUD) continues to be a health concern with a high rate of opioid related deaths being reported worldwide. Globally, approximately 62 million people reporting opioid use in 2019 and approximately 115,000 people died from an opioid overdose in 2017 worldwide (1). In Canada, 5,368 apparent opioid related deaths occurred between January to September of 2021 and an increasing trend of opioid related deaths exists from 2016-2020 (2). OUD is a chronic relapsing condition and Medication Assisted Treatments (MATs) are critical parts of the strategy to address the opioid epidemic and include the controlled administration of opioid agonist or antagonist (3,4). The World Health Organization recommends both methadone and buprenorphine/naloxone (also known as suboxone) as MATs (1,3).

Methadone Maintenance Treatment (MMT) has been reported to decrease opioid cravings and opioid use, with the treatment target aiming to help individuals control opioid use and regain stability (5–7). While methadone can be effective, there is a large variability in methadone effective dose; if individuals are on too low of a dose they may experience withdrawal symptoms and if too high of a dose, they may experience drowsiness, confusion and mental impairment (8). Of further concern, inappropriate dosing of methadone may lead to relapse or increase the individuals risk of overdose due to direct effects or interacting with other illicitly used opioids, and as such, continued opioid use is one of the most common risk factors for mortality among patients receiving MMT (9–11). Further, while MMT has been shown effect in reducing rates of relapse, there are a number of individuals who continue to use illicit opioid while in treatment and are at risk of opioid overdose (12,13). Thus, it is important to consider outcomes of MMT treatment in addition to methadone dose such as continued opioid use, relapse and opioid overdose.

Due to the variability in individual methadone dose tolerance and risk, interest in a genetic predisposition to MMT outcomes has been the focus of research (14–18). The most common Single Nucleotide Polymorphisms (SNPs) associated with MMT outcomes (including opioid addiction, methadone dose, methadone metabolism and plasma concentrations, opioid cessation, response to treatment and continued opioid use) corresponds to *OPRM1*, *OPRD1*, *ABCB1*, and *CYP2B6* (16,19,20). However, many of the genetic studies assessing the pharmacogenetics of MAT are candidate gene studies with small sample size, with a recent systematic review identifying 5 genome-wide association studies investigating MAT outcomes showing inconsistent findings (19).

With the few studies reporting Genome-wide Association Study (GWAS) results of MAT outcomes and the risk of mortality faced by this population, we aim to examine if novel genetic variants are associated with MMT outcomes, including continued opioid use, relapse, methadone dose and opioid overdose. Further, we aim to determine the known genetic variants associated with methadone dose in the literature are associated with MMT outcomes in our population through a Polygenic Risk Score (PRS).

5.1.2.1 Objectives

This study aims to identify novel genetic variants and test the association between genetic variants previously associated with MMT outcomes in a clinical sample. The study objectives are to:

1. Conduct a GWAS to identify novel genetic variants with various MMT outcomes
2. Investigate sex differences of any genetic variants approaching or reaching genome-wide significance
3. Conduct a PRS based on a large previously published GWAS on MMT

5.1.3 Methods

In accordance with the Strengthening the Reporting of Genetic Association studies (STREGA) guidelines, an extension of Strengthening the Reporting of Observational studies in Epidemiology (STROBE) statement, an accompanying STREGA checklist can be found in Supplementary File 1 (21).

5.1.3.1 Study Design and Setting

Data were collected as part of the GENetics of Opioid Addiction (GENOA) and Pharmacogenetics of Opioid Substitution Treatment Response (POST) programs. The GENOA and POST studies are prospective cohort studies conducted in collaboration with the Canadian Addiction Treatment Centre (CATC) and McMaster University. The GENOA and POST studies were designed to identify factors associated with opioid use and treatment outcomes including genetic risk factors in patients diagnosed with OUD and receiving treatment (22). Details on the GENOA and POST study have been previously reported (22). Briefly, participants (n_{GENOA}=1,536, n_{POST}=2,544) were recruited from 76 CATC sites across Ontario, Canada (GENOA; 2013-2016, POST; 2018-present). At study recruitment, participants completed an extensive interview with a trained researcher and were asked to provide a DNA sample. As previously reported, participants were followed for a 12-month period through their electronic medical record which documented their weekly or biweekly Urine Drug Screens (UDS). At study recruitment, participants UDS for 3-months prior to study entry was collected, thus a total of 15-months of UDS were obtained. Participants were given a small coffee shop gift card after every face-to-face interview in appreciation for participating (23,24).

The GENOA and POST studies were reviewed and approved by the Hamilton Integrated Research Ethics Board (GENOA: 11-056, POST: 4556).

5.1.3.2 Participants

While GENOA and POST studies had similar methods, differences in inclusion criteria exits. For the GENOA study, patients were eligible to participate if they were 18 years or older and met the criteria for Diagnostic and Statistical Manual – fourth edition (DSM-IV) opioid addiction requiring treatment (later replaced in Diagnostic and Statistical Manual – fifth edition (DSM-5) as Opioid Use Disorder (OUD)) (25). Patients were excluded if they did not speak English or refused to provide a blood sample (for

DNA). For the POST study, patients were eligible to participate if they were 16 years or older and met the criteria for DSM-5 Opioid Use Disorder (OUD). Patients were excluded if they did not speak English or refused to provide a saliva sample (for DNA).

5.1.3.3 Eligibility criteria

In addition to meeting eligibility for the GENOA or POST study criteria, further inclusion criteria for this study included being on MMT. For the measures of continued opioid use and relapse participants had to have UDSs assessing for the presence of opioids for a minimum duration of 3 months and 6 months, respectively. For continued opioid use and relapse, participants were excluded for self-reporting a current prescription for opioids due to the uncertainty for UDS screens for opioids would be from licit or illicit use. For the measures of methadone dose and opioid overdose participants were excluded if they did not self-report methadone dose or if they have ever had an opioid overdose, respectively. Participants in the GENOA study were not asked about lifetime opioid overdoes, and were therefore excluded from the opioid overdose analyses.

5.1.3.4 Variables and Quantitative variables

Outcomes measured in the study include the following:

1. Continued opioid use; defined as any opioid positive UDS observed over a duration of 3 to 15 months, measured as a binary variable
2. Relapse; defined as an event of an opioid positive UDS following at least 3 months of opioid negative UDSs, measured as a binary variable
3. Methadone dose; defined as the amount of methadone a patient is administered at the time of study recruitment in milligrams, measured as a continuous variable
4. Opioid overdose; defined as any self-reported opioid overdose reported at the time of study recruitment, measured as a binary variable

Covariates for the measures of continued opioid use, relapse and opioid overdose that were accounted for in the statistical models included: sex, age in years, dose of methadone in milligrams, duration on MMT in months and principal components accounting for differences due to population stratification. Covariate for the measure of methadone dose that were accounted for in the statistical models included: sex, age in years, duration on MMT in months, weight in kilograms and principal components.

5.1.3.5 Data Sources/Measurement

Participant's urine sample was tested based on reported or suspected drug use and were analyzed and reported as the number of positive screens for the drug detected in the test using the FaStep Assay (Trimedic Supply Network Ltd, Concord, Ontario, Canada) (26). Methadone dose and opioid overdose was self-reported by participants at study recruitment. Methadone dose was transformed by dividing by 10 for clinical interpretation, as changes in 10 milligrams of methadone was deemed as clinically meaningful compared to changes in 1 milligrams of methadone.

5.1.3.6 Quality Control Checks

The GENOA study collected blood samples and the POST study collected saliva samples at study recruitment for the extraction of DNA. The blood and saliva samples were sent to Génome Québec Innovation Centre for genotyping using GenomeStudio and the Illumina Global Screening Array – 24 v1.0 (27–29). R version 3.3.3 was used for quality control checks and variant quality control procedures were applied using PLINK v1.90 (30,31).

Samples were excluded if they have a low variant call rate (<99%), inconsistencies between self-reported vs. genetically determined sex or ancestry, if they exhibit excess heterozygosity suggestive of sample contamination (exchange of DNA between two or more samples), if they were a duplicate of if they were determined to be a first-degree or second-degree relative through a heritability analysis. Variants were excluded if they have a low call rate across samples (<99%), or if the minor allele was less than 0.05. As participants were predominantly European Ancestry (n=2,958), only those who were of European Ancestry were selected for imputation. Imputation was completed by TOPMed Imputation Server (Version R2) using the software Eagle v2.4 and Minimac4 for phasing (32–35). Post-imputation filtering excluded SNPs with Rsq quality metrics of less than 0.3 and minor allele frequencies lower than 0.05. Further information on the quality control steps were previously reported (22).

5.1.3.7 Bias

Although measures were taken to identify and mitigate areas of bias, potential sources of bias remained. While outcomes of continued opioid use and relapse were defined through UDSs to provide an objective measure, methadone dose and opioid overdose were self-report, allowing for potential social desirability or recall biases. Sex-differences have been identified in social desirability biases, such that differing response may have occurred within males and females based on what response seemed more desirable (36). Further, it is possible that participants did not accurately recall what their current methadone dose is or if they have experienced an opioid overdose. Additionally, the current study may be biased by volunteer bias. In addition to bias that exists within individuals for the participation of research, biases exist in participants who are willing to participate in genetic studies, thus the sample population may not be representative of the entire OUD population receiving treatment (37). In addition, study results are not generalizable outside of the European Ancestry or those receiving other MAT for OUD other than MMT. Lastly, due to the observational nature of this study, it is not possible to control for all extraneous confounding variables that may exist.

5.1.3.8 Sample Size

Of the total participants enrolled in the GENOA and POST study with genetic samples available (n=4,621), 2251 participants of European Ancestry passed the genetic quality control steps and were used for this study.

5.1.3.9 Statistical Methods

Descriptive statistics were reported on the total sample, by sex, to describe the demographic and clinical characteristics. Continuous variables were expressed as means with standard deviations, while categorical variables were expressed as counts.

Separate regression analyses were preformed to test the association between the outcomes and genetic variants. Logistic regressions were conducted for the outcomes of continued opioid use, relapse, and opioid overdose and a linear regression was used for the outcome of methadone dose. All aforementioned covariates were adjusted for in their respective analyses through the use of an additive model for genotype coding. Identical regression analyses as above were conducted separately for male and female subsets, as well as an interaction model, for SNPs approaching, or meeting, genome-wide significance.

For the PRS discovery statistics of SNPs reaching the threshold of $P < 5.0 \times 10^{-5}$ from a GWAS investigating daily methadone dose, using only the European Ancestry summary statistics (38). The GWAS summary statistics were converted from GRCh build 37 to 38 to match our data using the UCSC*liftOver* tool (39). The GWAS summary statistics results were pruned using PRSice-2 using an R^2 threshold of 0.5 and within 250kb. Subset of SNPs were selected at decreasingly liberal P-value thresholds (0.0001 and $1x10^{-5}$), based on the previously reported GWAS data availability. PRSs were calculated separately for each outcome, with all aforementioned covariates adjusted for in their respective analyses.

Samples with missing outcome values were excluded from the analysis. Missing values for the covariates of each analysis were imputed via mean substitution, from the averages of the values calculated per analysis using R studio 3.3.3 (31,40). For the outcome of continued opioid use, methadone dose and duration on MMT were imputed and for the outcome of methadone dose, duration on MMT and weight were imputed using the method mentioned above.

All statistical analyses were performed on PLINK v1.09, R studio 3.3.3 and PRSice-2 (30,31,40–42).

5.1.4 Results

5.1.4.1 Participants

Of the 2,251 participants included in this study, 2 were missing their methadone dose amount and 94 had no UDS reported. Of the remaining 2,157 participants with UDS, 169 only had one time-point and were excluded from the relapse outcome. Finally, only 1,327 of 2,251 participants reported if they had previously experienced an opioid overdose. Therefore, 2,157 were included in the continued opioid use analysis, 1,988 in the relapse analysis, 2,249 in the methadone dose analysis and 1,327 in the opioid overdose analysis.

5.1.4.2 Descriptive Data

The current study included 924 participants from the GENOA study and 1,327 from the POST study. Majority of study participants were male (57.53%) and had a mean age of 39.26 years. It was most common for participants to report to never been married, unemployed and have less than a grade 12 education. Participants reported an average methadone dose of 71.76 mg/day and have been on treatment for 53.53 months, or approximately 4.5 years, and 23.56% reported currently having a prescription to opioids, other than methadone for medical indications, namely, pain conditions. Finally, 6,377,206 SNPs passed the quality control steps and were included in the GWAS and of the 293 SNPs available from the previously reported GWAS summary statistics, 39 SNPs were included in the PRSs. The full list of participants demographics can be found in Table 1.

Table 9. Sample demographics

	Total	Male	Female
N (%)	2251	1295 (57.53)	956 (42.47)
GENOA	924	539 (58.33)	385 (41.67)
POST	1327	756 (56.97)	571 (43.03)
Age in years, Mean (SD)	39.26 (11.11)	40.06 (11.17)	38.19 (10.95)
Marital status, N (%)[a]			
Never Married	1030 (48.75)	623 (51.70)	407 (44.82)
Common law	425 (20.11)	222 (18.42)	203 (22.36)
Currently Married	194 (9.18)	111 (9.21)	83 (9.14)
Separated	202 (9.56)	99 (8.22)	103 (11.34)
Divorced	195 (9.23)	120 (9.96)	75 (8.26)
Widowed	67 (3.17)	30 (2.59)	37 (4.07)
Currently employed, N (%)[b]	717 (33.93)	487 (40.38)	230 (25.36)
Education, N (%)[c]			
Less than grade 9	516 (24.47)	296 (24.56)	220 (24.34)
Grade 9-12	974 (46.18)	606 (50.29)	368 (40.71)
Trade School	66 (3.13)	48 (3.98)	18 (1.99)
College/University/Graduate School	553 (26.22)	255 (21.16)	298 (32.96)
Methadone Dose in mg/day, Mean (SD)[d]	71.76 (43.07)	74.18 (44.57)	68.47 (40.75)
Duration on MMT in months, Mean (SD)[e]	53.53 (58.92)	53.50 (56.89)	53.57 (61.61)
Current opioid prescription, N (%)[f]	53 (23.56)	31 (2.39)	22 (2.30)
Continued opioid use, N (%)[g]	1716 (79.55)	993 (80.02)	723 (78.93)

Relapse, N (%)[h]	631 (31.74)	362 (31.81)	269 (31.65)
Reported opioid overdose in lifetime, N (%)[i]	437 (32.93)	244 (32.28)	193 (33.80)

[a]Data available for n_{Total}= 2,113, n_{Male}= 1,205, n_{Female}=908
[b]Data available for n_{Total}= 2,113, n_{Male}= 1,206, n_{Female}=907
[c]Data available for n_{Total}= 2,109, n_{Male}= 1,205, n_{Female}= 904
[d]Data available for n_{Total}= 2,249, n_{Male}= 1,294, n_{Female}= 955
[e]Data available for n_{Total}= 2,242, n_{Male}= 1,290, n_{Female}= 952
[f]Data available for n_{Total}= 2,250, n_{Male}= 1,295, n_{Female}= 955
[g]Data available for n_{Total}= 2,157, n_{Male}= 1,241, n_{Female}= 916
[h]Data available for n_{Total}= 1,988, n_{Male}= 1,138, n_{Female}= 850
[i]Data available for n_{Total}= 1,327, n_{Male}= 756, n_{Female}= 571

5.1.4.3 Main Results

We identified one SNP approaching genome-wide significance for each outcome; for continued opioid use rs5868616 on chromosome 5 (8.87×10^{-7}), for relapse rs10912116 on chromosome 1 ($4.40\text{x}10^{-7}$), for methadone dose rs6670338 on chromosome 1 ($5.79\text{x}10^{-7}$), and for opioid overdose rs12777585 on chromosome 10 ($2.60\text{x}10^{-6}$). For all outcomes, using the top SNP as a covariate in the analysis did not result in any further significant results, thus no other strong signals within the region exist. Results on the top SNPs for each respective outcome can be found in Table 2. Manhattan plots and QQ plots for the respective study outcomes can be found in Figure 1-8.

Table 10. Top SNPs from each GWAS

Outcome	Chr	SNP	BP	A1	OR/BETA	P
Continued Opioid Use	5	rs5868616	71874588	GT	1.624	8.87×10^{-7}
Relapse	1	rs10912116	187657898	T	0.684	$4.40\text{x}10^{-7}$
Methadone Dose	1	rs6670338	91764748	G	0.043	$5.79\text{x}10^{-7}$
Opioid Overdose	10	rs12777585	116685252	C	1.551	$2.60\text{x}10^{-6}$

Chr=chromosome, SNP=single nucleotide polymorphism, BP=base pair, A1=reference allele, OR=odds ratio, BETA= beta coefficient

Results from the sex-stratified association analyses between the SNPs approaching GWAS significance for each respective outcome are reported in Table 3, using the p-value of less than 0.05 to indicate statistical significance. The GT allele of rs5868616 was significant associated with an increased odds of continued opioid use in both males [odds ratio (OR)=1.56, 95% confidence interval (CI)=1.19, 2.03, p=1.09E-3] and females [OR= 1.70, 95% CI=1.28, 2.25, p=2.65E-4], however the sex by SNP interaction was not significant. The T allele of rs10912116 was significantly associated

with a decreased odds of relapse in both males [OR= 0.77, 95% CI= 0.63, 0.94, p=8.87E-3] and females [OR= 0.59, 95% CI=0.47, 0.74, p=5.07E-6] and the sex by SNP interaction was also significant [OR= 0.74, 95% CI=0.55, 1.00, p=4.89E-2]. The A allele of rs6670338 was significantly associated with methadone dose in both males [Beta=0.04, standard error (SE)=0.01, p=8.21E-4] and females [Beta=0.05, SE=0.01, p=2.32E-4], however, the sex by SNP interaction was not significant. Finally, the G allele of rs12777585 was significantly associated with an increased odds of opioid overdose in both males [OR=1.67, 95% CI=1.31, 2.13, p=3.76E-5] and females [OR=1.40, 95% CI=1.06, 1.86, p=1.88E-2], however the sex by SNP interaction was not significant.

Table 11. SNPs and associated outcomes stratified by sex

Outcome	SNP	N	Reference Allele	OR/BETA	95 % CI/SE	P
Continued Opioid Use	**rs5868616**		GT			
	Males	1241		1.56	1.19, 2.03	**1.09E-3**
	Females	916		1.70	1.28, 2.25	**2.65E-4**
	Interaction	2157		1.10	0.75, 1.62	0.63
Relapse	**rs10912116**		T			
	Males	1138		0.77	0.63, 0.94	**8.87E-3**
	Females	850		0.59	0.47, 0.74	**5.07E-6**
	Interaction	1988		0.74	0.55, 1.00	**4.89E-2**
Methadone dose	**rs6670338**		A			
	Males	1294		0.04	0.01	**8.21E-4**
	Females	955		0.05	0.01	**2.32E-4**
	Interaction	2249		0.01	0.02	0.39
Opioid Overdose	**rs12777585**		G			
	Males	756		1.67	1.31, 2.13	**3.76E-5**
	Females	571		1.40	1.06, 1.86	**1.88E-2**
	Interaction	1327		0.85	0.59, 1.23	0.39

OR=odds ratio, BETA= beta coefficient, 95% CI = 95% confidence interval levels (lower, upper), SE=standard error

Results from the PRS best-fit model for continued opioid use, relapse, methadone dose and opioid overdose are reported in Table 4, and interpreted with the significance threshold of $p<0.05$. Using the p-value threshold of 1.0 E-5 for the previous GWAS summary statistics, a significant PRS was found for methadone dose (PRS R^2=3.45E-3). No other PRS association with other outcomes were significant. The bar plots depicting the model of fit across p-value thresholds from the previously reported GWAS for each outcome can be seen in Figures 9-12.

Table 12. PRS best model fit for each outcome

Outcome	Threshold	PRS R^2	Full R^2	Null R^2	Coefficient	SE	P	Number of SNPs
Continued opioid use	1.0 E-5	7.37E-4	3.39E-2	3.31E-2	2.07	2.07	0.32	8
Relapse	0.0001	6.70E-4	1.41E-2	1.35E-2	-4.27	4.38	0.33	39
Methadone dose	1.0 E-5	3.45E-3	5.20E-2	5.43E-2	-0.54	0.20	**5.42E-3**	8
Opioid overdose	0.0001	9.58E-4	2.43E-2	2.34E-2	5.12	5.38	0.34	39

PRS R^2=Variance explained by the PRS, Full R^2=Variance explained by the full model, Null R^2=Variance explained by the covariates, SE=standard error, Number of SNPs=number of SNPs included in the model

5.1.5 Discussion

5.1.5.1 Key Results

No SNPs reached genome-wide significance in any GWAS. However, rs5868616 on chromosome 5 for continued opioid use, rs10912116 on chromosome 1 for relapse, rs6670338 on chromosome 1 for methadone dose, and rs12777585 on chromosome 10 for opioid overdose were identified as approaching genome-wide significance. All SNPs were found to be significantly associated with their respective outcome in males and females separately ($P<0.05$), however sex did not modify the association between rs5868616, rs6670338 or rs12777585. The interaction was significant for rs10912116, where sex modifies the association between rs10912116 and relapse such that the odds of relapse higher for females. For the PRS, statistically significant ($P<0.05$) results were observed for methadone dose, however no other significant results were observed.

5.1.5.2 Interpretation

The GWASs did not replicate any known genetic associations of MMT outcomes from the literature. Two SNPs, rs5868616 on chromosome 5 and rs10912116 on chromosome 1, have no known associations with other traits or pathways in the literature. The SNP associated with methadone dose, rs6670338, is part of the transforming growth factor beta receptor 3 (*TGFBR3*), however traits previously associated with areas in this gene include systolic blood pressure, ischemic stroke, diabetes and other traits that are not related to mental health conditions or addiction (43). The SNP associated with opioid overdose, rs12777585, is part of the heat shock protein family A member 12A (*HSPA12A*), wherein SNPs in this region have been previously associated with externalizing behaviours (namely attention deficit hyperactivity disorder, substance use, and antisocial behaviours), educational attainment and smoking initiation (43–46). As OUD, or more specifically substance use behaviour as a whole, is classified as

externalizing as well as known associations with education attainment and smoking behaviour it is possible that this gene region could indicate these traits are not independent or reflect genuine pleiotropy (44,47–49). Further investigation into *HSPA12A* is required to determine the genetic association with opioid overdose.

With growing evidence suggesting that the majority of genetic variants have small effects which collectively contribute to the risk of a disease, it is important to consider other methods, such as PRS, as they are able to aggregate the effects of variants across the genome through creating a weighted sum of the number of risk alleles an individual carries (50–52). We found a statistically significant PRS for methadone dose, suggesting it can be applied to assess the individual level variability in methadone dose. However, the variability explained by genetics was minor, 3.45E-3, with majority of the variance reported explained by the covariates (Null R^2=5.43E-2), suggesting that much of the variance is due to other, non-genetic, factors such as the environment. The PRSs for continued opioid use, relapse and opioid overdose were not significant and, similar to the PRS for methadone dose, the variability explained by genetics was minor (less than 6.70E-4). It is important to note that the GWAS summary statistics from the literature investigated methadone dose, and thus SNPs contributing to individual variability in methadone dose may not contribute to genetic variability in the outcomes of continued opioid use, relapse and opioid overdose despite their clinical associations (38). Thus, it is also important to consider that the lack of significant findings within this study may be due to shared genetic contribution of various substances of abuse or externalizing behaviours (44,53).

Lastly, it is important to discuss the results of the sex-specific analyses as it has been previously reported that women are more likely to present to treatment with higher rates of psychiatric comorbidities, greater life instability, have higher relapse rates, and experience faster dependence progression rates and men are more likely to present to treatment with ongoing drug use and other risky drug-related behaviours than women (18,54,55). Further, biological differences between women and men have been reported, including neuroanatomy and neurochemistry, as well as psychological and behavioural differences such as in cognition, aggression and neurological diseases (56). Thus, while this study did not fine sex-specific differences within MMT outcomes, it is important to continue to investigate the potential genetic differences that may exist in sex-specific treatment outcomes.

5.1.5.3 Limitations and generalizability

In addition to the sources of bias discussed earlier, limitations exist within the study. First, the results from this study are limited to those of European Ancestry and therefore may not be generalizable to individuals from different ancestry backgrounds. Further, it is important to note the level of missing data. Due to the specific criteria of each outcome, multiple participants were lost either due to a lack of UDS or, for opioid overdose specifically, not being a part of the POST study where data collection slightly varied from the GENOA study. It is important to mention that some participants with missing UDS could have left treatment due to a relapse (and continued opioid use),

transferred to another clinic and reminded stable on MMT or entered another treatment facility. More importantly, it is important to note that we do not know their true outcome. Finally, it is important to note that methadone dose or UDS may not have been an accurate measure of treatment response as treatment outcomes can be complex and no one definition has been agreed upon to be the ultimate treatment response (57). In addition, study participants were enrolled in different stages of treatment, as such participants could have been at induction, treatment stabilization or taper stages, each of which vary the amount of methadone as well likelihood of opioid use based on opioid cravings (58).

Regardless, this study provides additional insight into the genetics of MMT response. While further research is required to understand the complexity of OUD and treatment outcomes, it is important to continue to investigate genetic differences in MMT response given the known individual level variability and future clinical implications of personalized care.

5.1.6 Other Information

5.1.6.1 Funding

This study was supported by CIHR (PJT-156306), which played no role in the study design, analysis, reporting or publication.

5.1.6.2 Authors Contributions

ZS is the guarantor. AH, ZS and AP conceptualized the genetic study. AH implemented the design, quality control steps, and methodology with the aid of CC, AL, AP and ZS. AP provided genetic statistical guidance and data clean up methods. DM and JH provided support and collaboration with the CATC clinics. AH prepared the first draft. FK, LM, LT, ZS and AP revised the study design, data analysis, and data interpretation. All authors reviewed and approved the final manuscript.

5.1.7 List of Abbreviations

Canadian Addiction Treatment Centre (CATC)
Confidence Interval (CI)
Diagnostic and Statistical Manual – fourth edition (DSM-IV)
Diagnostic and Statistical Manual – fifth edition (DSM-5)
GENetics of Opioid Addiction (GENOA)
Genome-Wide Association Study (GWAS)
Medication Assisted Treatment (MAT)
Methadone Maintenance Treatment (MMT)
Odds Ratio (OR)
Opioid Use Disorder (OUD)
Pharmacogenetics of Opioid Substitution Treatment Response (POST)
Polygenic Risk Score (PRS)
Single Nucleotide Polymorphism (SNP)

Standard Error (SE)
Strengthening the Reporting of Genetic Association studies (STREGA)
Strengthening the Reporting of Observational studies in Epidemiology (STROBE)
Urine Drug Screens (UDS)

5.1.8 References

1. World Health Organization. Opioid overdose [Internet]. 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/opioid-overdose
2. The Government of Canada. Opioid- and Stimulant-related Harms in Canada. 2022.
3. Morin KA, Eibl JK, Franklyn AM, Marsh DC. The opioid crisis: past, present and future policy climate in Ontario, Canada. Subst Abuse Treat Prev Policy. 2017;12(1):1–7.
4. Coffa D, Snyder H. Opioid use disorder: medical treatment options. Am Fam Physician. 2019;100(7):416–25.
5. Joseph H, Stancliff S, Langrod J. Methadone maintenance treatment (MMT): A review of historical and clinical issues. Mt Sinai J Med. 2000;67(5):347–64.
6. Marsch LA. The efficacy of methadone maintenance interventions in reducing illicit opiate use, HIV risk behavior and criminality: a meta-analysis. Addiction. 1998;93(4):515–32.
7. Eap CB, Buclin T, Baumann P. Interindividual variability of the clinical pharmacokinetics of methadone. Clin Pharmacokinet. 2002;41(14):1153–93.
8. Somogyi AA, Barratt DT, Ali RL, Coller JK. Pharmacogenomics of methadone maintenance treatment. Pharmacogenomics. 2014;15(7):1007–27.
9. Prevention C for DC and. Calculating total daily dose of opioids for safer dosage. 2017.
10. Goud A, Kiefer E, Keller MS, Truong L, SooHoo S, Riggs R V. Calculating maximum morphine equivalent daily dose from prescription directions for use in the electronic health record: a case report. JAMIA open. 2019;2(3):296–300.
11. Huang C, Lee C. Factors associated with mortality among heroin users after seeking treatment with methadone: a population-based cohort study in Taiwan. J Subst Abus Treat. 2013;44(3):295–300.
12. Oviedo-Joekes E, Brissette S, Marsh DC, Lauzon P, Guh D, Anis A, et al. Diacetylmorphine versus Methadone for the Treatment of Opioid Addiction. N Engl J Med [Internet]. 2009 Aug 20;361(8):777–86. Available from: https://doi.org/10.1056/NEJMoa0810635
13. Mattick RP, Breen C, Kimber J, Davoli M. Methadone maintenance therapy versus no opioid replacement therapy for opioid dependence. Cochrane Database Syst Rev [Internet]. 2009;(3). Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01765807/full
14. Crettol S, Monnat M, Eap CB. Could pharmacogenetic data explain part of the interindividual sensitivity to methadone-induced respiratory depression? Crit Care. 2007;11(1):1–2.
15. Li Y, Kantelip J-PJ-P, Gerritsen-van Schieveen P, Davani S. Interindividual variability of methadone response. 2008;12(2):109–24.
16. Berrettini W. A brief review of the genetics and pharmacogenetics of opioid use disorders. Dialogues Clin Neurosci. 2022;
17. Dennis BB, Bawor M, Thabane L, Sohani Z, Samaan Z. Impact of ABCB1 and

CYP2B6 genetic polymorphisms on methadone metabolism, dose and treatment response in patients with opioid addiction: a systematic review and meta-analysis. PLoS One. 2014;9(1).
18. Chawar C, Hillmer A, Lamri A, Kapczinski F, Thabane L, Pare G, et al. Implications of OPRM1 and CYP2B6 variants on treatment outcomes in methadone-maintained patients in Ontario: Exploring sex differences. PLoS One. 2021;16(12):e0261201.
19. Chawar C, Hillmer A, Sanger S, D'Elia A, Panesar B, Guan L, et al. A systematic review of GWAS identified SNPs associated with outcomes of medications for opioid use disorder. Addict Sci Clin Pract. 2021;16(1):1–14.
20. Fonseca F, Torrens M. Pharmacogenetics of methadone response. Mol Diagn Ther. 2018;22(1):57–78.
21. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):p.e296.
22. Hillmer A, Chawar C, Lamri A, Hudson J, Kapczinski F, Minuzzi L, et al. Genetics of cannabis use in opioid use disorder: A Polygenic Risk Score and Genome-Wide Association Study. 2022.
23. Zielinski L, Bhatt M, Sanger N, Plater C, Worster A, Varenbut M, et al. Association between cannabis use and methadone maintenance treatment outcomes: an investigation into sex differences. Biol Sex Differ. 2017;8(1):8.
24. Rosic T, Naji L, Panesar B, Chai DB, Sanger N, Dennis BB, et al. Are patients' goals in treatment associated with expected treatment outcomes? Findings from a mixed-methods study on outpatient pharmacological treatment for opioid use disorder. BMJ Open. 2021;11(1):e044017.
25. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychaitric Publishing; 2013.
26. Trimedic Supply Network Ltd. FaStep Assay [Internet]. Available from: https://www.trimedic-inc.com/wp-content/uploads/2018/04/Fastep-Package-Insert.pdf
27. Centre d'expertise et de services Génome Québec [Internet]. Available from: https://cesgq.com/
28. illumina. GenomeStudio [Internet]. 2022. Available from: https://support.illumina.com/array/array_software/genomestudio/downloads.html
29. illumina. Infinium Global Screening Array - 24 v1.0 [Internet]. 2017. Available from: https://grcf.jhmi.edu/wp-content/uploads/2017/12/infinium-commercial-gsa-data-sheet-370-2016-016.pdf
30. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
31. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
32. Loh P-R, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, et al.

Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48(11):1443.
33. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284–7.
34. TOPMed. TOPMed Imputation Server - Reference Panels [Internet]. Available from: https://topmedimpute.readthedocs.io/en/latest/reference-panels.html
35. Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature. 2021;590(7845):290–9.
36. Paunonen S V. Sex differences in judgments of social desirability. J Pers. 2016;84(4):423–32.
37. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2018;47(1):226–35.
38. Smith AH, Jensen KP, Li J, Nunez Y, Farrer LA, Hakonarson H, et al. Genome-wide association study of therapeutic opioid dosing identifies a novel locus upstream of OPRM1. Mol Psychiatry. 2017;22(3):346–52.
39. Kuhn RM, Haussler D, Kent WJ. The UCSC genome browser and associated tools. Brief Bioinform. 2013;14(2):144–61.
40. RStudio Team. RStudio: Integrated Development for R. Boston, MA: RStudio; 2020.
41. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience. 2019;8(7):giz082.
42. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4(1):s13742-015.
43. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019;47(D1):D1005–12.
44. Karlsson Linnér R, Mallard TT, Barr PB, Sanchez-Roige S, Madole JW, Driver MN, et al. Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. Nat Neurosci. 2021;24(10):1367–76.
45. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet. 2018;50(8):1112–21.
46. Siedlinski M, Cho MH, Bakke P, Gulsvik A, Lomas DA, Anderson W, et al. Genome-wide association study of smoking behaviours in patients with COPD. Thorax. 2011;66(10):894–902.
47. Dahlman D, Ohlsson H, Edwards AC, Sundquist J, Håkansson A, Sundquist K. Socioeconomic correlates of incident and fatal opioid overdose among Swedish people with opioid use disorder. Subst Abuse Treat Prev Policy. 2021;16(1):1–8.

48. Ellis MS, Kasper ZA, Cicero TJ. The impact of opioid use disorder on levels of educational attainment: perceived benefits and consequences. Drug Alcohol Depend. 2020;206:107618.
49. Rajabi A, Dehghani M, Shojaei A, Farjam M, Motevalian SA. Association between tobacco smoking and opioid use: A meta-analysis. Addict Behav. 2019;92:225–35.
50. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. 2020;15(9):2759–72.
51. Khera A V, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet. 2018;50(9):1219–24.
52. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. Genome Med. 2020;12(1):1–11.
53. Popescu A, Marian M, Drăgoi AM, Costea R-V. Understanding the genetics and neurobiological pathways behind addiction. Exp Ther Med. 2021;21(5):1–10.
54. Bawor M, Dennis BB, Bhalerao A, Plater C, Worster A, Varenbut M, et al. Sex differences in outcomes of methadone maintenance treatment for opioid use disorder: a systematic review and meta-analysis. Can Med Assoc Open Access J. 2015;3(3):E344–51.
55. Huhn AS, Berry MS, Dunn KE. Sex-based differences in treatment outcomes for persons with opioid use disorder. Am J Addict. 2019;28(4):246–61.
56. Ngun TC, Ghahramani N, Sánchez FJ, Bocklandt S, Vilain E. The genetics of sex differences in brain and behavior. Front Neuroendocrinol. 2011;32(2):227–46.
57. Dennis BB, Sanger N, Bawor M, Naji L, Plater C, Worster A, et al. A call for consensus in defining efficacy in clinical trials for opioid addiction: combined results from a systematic review and qualitative study in patients receiving pharmacological assisted therapy for opioid use disorder. Trials. 2020;21(1):1–16.
58. Schuckit MA. Treatment of opioid-use disorders. N Engl J Med. 2016;375(4):357–68.

5.1.8.1 Figures

Figure 1. Continued opioid use Manhattan Plot

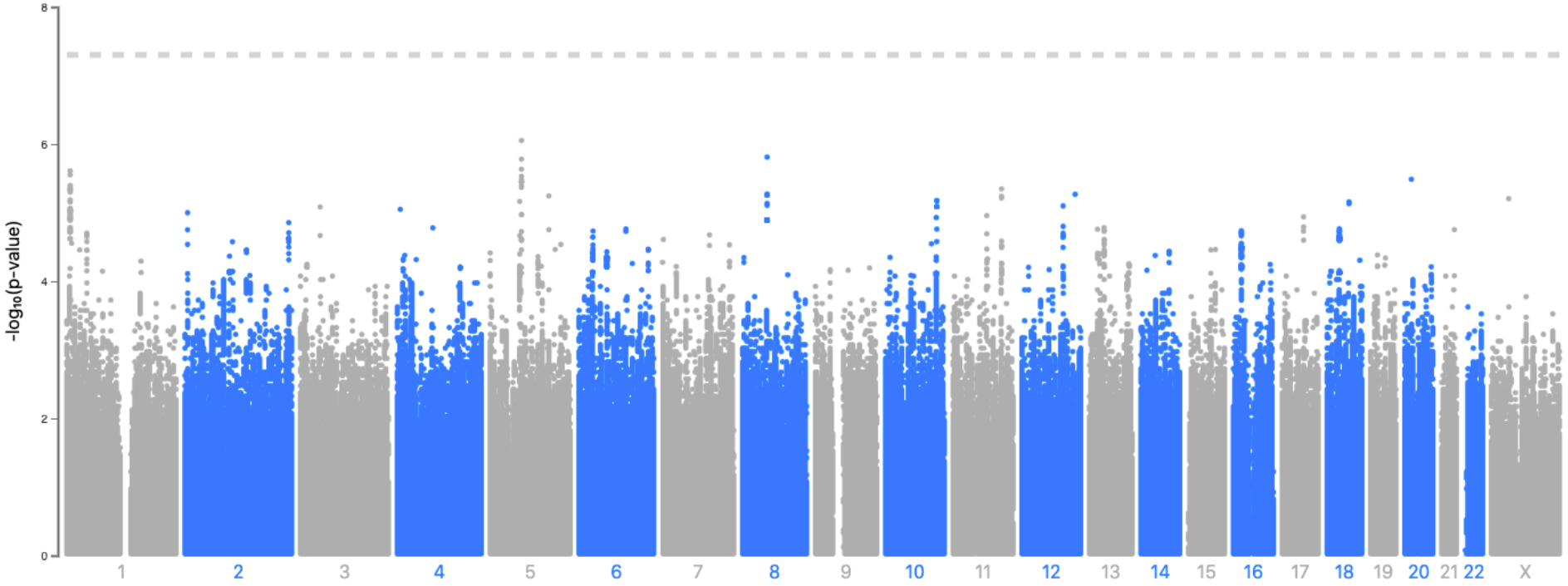
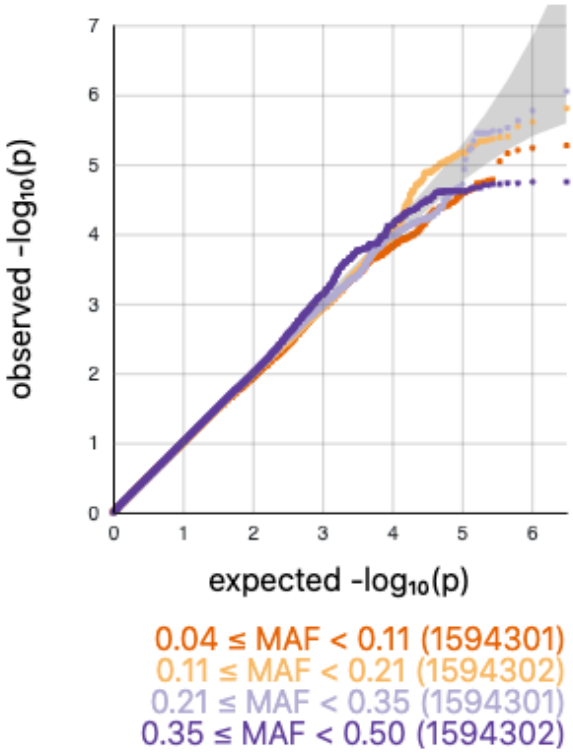


Figure 2. Continued opioid use QQ Plot



GC lambda 0.5: 0.994

Figure 3. Relapse Manhattan Plot

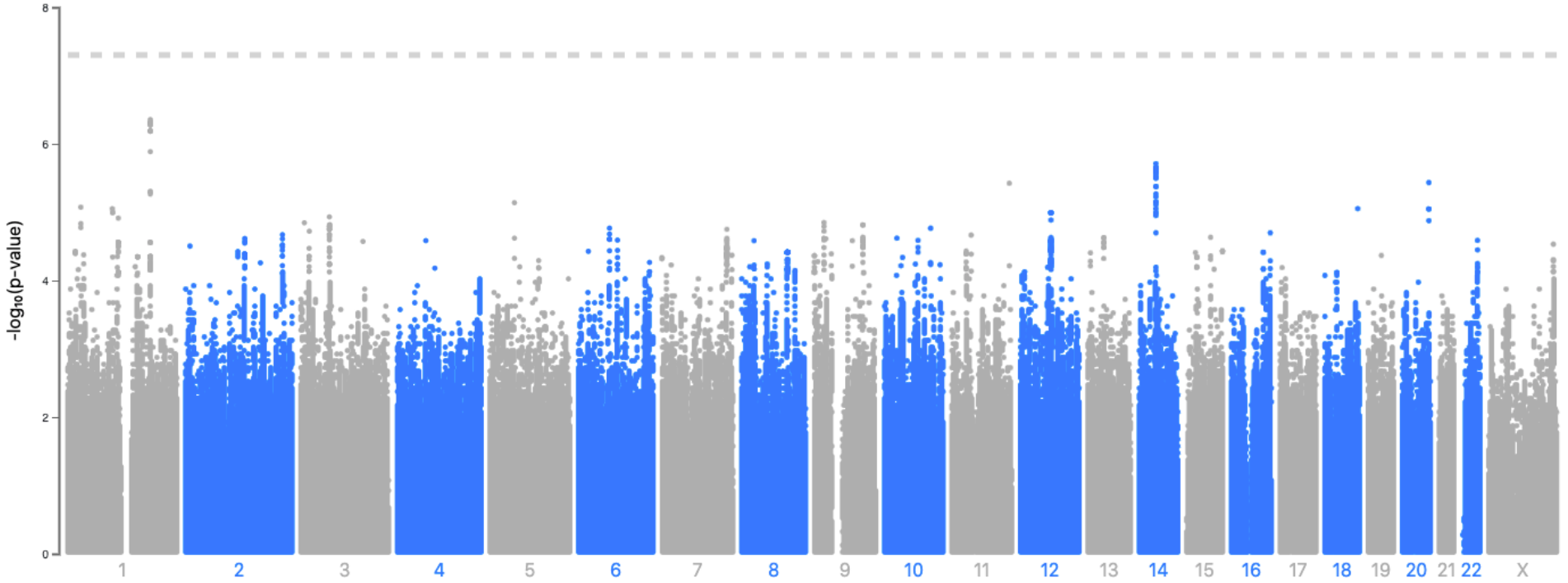
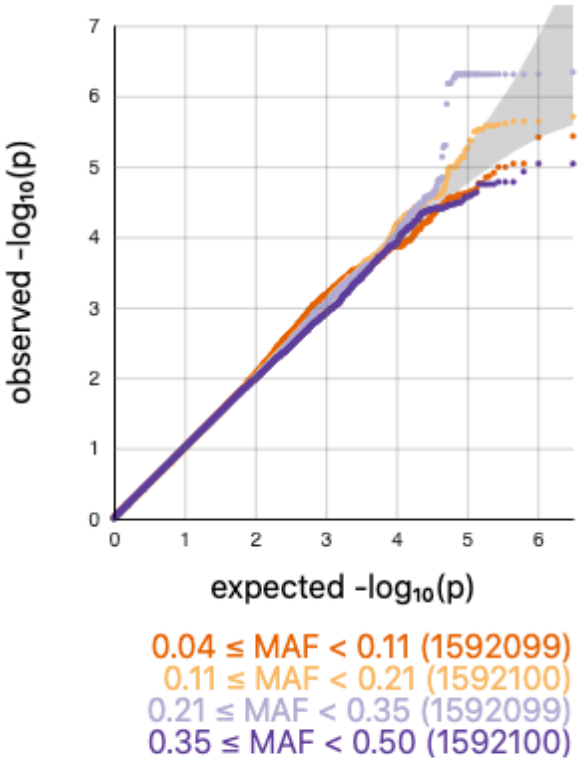


Figure 4. Relapse QQ Plot



GC lambda 0.5: 1.008

Figure 5. Methadone dose Manhattan Plot

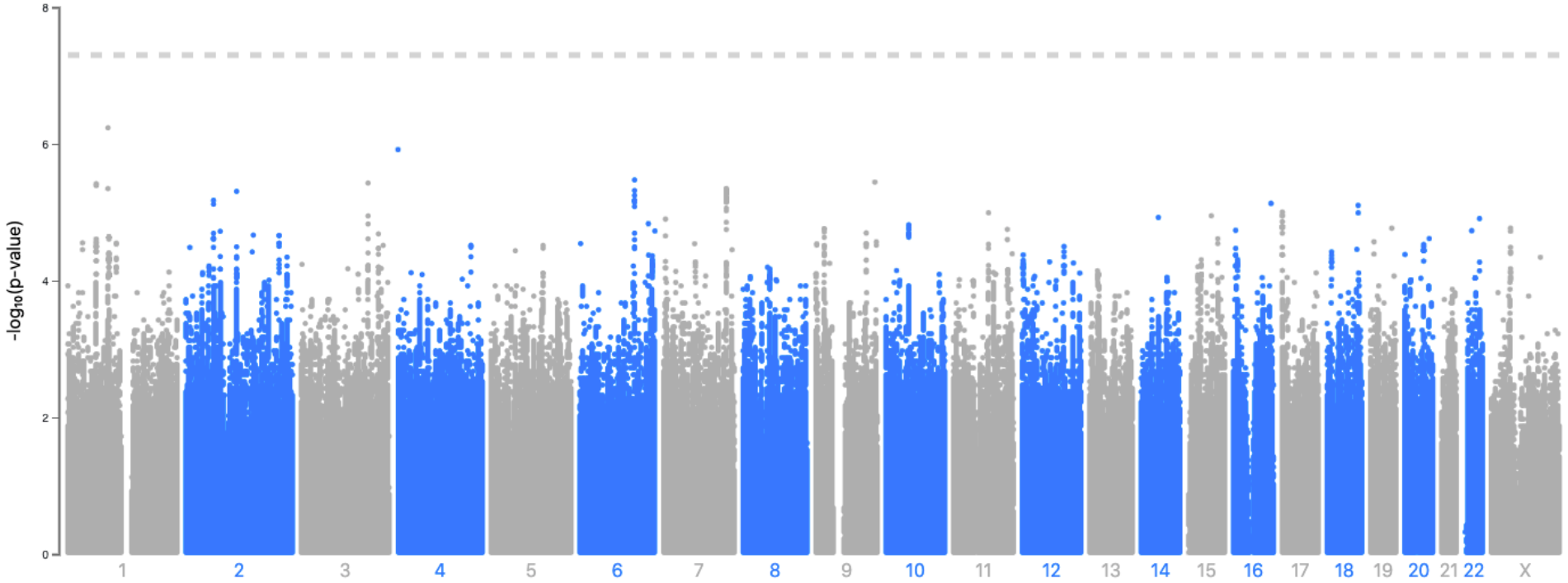


Figure 6. Methadone dose QQ Plot

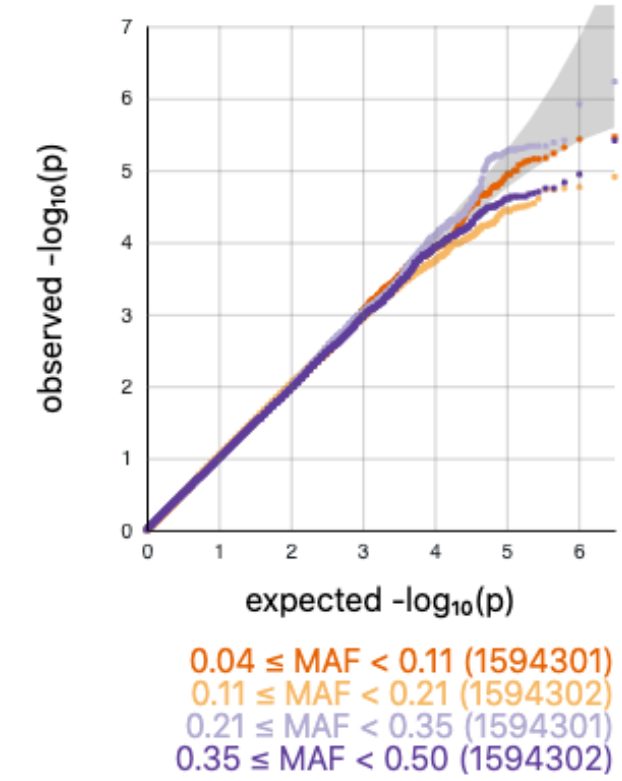


Figure 7. Opioid overdose Manhattan Plot

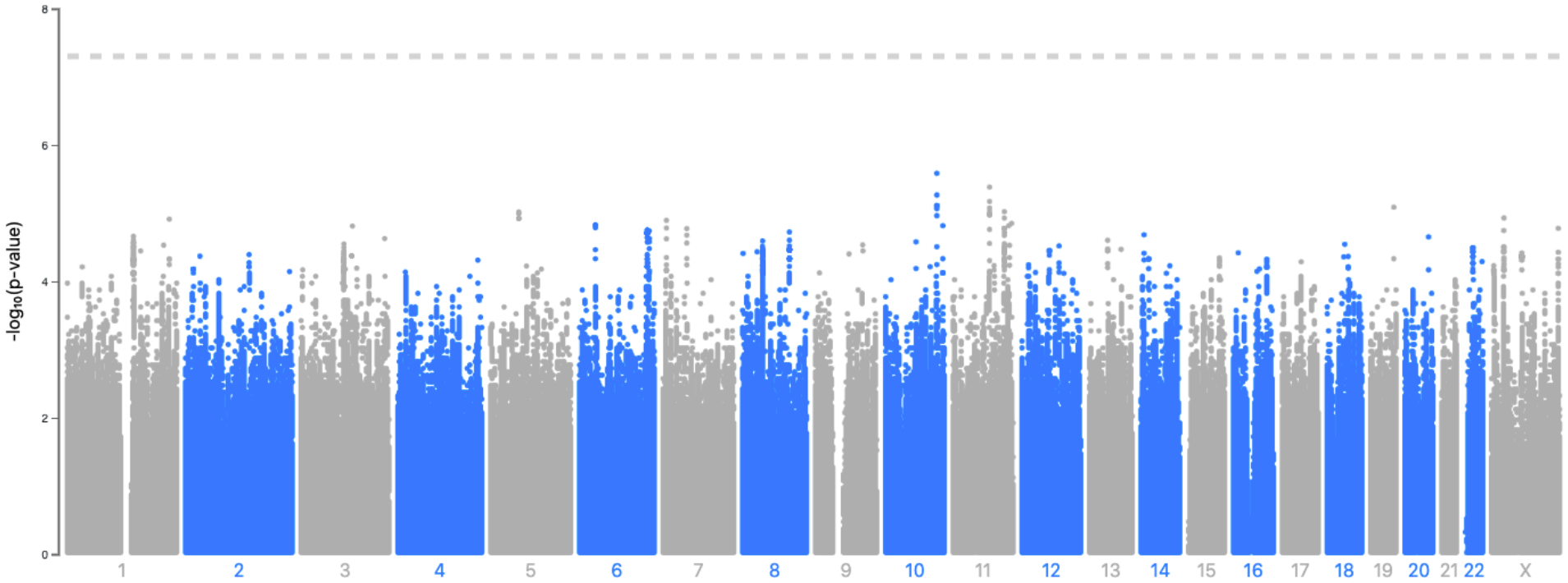


Figure 8. Opioid overdose QQ Plot

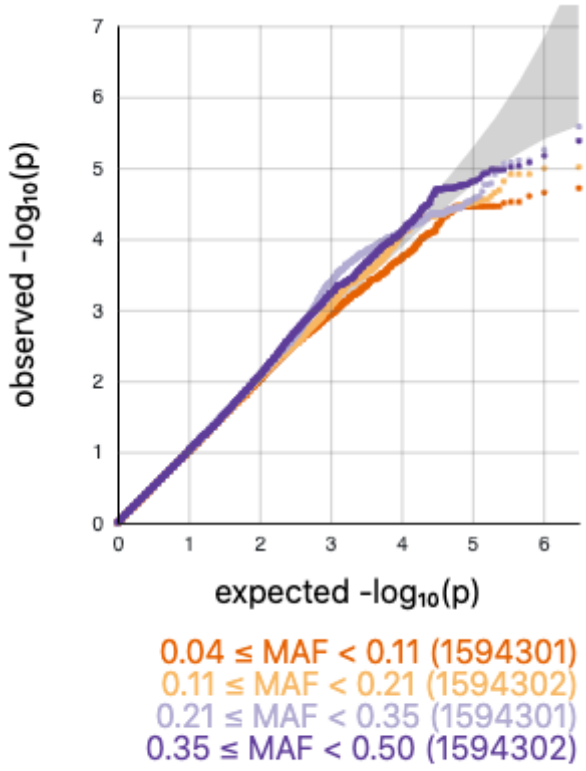


Figure 9. Continued opioid use PRS model fit

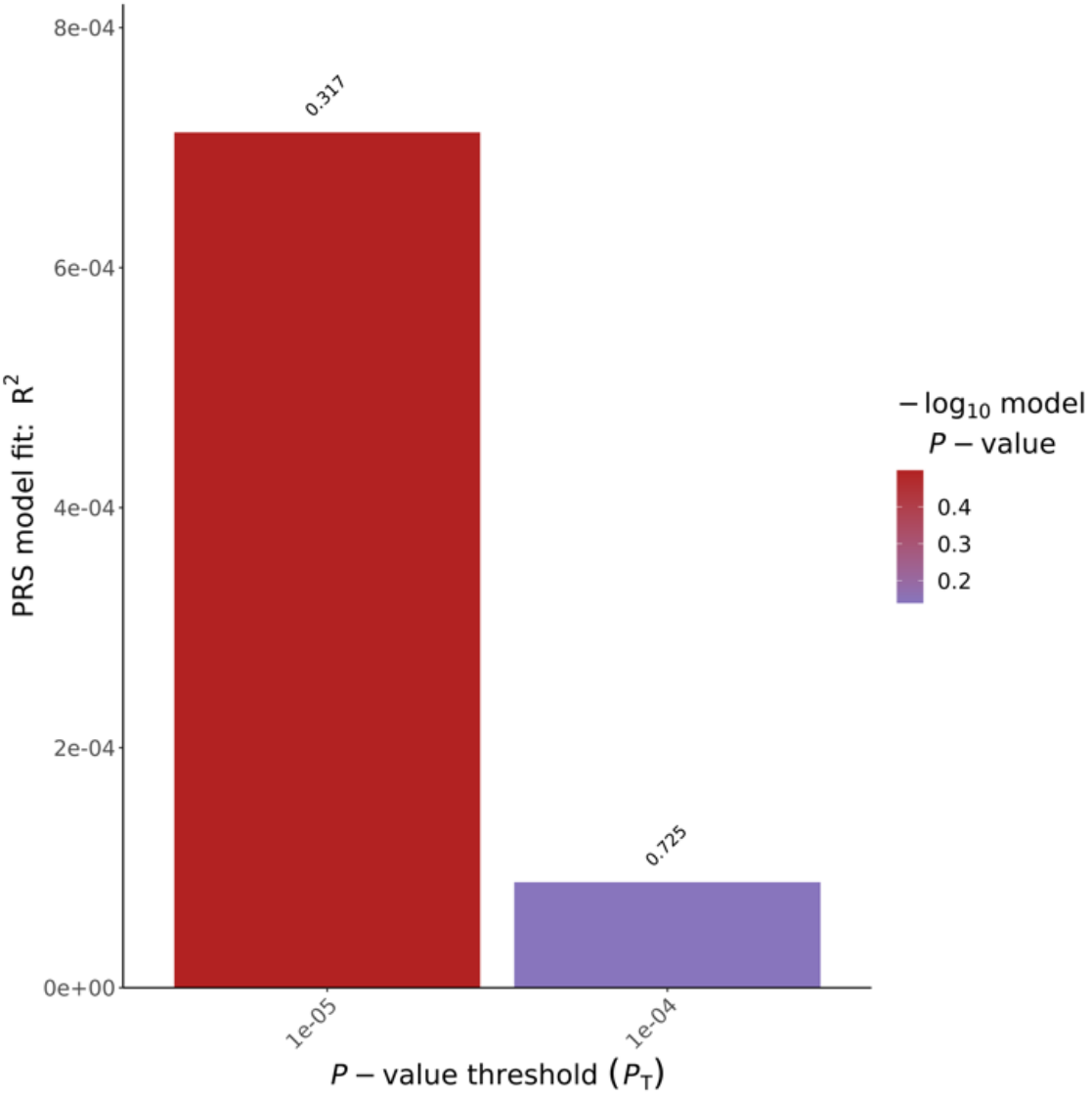


Figure 10. Relapse PRS model fit

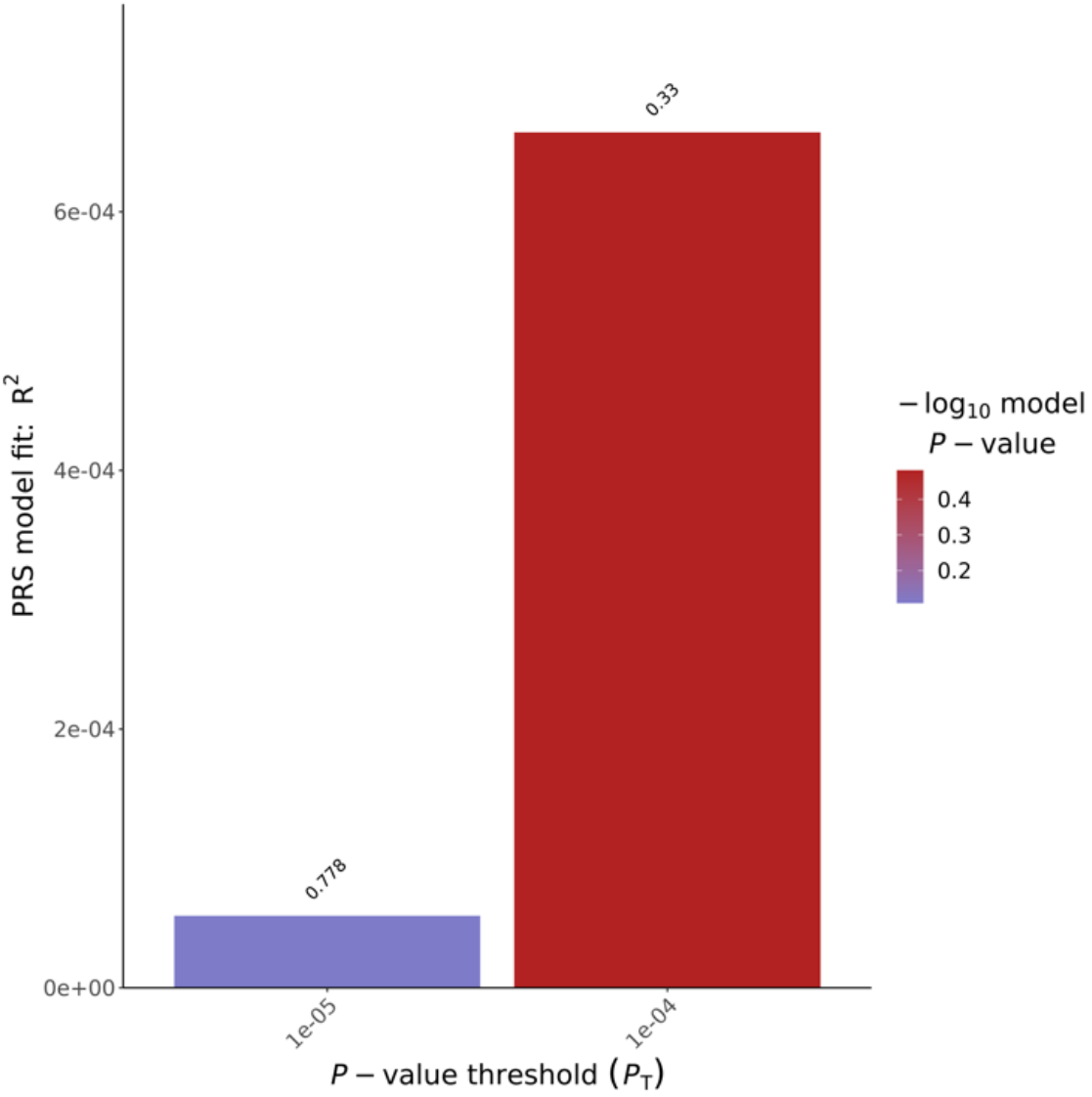


Figure 11. Methadone dose PRS model fit

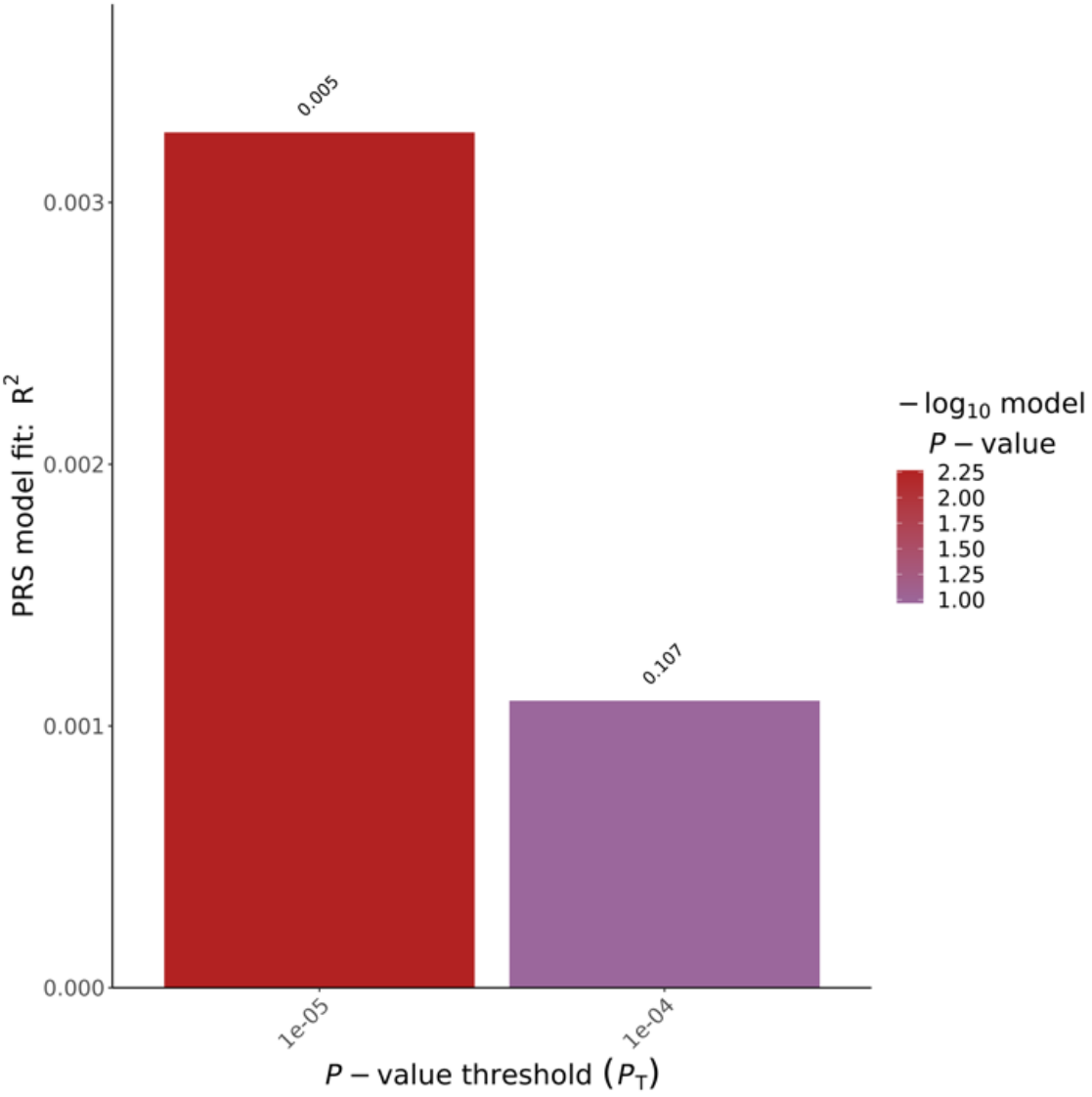
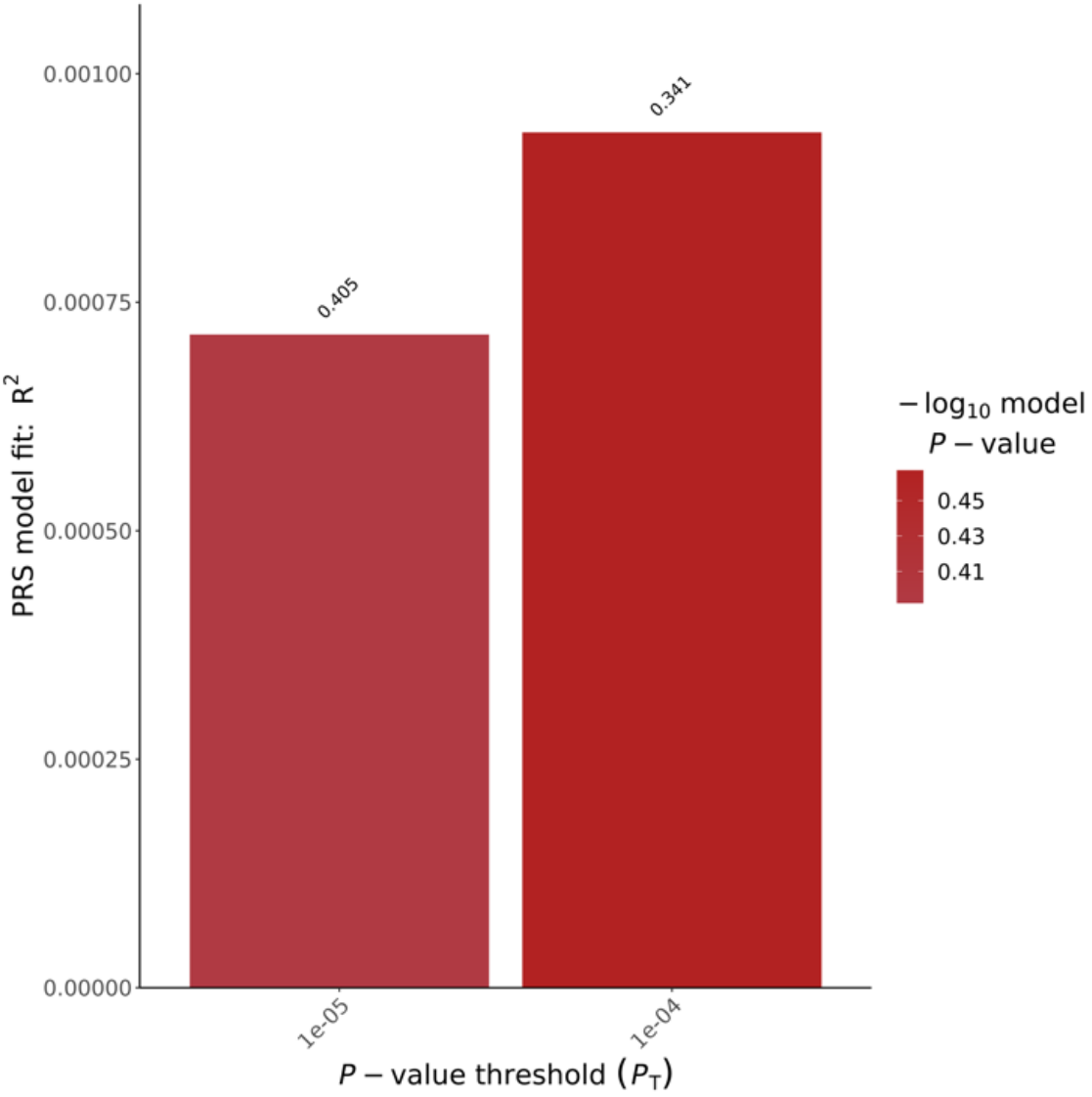


Figure 112. Opioid overdose PRS model fit



6 CHAPTER 6: Conclusion

6.1 Overview

Through this thesis, we have explored the genetics of cannabis use and MMT outcomes. First, we methodologically planned a systematic review and meta-analysis of genetic studies of cannabis use (Chapter 2) and implemented it, while carefully assessing the quality of the findings (Chapter 3). We identified 96 genetic variants associated with cannabis dependence, CUD, age of onset of cannabis use, or lifetime cannabis use and multiple genes of interest, including *ANKFN1*, *INTS7*, *PI4K2B*, *CSMD1*, *CST7*, *ACSS1*, and *SCN9A*. The results from the systematic review were able to summarize the current literature of GWASs investigating cannabis use, and highlight potential regions of interest to inform future genetic studies. Next, we employed GWAS and PRS methods to investigate the genetics of cannabis (Chapter 4) and MMT outcomes (Chapter 5). With GWAS investigating novel genetic variants, PRSs creating an individuals weighted genetic risk score for a trait and an exploration into sex-differences, we were able to provide further insight into the genetics of cannabis use and MMT outcomes, with possible future implications for personalized care. Chapter 4 did not identify any regions reaching genome-wide significance or sex-specific results, however, did find statistically significant PRS for regular cannabis with minor variance explained (R^2=2.50E-3). Similarly, Chapter 5 did not identify any regions of genome-wide significance or sex-specific results and found a statistically significant PRS score for methadone dose with minor variance explained (R^2=3.45E-3).

6.2 Overall Implications

The evidence provided in this systematic review significantly contributes to the genetics of cannabis use and MMT outcomes. The systematic review and meta-analysis provide an overview of the current literature on GWASs investigating cannabis use while maintaining transparency in reporting. The GWAS and PRS on cannabis use provided insight into the genetics of cannabis use within the OUD population, a unique population of interest in which cannabis use is prevalent. The GWAS and PRS on MMT outcomes provided insight into the genetics of continued opioid use, relapse, methadone dose and opioid overdose. Finally, the genetic studies presented were some of the first to introduce sex differences to the analysis and analyzing same-sex groups independently.

Importantly, this thesis conducted research on a vulnerable population, that is, individuals living with an OUD. While the results within this thesis reaching statistical significance are limited, this thesis does provide more information for opioid addiction clinicians that may help improve treatment for patients. We have identified background information on the importance of sex-differences for both cannabis use and methadone dose. Additionally, we provide important information on the genetic predispositions that might play a role in cannabis use and response to MMT. Finally, with the support of future research, we aid in the advancement of personalized medicine through evidence which demonstrates that an individual's genetic makeup may impact MMT treatment response. Personalized medicine involves utilizing individual characteristics, including

their genetics, to guide treatment plans with the aim of better outcomes (39). In cannabis use and response to MMT, personalized medicine could include informing individuals of their predisposed genetic risk to certain substances or reduce the likelihood of exposing individuals to a treatment that would be ineffective to them (40).

Finally, in addition to disseminating information to opioid addiction clinicians, this thesis aims to disseminate information to researchers and policy makers. Through publishing the systematic review protocol and full review, as well as intentions to submit the GWASs and PRSs for cannabis use and MMT outcomes, this work promotes research transparency and encourages future research.

6.3 Future Directions

While the PRSs conducted had a sufficient sample size, the GWASs were underpowered. Thus, although the GWASs provided SNPs approaching genome-wide significance, it is recommended for larger, more powerful, GWASs to be conducted within the OUD population. Additionally, given the low variance explained by the PRSs, it is recommended to test additional previous GWAS on traits associated with our outcomes, such as externalizing behaviours, to explore whether the lack of genetic variance explained may be due to genetic pleiotropy of different risk-taking traits. Finally, it is important to continue to test sex-differences within genetics due to the known phenotypic differences between males and females, and the possibility of personalized care.

7 Reference List

1. The Government of Canada. Cannabis Stats Hub. 2019.
2. Compton WM, Valentino RJ, DuPont RL. Polysubstance use in the US opioid crisis. Mol Psychiatry. 2021;26(1):41–50.
3. Bawor M, Dennis BB, Varenbut M, Daiter J, Marsh DC, Plater C, et al. Sex differences in substance use, health, and social functioning among opioid users receiving methadone treatment: A multicenter cohort study. Biol Sex Differ [Internet]. 2015;6(1):1–11. Available from: http://dx.doi.org/10.1186/s13293-015-0038-6
4. Saxon AJ, Calsyn DA, Greenberg D, Blaes P, Haver VM, Stanton V. Urine screening for marijuana among methadone-maintained patients. Am J Addict. 1993;2(3):207–11.
5. Reiman A. Cannabis as a substitute for alcohol and other drugs. Harm Reduct J. 2009;6(1):35.
6. Ashrafioun L, Bohnert KM, Jannausch M, Ilgen MA. Characteristics of substance use disorder treatment patients using medical cannabis for pain. Addict Behav. 2015;42:185–8.
7. Wasserman DA, Weinstein MG, Havassy BE, Hall SM. Factors associated with lapses to heroin use during methadone maintenance. Drug Alcohol Depend. 1998;52(3):183–92.
8. Proctor S, Copeland A, Kopak A, Hoffmann N, Herschman P, Polukhina N. Outcome predictors for patients receiving methadone maintenance treatment: findings from a retrospective multi-site study. J Substamce Use. 2016;21(6):601–13.
9. Lucas P, Walsh Z, Crosby K, Callaway R, Belle-Isle L, Kay R, et al. Substituting cannabis for prescription drugs, alcohol and other substances among medical cannabis patients: The impact of contextual factors. Drug Alcohol Rev. 2016;35(3):326–33.
10. Lucas P, Baron EP, Jikomes N. Medical cannabis patterns of use and substitution for opioids & other pharmaceutical drugs, alcohol, tobacco, and illicit substances; results from a cross-sectional survey of authorized patients. Harm Reduct J. 2019;16(1):9.
11. Lau N, Sales P, Averill S, Murphy F, Sato S, Murphy S. A safer alternative: Cannabis substitution as harm reduction. Drug Alcohol Rev. 2015;34(6):654–9.
12. Epstein DH, Preston KL. Does cannabis use predict poor outcome for heroin-dependent patients on maintenance treatment? Past findings and more evidence against. Addiction. 2003;98(3):269–79.
13. Peirce, J Mp, N Mr, J Mk, KKrasnansky, JStabile, et al. Correlates of stimulant treatment outcome across treatment modalities. Am J Drug Alcohol Abuse [Internet]. 2009;35(1 CC-Drugs and Alcohol):48-53. Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01621014/full
14. Cichewicz DL. Synergistic interactions between cannabinoid and opioid analgesics. Life Sci. 2004;74(11):1317–24.

15. Saxon AJ, Wells EA, Fleming C, Jackson TR, Calsyn DA. Pre-treatment characteristics, program philosophy and level of ancillary services as predictors of methadone maintenance treatment outcome. Addiction. 1996;91(8):1197–210.
16. Nirenberg TD, Cellucci T, Liepman MR, Swift RM, Sirota AD. Cannabis versus other illicit drug use among methadone maintenance patients. Psychol Addict Behav. 1996;10(4):222.
17. World Health Organization. Opioid overdose [Internet]. 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/opioid-overdose
18. Strang J, Volkow ND, Degenhardt L, Hickman M, Johnson K, Koob GF, et al. Opioid use disorder. Nat Rev Dis Prim. 2020;6(1):1–28.
19. The Government of Canada. Opioid- and Stimulant-related Harms in Canada. 2022.
20. Morin KA, Eibl JK, Franklyn AM, Marsh DC. The opioid crisis: past, present and future policy climate in Ontario, Canada. Subst Abuse Treat Prev Policy. 2017;12(1):1–7.
21. Coffa D, Snyder H. Opioid use disorder: medical treatment options. Am Fam Physician. 2019;100(7):416–25.
22. Joseph H, Stancliff S, Langrod J. Methadone maintenance treatment (MMT): A review of historical and clinical issues. Mt Sinai J Med. 2000;67(5):347–64.
23. Marsch LA. The efficacy of methadone maintenance interventions in reducing illicit opiate use, HIV risk behavior and criminality: a meta-analysis. Addiction. 1998;93(4):515–32.
24. Eap CB, Buclin T, Baumann P. Interindividual variability of the clinical pharmacokinetics of methadone. Clin Pharmacokinet. 2002;41(14):1153–93.
25. Somogyi AA, Barratt DT, Ali RL, Coller JK. Pharmacogenomics of methadone maintenance treatment. Pharmacogenomics. 2014;15(7):1007–27.
26. Prevention C for DC and. Calculating total daily dose of opioids for safer dosage. 2017.
27. Goud A, Kiefer E, Keller MS, Truong L, SooHoo S, Riggs R V. Calculating maximum morphine equivalent daily dose from prescription directions for use in the electronic health record: a case report. JAMIA open. 2019;2(3):296–300.
28. Huang C, Lee C. Factors associated with mortality among heroin users after seeking treatment with methadone: a population-based cohort study in Taiwan. J Subst Abus Treat. 2013;44(3):295–300.
29. Hillmer A, Chawar C, D'Elia A, Kapoor R, Butt M, Samaan Z. A systematic review and meta-anlaysis of genome-wide association studies and cannabis use. PROSPERO: International prospective register of systematic reivews. 2020.
30. Agrawal A, Lynskey MT. Candidate Genes for Cannabis Use Disorders: Findings, Challenges and Directions. Addiction. 2009;104(4):518–32.
31. Stringer S, Minică CC, Verweij KJH, Mbarek H, Bernard M, Derringer J, et al. Genome-wide association study of lifetime cannabis use based on a large meta-analytic sample of 32 330 subjects from the International Cannabis Consortium. Transl Psychiatry. 2016;6(3):e769.
32. Agrawal A, Chou YL, Carey CE, Baranger DAA, Zhang B, Sherva R, et al.

Genome-wide association study identifies a novel locus for cannabis dependence. Mol Psychiatry. 2018;23(5):1293–302.
33. Minică CC, Verweij KJH, Most PJ, Mbarek H, Bernard M, Eijk KR, et al. Genome-wide association meta-analysis of age at first cannabis use. Addiction [Internet]. 2018 Nov;113(11):2073–86. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=132211589&site=ehost-live
34. Pasman JA, Verweij KJH, Gerring Z, Stringer S, Sanchez-Roige S, Treur JL, et al. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. Nat Neurosci [Internet]. 2018;21(9):1161–70. Available from: http://dx.doi.org/10.1038/s41593-018-0206-1
35. Johnson EC, Demontis D, Thorgeirsson TE, Walters RK, Polimanti R, Hatoum AS, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. The Lancet Psychiatry. 2020;7(12):1032–45.
36. Berrettini W. A brief review of the genetics and pharmacogenetics of opioid use disorders. Dialogues Clin Neurosci. 2022;
37. Chawar C, Hillmer A, Sanger S, D'Elia A, Panesar B, Guan L, et al. A systematic review of GWAS identified SNPs associated with outcomes of medications for opioid use disorder. Addict Sci Clin Pract. 2021;16(1):1–14.
38. Fonseca F, Torrens M. Pharmacogenetics of methadone response. Mol Diagn Ther. 2018;22(1):57–78.
39. Hamburg MA, Collins FS. The path to personalized medicine. N Engl J Med. 2010;363(4):301–4.
40. Oslin D. Personalized addiction treatment: how close are we? Alcohol Alcohol. 2011;46(3):231–2.

8 Appendix

8.1 CHAPTER 2 Additional File 1

PRISMA-P 2015 Checklist

This checklist has been adapted for use with systematic review protocol submissions to BioMed Central journals from Table 3 in Moher D et al: Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 2015 **4**:1

An Editorial from the Editors-in-Chief of *Systematic Reviews* details why this checklist was adapted - Moher D, Stewart L & Shekelle P: Implementing PRISMA-P: recommendations for prospective authors. *Systematic Reviews* 2016 **5**:15

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
ADMINISTRATIVE INFORMATION					
Title					
Identification	1a	Identify the report as a protocol of a systematic review	☒	☐	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	☐	☐	NA
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number in the Abstract	☒	☐	2
Authors					
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	☒	☐	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	☒	☐	15
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	☒	☐	15
Support					
Sources	5a	Indicate sources of financial or other support for the review	☒	☐	14
Sponsor	5b	Provide name for the review funder and/or sponsor	☐	☐	NA
Role of sponsor/funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	☐	☐	NA
INTRODUCTION					
Rationale	6	Describe the rationale for the review in the context of what is already known	☒	☐	3-4

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	☒	☐	4-5
METHODS					
Eligibility criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics (e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	☒	☐	6-7
Information sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	☒	☐	7
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	☒	☐	Additional File 2
STUDY RECORDS					
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	☒	☐	8
Selection process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	☒	☐	8
Data collection process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	☒	☐	8
Data items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	☒	☐	8-9
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	☒	☐	9-10
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	☒	☐	10
DATA					
Synthesis	15a	Describe criteria under which study data will be quantitatively synthesized	☒	☐	10-11
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I^2, Kendall's tau)	☒	☐	11

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-regression)	☒	☐	11
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	☒	☐	11
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	☒	☐	11-12
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	☒	☐	12

8.2 CHAPTER 2 Additional File 2

Search Strategy

MEDLINE	15. Genome-Wide Association Study/ 16. Genotyping Techniques/ 17. Genome, Human/ 18. Genetic Variation/ 19. genetics/ or exp human genetics/ 20. (human* adj2 (genotyp* or genome* or genetic*)).ti,ab,kw,kf. 21. (GWS or GWAS or GWA).mp. 22. genome wide.ti,ab,kw,kf. 23. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 24. exp Cannabis/ 25. ((cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) adj2 (overdose* or use* or using or misuse* or abus* or dependence* or addict*)).ti,ab,kw,kf. 26. 10 or 11 27. 9 and 12 *28. Limit 13 to humans*
Web of Science	8. TS=(genome-wide association study or genome-wide association or GWAS or GWA or genome wide) 9. TS=(human NEAR/2 genome) 10. TS=((cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) NEAR/2 (overdose* or use* or using or misuse* or abus* or dependence* or addict*)) 11. TS=(cannabis* or marijuana* or marihuana*) 12. #1 OR #2 13. #3 OR #4 14. #5 and #6
EMBASE	15. Genome-Wide Association Study/ 16. Genotyping Techniques/

	17. Genome, Human/ 18. Genetic Variation/ 19. genetics/ or exp human genetics/ 20. (human* adj2 (genotyp* or genome* or genetic*)).ti,ab,kw. 21. (GWS or GWAS or GWA).mp. 22. genome wide.ti,ab,kw. 23. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 24. exp Cannabis/ 25. ((cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) adj2 (overdose* or use* or using or misuse* or abus* or dependence* or addict*)).ti,ab,kw. 26. 10 or 11 27. 9 and 12 *28. Limit 13 to human*
CINAHL	7. genome-wide association study or genome-wide association or GWAS or GWA or genome wide or genome 8. cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) 9. overdose* or use* or using or misuse* or abus* or dependence* or addict* 10. S2 and S3 11. S1 and S4 12. *Limit to Human*
GWAS Catalog	Terms Searched: - Cannabis - Cannabis dependence - Marihuana - Marijuana - Cannabinoids - Hash - Kush - Weed - Pot - THC - CBD
GWAS Central	Terms Searched:

	- Cannabis - Cannabis dependence - Marijuana - Marihuana - Cannabinoids - Hash - Kush - Weed - Pot - THC - CBD
NIH Database of Genotypes and Phenotypes	Terms Searched: - Cannabis - Cannabis dependence - Marijuana - THC - Marihuana - Cannabinoids - Hash - Kush - Weed - Pot - CBD

8.3 CHAPTER 2 Published Protocol

Hillmer *et al. Systematic Reviews* (2020) 9:190
https://doi.org/10.1186/s13643-020-01442-2

Systematic Reviews

PROTOCOL Open Access

Genetic determinants of cannabis use: a systematic review protocol

Alannah Hillmer[1], Caroul Chawar[1], Stephanie Sanger[2], Alessia D'Elia[1], Mehreen Butt[3], Raveena Kapoor[4], Flavio Kapczinski[5], Guillaume Pare[6], Lehana Thabane[7] and Zainab Samaan[5*]

Abstract

Background: With the legalization of cannabis in Canada, there is an increase trend in use. Cannabis has been known to have several health implications, one of which is the development of cannabis use disorder (CUD). CUD is more common in males than females, as well as in certain ethnic groups such as Native Americans. Additionally, both environmental and genetic risk factors have been found for cannabis use. The objective of this systematic review will be to summarize the genetic variants associated with cannabis use which have reached borderline genome-wide significance.

Methods: This systematic review will incorporate articles that have performed a genome-wide association study (GWAS) investigating cannabis use. MEDLINE, Web of Science, EMBASE, GWAS Catalog, GWAS Central, and NIH Database of Genotype and Phenotype will be searched using a comprehensive search strategy. The quality of genetic association studies (Q-Genie) tool will be utilized to assess the quality of the included studies. All screening and data extraction will occur independently by two authors. If feasible, a random-effects meta-analysis will be conducted on pooled odds ratios of single nucleotide polymorphisms reaching borderline genome-wide significance.

Discussion: This systematic review will synthesize available GWAS on cannabis use. Results from this review will inform and direct further investigation of genetic variants associated with cannabis use.

Systematic review registration: PROSPERO CRD42020176016

Keywords: Systematic review, Cannabis, Genetics, Genome-wide

Background

On October 17, 2018, the Cannabis Act came into effect in Canada allowing for the legal growth of cannabis plants as well as the recreational possession and consumption of cannabis for those who are 18 years or older [1]. In response to the Cannabis Act, Statistics Canada has introduced a National Cannabis Survey which has been conducted every 3 months since February 2018. The NCS showed that nearly 17% of Canadians aged 15 years and older reported using cannabis within a 3-month period between mid-August and mid-September of 2019, a rate that was consistent with the rate of the year prior, when cannabis was an illicit substance. However, in the fourth quarter of 2019, cannabis use was increased when compared to the fourth quarter of 2018. Additionally, regardless of the year of study, cannabis consumption rates continue to be higher among males than females [2].

Cannabis use disorder (CUD) is defined as a problematic pattern of cannabis use leading to clinically significant impairment or distress. In 2013, the Diagnostic and Statistical Manual reported that CUD is prevalent in 3.4% of youth aged 12 to 17 years old and 1.5% of adults age 18 years or older. Trends of CUD also differ among

* Correspondence: samaanz@mcmaster.ca
[5]Department of Psychiatry and Behavioural Neurosciences, McMaster University, 100 West 5th St., Hamilton, ON L8N 3 K7, Canada
Full list of author information is available at the end of the article

sex and ethnicities. Rates of CUD are higher in males compared to females and rates of CUD are higher in Native American and Alaska Natives compared to other ethnic groups [3]. Results from a meta-analysis on twin studies estimated the heritability for cannabis use initiation to be 40–48% and 51–59% for problematic cannabis use, suggesting a genetic component to cannabis use and CUD [4]. A genome-wide association study (GWAS) combined five cohorts identifying several genes and single nucleotide polymorphisms (SNPs) associated with cannabis use and dependence [5]. A cluster of correlated SNPs in a novel region of chromosome 10 were identified at genome-wide significant levels in participants of European descent [5]. However, of three meta-analyses conducted on cannabis use in the literature, only one study identified a significate association [6–8]. One region on chromosome 16 was significantly associated with age of first cannabis use, with the strongest association for the intronic variant rs1574587 [7].

Interestingly, one study investigated the genetic and environmental risk factors for cannabis availability reported variation in cannabis initiation and symptoms of cannabis use disorder. Cannabis availability and initiation had a correlation of 0.48 and cannabis availability and symptoms of cannabis use disorder had a correlation of 0.23. Additionally, much of the variation associated with problematic use can be explained by shared environmental risk in cannabis availability leading to initiation and the genetic non-shared environmental risks for cannabis initiation [9]. These findings are of specific interest to Canada and other countries with legalization of cannabis is already in effect or being considered, as cannabis is increasingly more available since the legalization.

With cannabis availability increasing, and known heritability of CUD, it is important to understand the genetic risk factors associated with cannabis use. While meta-analyses of GWASs provide regions of interest, no known systematic review exists that summarizes identified genes and/or SNPs that have reached genome-wide significance for cannabis use. It is important to provide a summary of the literature which includes recent GWASs in the context of cannabis legalization. Further, understanding the genetic basis of cannabis use will assist health care workers in making science-informed decisions regarding the recommendation of recreational use and prescription of cannabis.

Objectives

The main goal of this systematic review is to identify genetic variants from genome-wide association studies (GWASs) associated with cannabis use. Though genetic variants most commonly reported by GWASs are SNPs, this review will be inclusive of any other genetic markers reported in GWASs. We will summarize the results of GWASs which meet our inclusion criteria, and if possible, we will meta-analyze genetic variants that are reported in more than one primary study.

Primary objectives of this systematic review include the following:

1. Identify genetic variants associated with current cannabis use. Current cannabis use is defined by either self-report or positive urine drug screens within 1 month of the study being conducted.
2. Identify genetic variants associated with lifetime cannabis use. Lifetime cannabis use is defined by any self-reported or positive urine drug screens of cannabis use within one's lifetime.
3. Identify genetic variants associated with CUD. CUD is defined by any diagnostic and classification systems used to diagnosis CUD or questionnaires validated to assess CUD.

Secondary objectives of this systematic review include the following:

1. Identify genetic variants associated with the adverse outcomes of cannabis use including psychiatric (cognitive impairment, psychotic symptoms, depression, anxiety, suicidal behavior) and non-psychiatric (chronic bronchitis, lung infections, chronic cough, increased risk of motor vehicle accidents) [10–12].
2. When feasible, perform subgroup summaries by sex or ethnic differences.

Methods and analysis

This protocol is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) statement [13] (see PRISMA-P checklist in Additional file 1). This protocol was registered within the International Prospective Register of Systematic Reviews (PROSPERO) (registration number: CRD42020176016).

Eligibility criteria

GWAS studies presenting original data on associations between cannabis use and genetic polymorphisms using any study design (i.e., case-control, cohort, etc.) will be included in this systematic review. All other types of studies will be excluded. Studies in any setting will be included and no restriction will be placed on age, sex, ethnic background, or language. Additionally, articles that do not present sufficient data to calculate the odds ratio (OR) with a 95% confidence interval will be excluded from quantitative analyses if data cannot be obtained after contacting the studies' authors and the calculations cannot be made with the available published information. However,

we will include these studies in the qualitative description of the review findings.

We will include studies investigating cannabis use disorder as defined by the Diagnostic and Statistical Manual-5 (DSM-5) or other diagnostic and classification systems such as the International Statistical Classification of Diseases and Related Health Problems-10 (ICD-10) or specific diagnostic scales designed to screen and diagnose dependence or use disorder of cannabis as well as any studies measuring any use of cannabis. We define cannabis use based on the included studies' definitions and accept the following definition: current cannabis use is defined as either self-report or positive urine drug screens within 1 month of the study being conducted and lifetime cannabis use is defined as any self-reported or positive urine drug screens of cannabis use within one's lifetime [14]. Clinical diagnoses and questionnaires validated to assess CUD will also be accepted. All studies not investigating current cannabis use, lifetime cannabis use, or CUD will be excluded. In the case of polymorphisms reported in duplicate publications from the same study population, the article that is the most recent will be included.

Information sources

A Health Science Librarian was consulted to develop a comprehensive search strategy. No language restriction will be placed on the search strategy, though studies will be limited to human studies. MEDLINE, Web of Science, EMBASE, GWAS Catalog, GWAS Central, and NIH Database of Genotype and Phenotype will be searched using the agreed-upon strategy, modified for each database. The search strategy will include all terms relevant to cannabis and genome-wide association studies. Databases will be searched from inception onwards. Sources of gray literature including dissertations and theses, clinical guidelines, and reports from regulatory agencies will be searched. Reference lists of relevant systematic reviews and all included studies will be checked to identify additional articles.

Search strategy

Draft search strategies for multiple electronic databases are provided in Additional file 2.

Study records

Data management

All of the references will be managed and organized through Zotero [15]. Covidence will be used for the management of this systematic review at the title and abstract, full text, and data extraction stages [16]. Prior to the formal screening process, a calibration will take place to pilot and refine the screening process. Training will be given to all team members on using Covidence prior to starting the review.

Selection process

Two independent reviewers will screen titles and abstracts for inclusion criteria. Full-text review will also be completed independently by two reviewers. Disagreements between reviewers will be resolved by consensus or including a third reviewer. We will record the reason for excluding studies at the full-text review stage.

Data collection process

Data extraction will take place independently and in duplicate for each eligible study. Standardized full-text data extraction forms will be constructed. The data extraction form will be pilot tested by two independent reviewers to determine the feasibility of this review and ensure all details are captured. In the event of missing data, we will contact study authors to obtain missing information where possible. All contact with the authors will be documented.

Data items

We will extract the following information: author, year of study, country, cohort population used, number of participants (separated by those included in the cannabis use group and non-cannabis use group), control population, the ethnicity of participants, mean age, sex ratio, the measure of cannabis use disorder or cannabis use or definition of cannabis use, inclusion and exclusion criteria, how cannabis use was reported (i.e., self-report, drug urine screens), frequency of cannabis use, and finally any genetic variants which reached the significance threshold set of $p \leq 10^{-7}$. Genome-wide significance is generally considered any SNP with a p value less than 5×10^{-8}; however, SNPs reaching borderline significance, $p < 10^{-7}$, will also be extracted as borderline significance has been found to be generally replicable [17].

Outcomes and prioritization

The main aim of the systematic review will be to assess variants reaching the given threshold associated with cannabis use outcomes from the primary studies included in this review.

The primary outcomes are as follows:

1. Current cannabis use is defined as either self-reported cannabis use or positive cannabis urine drug screens within 1 month of the study being conducted.
2. Lifetime cannabis use is defined as self-reported ever used cannabis during the individual's lifetime.
3. CUD is defined by a diagnosis from the DSM-5 or other diagnostic and classification system such as

the ICD-10 or specific diagnostic scales designed to screen and diagnose dependence or use disorder of cannabis.

For each of the outcomes above, we will collect information on each outcome as reported in the primary studies meeting the eligibility criteria, including dichotomous use of cannabis, percent positive urine screens, questionnaires, and diagnostic classification.

The secondary outcomes are as follows:

1. Adverse outcomes of cannabis use including psychiatric and non-psychiatric outcomes. We will collect data as reported in the primary studies included such as comorbid diagnosis and additional medication condition.
2. We will collect information from the included primary studies on sex and ethnic groups within the study. We will provide a qualitative summary and, if feasible, conduct a subgroup meta-analysis of genetic variants within specific ethnic groups.

Risk of bias in individual studies

Quality assessment will be completed in duplicate for each study included. The quality of genetic association studies (Q-Genie) tool [version 1.1] will be used. Disagreements of quality assessments will be resolved through discussion [18]. If a consensus is not reached through discussion, a third author will be consulted to resolve the disagreement.

Data synthesis

Studies included in this systematic review will undergo qualitative synthesis. Summary tables will be used which will include the sample size, size of cannabis group and non-cannabis group, sex distribution, mean age, study design, ethnic population, and outcome (current cannabis use, lifetime cannabis use, or CUD). A separate table will be used to display any variants reaching borderline genome-wide significance, the corresponding study it was reported in, the corresponding chromosome and position, minor allele, gene/locus, population size, outcome associated, measure, measure of association value, measure of variability, ethnicity, and p value reported.

Heterogeneity between the studies will be assessed through the I^2 statistic with a 95% confidence interval. We will also report summary tables including the study design, population, and cannabis use measure/definition to describe heterogeneity qualitatively. If appropriate, a random-effects meta-analysis will be conducted on pooled odds ratios for the main outcome previously mentioned. If appropriate, the a random-effects meta-analysis will be conducted on pooled odds ratio for the secondary outcomes previously mentioned as well as a subgroup analyses of the participants' sex and ethnicities. Subgroup analyses by participant's sex account for any differences in cannabis use between sexes which has been previously reported in the literature [19–21]. Additionally, due to genetic differences between ethnicities, genetic associations may be more predominant in certain ethnic groups than others, as such a subgroup analysis will be conducted, if feasible [22]. Studies excluded from the quantitative analysis will be listed and an exclusion reason will be given.

If quantitative methods of analysis are not feasible for both the primary or secondary outcomes due to either low heterogeneity found by the I^2 statistic or qualitative synthesis or no two study reports the same genetic variant, only qualitative synthesis results will be reported. We will not conduct a meta-analysis of individual participant data.

Meta-bias

To help mitigate publication bias conference, abstracts will included, manual searches of references lists will be conducted, and Cochrane Clinical Trail Protocols Registry and ClinicalTrails.gov databases will be searched for relevant clinical trial protocols. Additionally, the GWAS catalog will be manual searched for borderline significant variants associated with current cannabis use, lifetime cannabis use, or CUD to ensure all variants are captured within this review. Authors of conference abstracts will be contacted to determine the stage of the research project and all correspondence will be documented. If the published work was not captured by the search strategy and deemed eligible by two independent reviewers, it will be included. Two independent reviewers will search the references lists of all included studies. Any identified references, deemed eligible by two independent reviewers, will be included.

Confidence in cumulative estimate

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) will be used to assess the strength of evidence. GRADE scores according to the risk of bias, publication bias, consistency, directness, and precision. A score of high-, moderate, low-, or very low-quality evidence will be assigned and summarized in a table [23].

Presenting and reporting of results

The full review will follow the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines with special consideration to the Human Genome Epidemiology Network (HuGENet) guidelines [24]. Although HuGENet reviews typically focus on a single gene, we will present information on each genetic variant-phenotype association reported which will

include the study details, population, findings, and source of data.

Discussion

A lack of consistent evidence exists in the current literature for genetic variants associated with cannabis use. In addition, this is the first known systematic review to synthesize the available evidence on genetic variants associated with cannabis use. The proposed systematic review aims to identify all genetic variants that have reached borderline genome-wide significance associated with cannabis use and CUD. The proposed systematic review will provide an overview of the current literature on the genetics of cannabis, aiding in the genetic understanding of cannabis use. Understanding the genetic contribution to cannabis use and its effects such as cannabis use disorder has the potential to aid medical practitioners in making decisions related to cannabis use for medical reasons and the associated potential risks. Additionally, variants reaching borderline genome-wide significance will be examined in the context of their known or biologically plausible relevance to further our understanding.

Anticipated limitations of this review existed at both the study and review level. Limitations at the study level may include a lack of reporting quality control steps, reporting of variants within linkage disequilibrium, small sample size, and a lack of reporting variants that failed to reach genome-wide significance ($p < 5 \times 10^{-8}$) but may have reached borderline significance levels ($p < 10^{-7}$). At the review level, limitations exist in the expected high heterogeneity, differing outcomes for cannabis use reported in the literature and the exclusion of meta-analysis and candidate gene studies.

On completion of the systematic review, we will publish in a peer-review academic journal to reach both clinical and academic experts in the field. This systematic review will then inform and direct the further investigation of genetic variants associated with cannabis through candidate gene studies.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13643-020-01442-2.

Additional file 1. "PRISMA-P 2015 Checklist" and contains PRISMA-P checklist.

Additional file 2. Search strategy.

Abbreviations

CUD: Cannabis use disorder; DSM-5: Diagnostic and Statistical Manual 5th edition; GRADE: Grading of Recommendations Assessment, Development and Evaluation; GWAS: Genome-wide association study; GWASs: Genome-wide association studies; HuGENet: The Human Genome Epidemiology Network; ICD-10: International Statistical Classification of Diseases and Related Health Problems -10; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols; Q-Genie: The quality of genetic association studies

Amendments

If amendments to this protocol are made, they will be documented and communicated to the journal. A data of amendment, description, and rationale will accompany each amendment.

Authors' contributions

ZS is the guarantor. AH and CC drafted the manuscript. AH, CC, and ZS contributed to the development of the selection criteria, the risk of bias assessment strategy, and data extraction criteria. SS provided expertise in developing the search strategy. All authors read, provided feedback, and approved the final manuscript.

Funding

This study is supported in part by CIHR (grant number PJT-156306). The funding source has no role in the study design, analysis, reporting, or the decision to publish the study.

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Neuroscience Graduate Program, Department of Psychiatry and Behavioural Neurosciences, McMaster University, 100 West 5th St., Hamilton, ON L8N 3 K7, Canada. [2]Health Science Library, McMaster University, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada. [3]Integrated Science Program, McMaster University, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada. [4]Health Sciences Program, McMaster University, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada. [5]Department of Psychiatry and Behavioural Neurosciences, McMaster University, 100 West 5th St., Hamilton, ON L8N 3 K7, Canada. [6]Population Health Research Institute, McMaster University, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada. [7]Department of Health Research Method, Evidence & Impact, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada.

Received: 20 April 2020 Accepted: 31 July 2020
Published online: 20 August 2020

References

1. The Government of Canada. Cannabis legalization and regulation. 2018.
2. Statistics Canada. National Cannabis Survey, third quarter 2019. 2019.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington: American Psychaitric Publishing; 2013.
4. Verweij KJH, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG, et al. Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. Addiction. 2010;105(3):417–30.
5. Agrawal A, Chou YL, Carey CE, Baranger DAA, Zhang B, Sherva R, et al. Genome-wide association study identifies a novel locus for cannabis dependence. Mol Psychiatry. 2018;23(5):1293–302.
6. Verweij K, Vinkhuyzen A, Benyamin B, Lynskey M, Quaye L, Agrawal A, et al. The genetic etiology of cannabis use initiation: a meta-analysis of genome-wide association studies and a SNP-based heritability estimation. Addict Biol. 2013;18(5):846–50.
7. Minică CC, Verweij KJH, Most PJ, Mbarek H, Bernard M, Eijk KR, et al. Genome-wide association meta-analysis of age at first cannabis use. Addiction. 2018;113(11):2073–86 Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=132211589&site=ehost-live.
8. Stringer S, Minică CC, Verweij KJH, Mbarek H, Bernard M, Derringer J, et al. Genome-wide association study of lifetime cannabis use based on a large

meta-analytic sample of 32 330 subjects from the International Cannabis Consortium. Transl Psychiatry. 2016;6(3):e769.
9. Gillespie NA, Neale MC, Kendler KS. Pathways to cannabis abuse: a multi-stage model from cannabis availability, cannabis initiation and progression to abuse. Addiction. 2009;104(3):430–8.
10. The Government of Canada. Cannabis health effects. 2019.
11. Hall W. What has research over the past two decades revealed about the adverse health effects of recreational cannabis use? Addiction. 2015;110(1):19–35.
12. Gobbi G, Atkin T, Zytynski T, Wang S, Askari S, Boruff J, et al. Association of cannabis use in adolescence and risk of depression, anxiety, and suicidality in young adulthood: a systematic review and meta-analysis. JAMA Psychiatry. 2019;76(4):426–34.
13. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015;349:g7647.
14. National Academies of Sciences Engineering and Medicine. The health effects of cannabis and cannabinoids: the current state of evidence and recommendations for research. Washington, DC: National Academies Press; 2017.
15. Center for History New Media. Zotero: The next-generation research tool: George Mason University; 2009. Available from: http://www.zotero.org/.
16. Veritas Health Innovation. Covidence systematic review software. Melbourne. Available from: https://www.covidence.org/.
17. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. Genet Epidemiol Off Publ Int Genet Epidemiol Soc. 2008;32(3):227–34.
18. Sohani ZN, Meyre D, de Souza RJ, Joseph PG, Gandhi M, Dennis BB, et al. Assessing the quality of published genetic association studies in meta-analyses: the quality of genetic studies (Q-Genie) tool. BMC Genet. 2015;16:50.
19. Cooper ZD, Haney M. Investigation of sex-dependent effects of cannabis in daily cannabis smokers. Drug Alcohol Depend. 2014;136:85–91.
20. Lev-Ran S, Imtiaz S, Taylor BJ, Shield KD, Rehm J, Le Foll B. Gender differences in health-related quality of life among cannabis users: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Drug Alcohol Depend. 2012;123(1–3):190–200.
21. Van Gastel WA, MacCabe JH, Schubart CD, Van Otterdijk E, Kahn RS, Boks MPM. Cannabis use is a better indicator of poor mental health in women than in men: a cross-sectional study in young adults from the general population. Community Ment Health J. 2014;50(7):823–30.
22. Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, et al. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. Int J Methods Psychiatr Res. 2018;27(2):e1608.
23. Guyatt GH, Oxman AD, Schünemann H, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol. 2011;64(4):380–2.
24. Little J, Higgins JPT, Bray M, Ioannidis J, Khoury M, Manolio T, et al. The HuGENet™ HuGE review handbook, version 1.0. Ottawa, Ontario: Canada HuGENet Canada Coord Cent; 2006.

Publisher's Note

8.4 CHAPTER 3 Supplementary File 1

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	5
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	5

Page 1 of 2

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	5
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	6
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	7
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	7-8
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	NA
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	8-9
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	9-10
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	10
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	10

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

Page 2 of 2

8.5 CHAPTER 3 Published Manuscript

Hillmer *et al. BMC Med Genomics* (2021) 14:203
https://doi.org/10.1186/s12920-021-01035-5

BMC Medical Genomics

RESEARCH Open Access

Genetic basis of cannabis use: a systematic review

Alannah Hillmer[1], Caroul Chawar[1], Stephanie Sanger[2], Alessia D'Elia[1], Mehreen Butt[3], Raveena Kapoor[4], Flavio Kapczinski[1], Lehana Thabane[5] and Zainab Samaan[1*]

Abstract

Background: With the increase in cannabis use rates, cannabis use disorder is being reported as one of the most common drug use disorders globally. Cannabis use has several known physical, psychological, and social adverse events, such as altered judgement, poor educational outcomes, and respiratory symptoms. The propensity for taking cannabis and the development of a cannabis use disorder may be genetically influenced for some individuals. Heritability estimates suggest a genetic basis for cannabis use, and several genome-wide association studies (GWASs) have identified possible regions of association, albeit with inconsistent findings. This systematic review aims to summarize the findings from GWASs investigating cannabis use and cannabis use disorder.

Methods: This systematic review incorporates articles that have performed a GWAS investigating cannabis use or cannabis use disorder. MEDLINE, Web of Science, EMBASE, CINAHL, GWAS Catalog, GWAS Central, and NIH Database of Genotype and Phenotype were searched using a comprehensive search strategy. All studies were screened in duplicate, and the quality of evidence was assessed using the quality of genetic association studies (Q-Genie) tool. All studies underwent qualitative synthesis; however, quantitative analysis was not feasible.

Results: Our search identified 5984 articles. Six studies met our eligibility criteria and were included in this review. All six studies reported results that met our significance threshold of $p \leq 1.0 \times 10^{-7}$. In total 96 genetic variants were identified. While meta-analysis was not possible, this review identified the following genes, *ANKFN1*, *INTS7*, *PI4K2B*, *CSMD1*, *CST7*, *ACSS1*, and *SCN9A*, to be associated with cannabis use. These regions were previously reported in different mental health conditions, however not in relation to cannabis use.

Conclusion: This systematic review summarized GWAS findings within the field of cannabis research. While a meta-analysis was not possible, the summary of findings serves to inform future candidate gene studies and replication efforts.

Systematic Review Registration PROSPERO CRD42020176016.

Keywords: Systematic review, Cannabis, Genetics, Genome-wide Association Study

Introduction

Rationale

Over the past two decades cannabis use and dependence are estimated to have increased, with cannabis use disorder (CUD) reported as one of the most common drug use disorders globally [1]. In Canada, it has been reported that nearly 17 percent of Canadians aged 15 years and older reported using cannabis between

*Correspondence: samaanz@mcmaster.ca
[1] Neuroscience Graduate Program, Department of Psychiatry and Behavioural Neurosciences, McMaster University, 100 West 5th St., Hamilton, ON L8N 3K7, Canada
Full list of author information is available at the end of the article

October and December of 2019, an increase from 14 percent between January to March of 2018. Additionally, cannabis consumption rates are higher among males than females [2]. Concerningly, cannabis has been associated with substantial adverse effects. Like other drugs, cannabis can result in cravings, dependence, and drug-seeking behaviour [3, 4]. During intoxication, cannabis can interfere with memory, motor coordination, altered judgement, and at higher doses, paranoia or psychosis [3]. Further, repeated use of cannabis can have long lasting effects, including altered brain development, poor education outcome, cognitive impairment, diminished life satisfaction and achievement, poor professional and social achievements, symptoms of chronic bronchitis and increased risk of chronic psychotic disorders [3, 5].

Heritability estimates for cannabis use initiation varied from 30 to 48%, and from 51 to 59% for problematic cannabis use, suggesting a genetic component exists [6]. Genome-wide association study (GWAS) meta-analyses have identified possible regions of association on chromosome 3 for lifetime cannabis use (*CADM2*), chromosome 10 for CUD (rs77300175), and chromosome 16 for age of first cannabis use (*ATP2C2*) [7–9]. Moreover, candidate gene studies have detected some significant associations with cannabis use on the *CNR1*, *GABRA2*, *FAAH*, and *ABCB1* genes, but as with genome-wide association studies (GWASs), replication of these associations has been inconsistent [10].

GWASs provide a 'hypothesis-free' method of identifying novel variant-trait associations, leading to the discovery of novel biological mechanisms and diverse clinical applications [11]. As such, in this systematic review, we will summarize GWAS findings relevant to cannabis use or CUD outcomes and discuss future directions.

Objectives:

The main goal of this systematic review is to identify genetic variants from GWASs associated with cannabis use.

Primary objectives of this systematic review include the following:

1. Identify genetic variants associated with current cannabis use. Current cannabis use is defined by either self-report or positive urine drug screens within 1 month of the study being conducted.
2. Identify genetic variants associated with lifetime cannabis use. Lifetime cannabis use is defined by any self-reported or positive urine drug screens of cannabis use within one's lifetime.
3. Identify genetic variants associated with CUD. CUD is defined by any diagnostic and classification systems used to diagnose CUD or questionnaires validated to assess CUD.

Secondary objectives of this systematic review include the following:

1. Identify genetic variants associated with the adverse outcomes of cannabis use, including psychiatric (cognitive impairment, psychotic symptoms, depression, anxiety, suicidal behavior) and non-psychiatric (chronic bronchitis, lung infections, chronic cough, increased risk of motor vehicle accidents) [12–14].
2. When feasible, perform subgroup summaries by sex or ethnic differences.

Methods

This systematic review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [15] (see PRISMA checklist in Additional file 1). The Human Genome Epidemiology Network (HuGENet) guideline was used to supplement the PRISMA guideline. While this review does not conform with the HuGENet guideline expectations of reporting on candidate gene study findings, the HuGENet is used to uphold the standard of reporting research specific to genetic association studies [16].

Protocol and registration

The protocol for this systematic review has been registered within the International Prospective Register of Systematic Reviews (PROSPERO) (registration number: CRD42020176016) [17]. The full protocol has been published in the journal of Systematic Reviews [18].

Eligibility criteria

This review investigates GWASs presenting original data on associations between cannabis use and genetic polymorphisms using any study design (i.e. case–control, cohort, etc.). We include studies investigating CUD as well as any studies measuring any use of cannabis. Studies that investigated CUD as defined by any version of the Diagnostic and Statistical Manual (DSM) or other diagnostic and classification systems such as the International Statical Classification of Diseases and Related Health Problems-10 (ICD-10) were included. We define cannabis use based on the included studies' definitions and accept the following definitions: current cannabis use is defined as either self-report or positive urine drug screens within one month of the study being conducted, and lifetime cannabis use is defined as any self-reported or positive urine drug screens of cannabis use within one's lifetime [19]. All other studies that did not perform a GWAS and investigate cannabis use or CUD were excluded. No restrictions were placed on the study setting or participant's age, sex, ethnic background or

language. Further details on the inclusion criteria can be found in the study protocol [18].

Information sources and search strategy

A Health Science Librarian was consulted to develop a comprehensive search strategy. OVID MEDLINE 1946-Present, Web of Science 1976-Present, OVID EMBASE 1974-Present, EBSCOHost CINHAL 1981-Present, GWAS Catalog, GWAS Central, and NIH Database of Genotype and Phenotype databases were searched using the established strategy, modified for each database. All databases were searched from inception to February 2nd, 2021. The search strategy included all terms relevant to genome-wide association studies and cannabis. The search strategies for each electronic database are provided in Table 1.

Study selection and data collection process

Calibration was completed prior to the formal screening process. Title and abstract screening, full-text screening and data extraction phases were completed in duplicate through Covidence [20]. Conflict resolution at the title and abstract and full-text stages was performed by a senior reviewer (AH or CC), blind to the reviewer's vote. Disagreements at the data extraction stage was resolved by the consensuses of the two reviewers. The reason for study exclusion was recorded at the full-text stage.

Data items

Data extracted included baseline participant characteristics, the measure of cannabis used, relevant and significant measured outcomes, statistical measures, and reported study limitations and conflicts. For this review, the threshold of significance of genetic variants reaching $p \leq 10^{-7}$ was set, as some GWAS results with this significance level have been shown to be replicable within the literature [21].

Risk of bias within studies and data analysis

Quality assessment was completed in duplicate for each included study using the Quality of Genetic Association Studies (Q-Genie) tool [Version 1.1] [22]. Disagreements of quality assessment was resolved through discussion between the two reviewers, and the first author reviewed and confirmed all quality assessments.

Summary measures and synthesis of results

A random-effects meta-analysis through pooled odds ratios was planned to quantitatively assess the data. However, these measures were not appropriate as data extracted from each study were unique and could not be combined. For the aforementioned reasons, a heterogeneity test, and a subgroup meta-analyses could not be completed.

Risk of bias across studies

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) was used to assess the strength of evidence, with specific consideration of prognostic factors [23, 24]. GRADE scores assess outcomes according to the risk of bias, publication bias, consistency, directness, and precision [23].

Results

Study selection

The search strategy, along with hand-searching, yielded 5984 studies. After removing duplicates through the Zotero reference manger and Covidence, 4344 studies were unique and screened for eligibility at the title and abstract phase [20, 25]. Of the 69 studies eligible for full-text screening, 6 studies were included in this review and underwent data extraction and quality assessment.

Studies frequently failed to meet the eligibility criteria for inclusion for the following reasons (i) conducted a GWAS meta-analysis, (ii) conducted a candidate gene study or (iii) were investigating a factor associated with cannabis use (i.e. aggression) rather than cannabis use itself.

Please see the PRISMA flow diagram in Fig. 1.

Study characteristics

Individual study characteristics are reported in Table 2. Two studies were case–control, two were cohort, one was case-cohort, and another was case-cohort and cohort. Interestingly, the first GWAS in the field of cannabis use was published in 2011 and the most recent conducted in 2019 [26, 27]. All studies used data from large study datasets. Three studies utilized the Study of Addiction: Genetics and Environment (SAGE) [4, 26, 28]. The International Cannabis Consortium (ICC), UKBiobank, and 23andMe were utilized in one study which performed three independent GWAS on the aforementioned datasets [9]. Another study combined the Yale-Penn and the International Consortium on the Genetics of Heroin Dependence (ICGHD) to perform a single GWAS [4]. Finally, one study utilized the Integrative Psychiatric Research (iPSYCH) [27] and another the Netherlands twin registry [29]. Studies varied in size from 3053 to 51,372 participants. Of the studies which reported participants' sex and age, three studies had a population comprised of mostly female participants [9, 26, 28, 29], while only one reported majority male [4]. The mean age of study participants varied from mid-thirties to mid-fifties. Three studies reported on participants of European or African American

Table 1 Search strategy

OVID MEDLINE	1. Genome-Wide Association Study/ 2. Genotyping Techniques/ 3. Genome, Human/ 4. Genetic Variation/ 5. genetics/ or exp human genetics/ 6. (human* adj2 (genotyp* or genome* or genetic*)).ti,ab,kw,kf 7. (GWS or GWAS or GWA).mp 8. genome wide.ti,ab,kw,kf 9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 10. exp Cannabis/ 11. ((cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) adj2 (overdose* or use* or using or misuse* or abus* or dependence* or addict*)).ti,ab,kw,kf 12. 10 or 11 13. 9 and 12 *14. Limit 13 to humans*
Web of Science	TS = (genome-wide association study or genome-wide association or GWAS or GWA or genome wide) TS = (human NEAR/2 genome) TS = ((cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) NEAR/2 (overdose* or use* or using or misuse* or abus* or dependence* or addict*)) TS = (cannabis* or marijuana* or marihuana*) #1 OR #2 #3 OR #4 #5 and #6
OVID EMBASE	Genome-Wide Association Study/ Genotyping Techniques/ Genome, Human/ Genetic Variation/ genetics/ or exp human genetics/ (human* adj2 (genotyp* or genome* or genetic*)).ti,ab,kw. (GWS or GWAS or GWA).mp. genome wide.ti,ab,kw. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 exp Cannabis/ ((cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) adj2 (overdose* or use* or using or misuse* or abus* or dependence* or addict*)).ti,ab,kw. 10 or 11 9 and 12 *Limit 13 to human*
EBSCOHost CINAHL	genome-wide association study or genome-wide association or GWAS or GWA or genome wide or genome cannabis* or marijuana* or cannabinoids* or marihuana* or hash* or kush* or weed* or pot* or THC* or CBD*) overdose* or use* or using or misuse* or abus* or dependence* or addict* S2 and S3 S1 and S4 *Limit to Human*
GWAS Catalog	Terms Searched: Cannabis Cannabis dependence Marihuana Marijuana Cannabinoids Hash Kush Weed Pot THC CBD

include the study details, population, findings, and source of data.

Discussion

A lack of consistent evidence exists in the current literature for genetic variants associated with cannabis use. In addition, this is the first known systematic review to synthesize the available evidence on genetic variants associated with cannabis use. The proposed systematic review aims to identify all genetic variants that have reached borderline genome-wide significance associated with cannabis use and CUD. The proposed systematic review will provide an overview of the current literature on the genetics of cannabis, aiding in the genetic understanding of cannabis use. Understanding the genetic contribution to cannabis use and its effects such as cannabis use disorder has the potential to aid medical practitioners in making decisions related to cannabis use for medical reasons and the associated potential risks. Additionally, variants reaching borderline genome-wide significance will be examined in the context of their known or biologically plausible relevance to further our understanding.

Anticipated limitations of this review existed at both the study and review level. Limitations at the study level may include a lack of reporting quality control steps, reporting of variants within linkage disequilibrium, small sample size, and a lack of reporting variants that failed to reach genome-wide significance ($p < 5 \times 10^{-8}$) but may have reached borderline significance levels ($p < 10^{-7}$). At the review level, limitations exist in the expected high heterogeneity, differing outcomes for cannabis use reported in the literature and the exclusion of meta-analysis and candidate gene studies.

On completion of the systematic review, we will publish in a peer-review academic journal to reach both clinical and academic experts in the field. This systematic review will then inform and direct the further investigation of genetic variants associated with cannabis through candidate gene studies.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13643-020-01442-2.

Additional file 1. "PRISMA-P 2015 Checklist" and contains PRISMA-P checklist.

Additional file 2. Search strategy.

Abbreviations

CUD: Cannabis use disorder; DSM-5: Diagnostic and Statistical Manual 5th edition; GRADE: Grading of Recommendations Assessment, Development and Evaluation; GWAS: Genome-wide association study; GWASs: Genome-wide association studies; HuGENet: The Human Genome Epidemiology Network; ICD-10: International Statistical Classification of Diseases and Related Health Problems -10; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols; Q-Genie: The quality of genetic association studies

Amendments

If amendments to this protocol are made, they will be documented and communicated to the journal. A data of amendment, description, and rationale will accompany each amendment.

Authors' contributions

ZS is the guarantor. AH and CC drafted the manuscript. AH, CC, and ZS contributed to the development of the selection criteria, the risk of bias assessment strategy, and data extraction criteria. SS provided expertise in developing the search strategy. All authors read, provided feedback, and approved the final manuscript.

Funding

This study is supported in part by CIHR (grant number PJT-156306). The funding source has no role in the study design, analysis, reporting, or the decision to publish the study.

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Neuroscience Graduate Program, Department of Psychiatry and Behavioural Neurosciences, McMaster University, 100 West 5th St., Hamilton, ON L8N 3 K7, Canada. [2]Health Science Library, McMaster University, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada. [3]Integrated Science Program, McMaster University, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada. [4]Health Sciences Program, McMaster University, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada. [5]Department of Psychiatry and Behavioural Neurosciences, McMaster University, 100 West 5th St., Hamilton, ON L8N 3 K7, Canada. [6]Population Health Research Institute, McMaster University, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada. [7]Department of Health Research Method, Evidence & Impact, 1280 Main St. W, Hamilton, ON L8S 4 L8, Canada.

Received: 20 April 2020 Accepted: 31 July 2020
Published online: 20 August 2020

References

1. The Government of Canada. Cannabis legalization and regulation. 2018.
2. Statistics Canada. National Cannabis Survey, third quarter 2019. 2019.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington: American Psychaitric Publishing; 2013.
4. Verweij KJH, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG, et al. Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. Addiction. 2010;105(3):417–30.
5. Agrawal A, Chou YL, Carey CE, Baranger DAA, Zhang B, Sherva R, et al. Genome-wide association study identifies a novel locus for cannabis dependence. Mol Psychiatry. 2018;23(5):1293–302.
6. Verweij K, Vinkhuyzen A, Benyamin B, Lynskey M, Quaye L, Agrawal A, et al. The genetic etiology of cannabis use initiation: a meta-analysis of genome-wide association studies and a SNP-based heritability estimation. Addict Biol. 2013;18(5):846–50.
7. Minică CC, Verweij KJH, Most PJ, Mbarek H, Bernard M, Eijk KR, et al. Genome-wide association meta-analysis of age at first cannabis use. Addiction. 2018;113(11):2073–86 Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=132211589&site=ehost-live.
8. Stringer S, Minică CC, Verweij KJH, Mbarek H, Bernard M, Derringer J, et al. Genome-wide association study of lifetime cannabis use based on a large

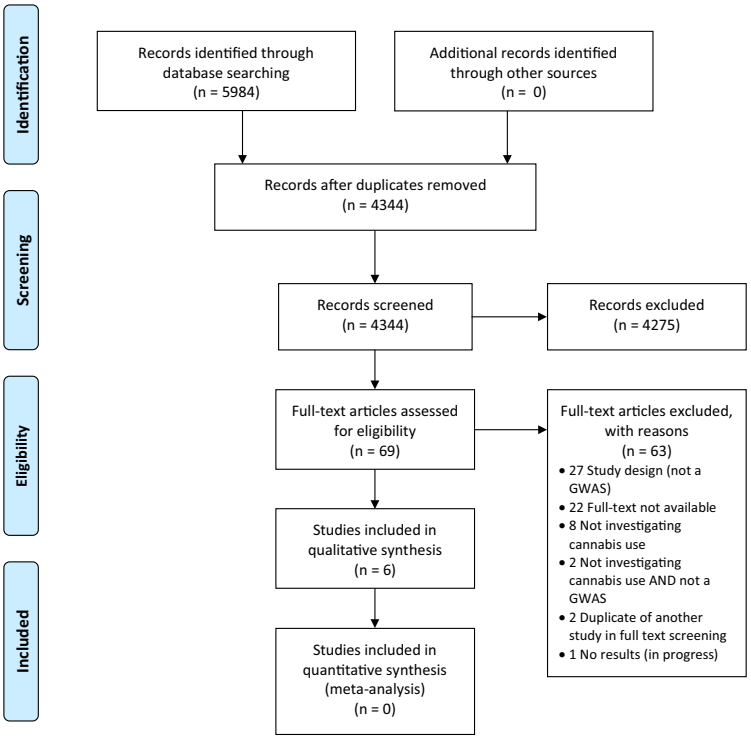


Fig. 1 PRISMA Flow Dagram of Study Inclusion

onset of cannabis use. Identified SNPs were found on chromosomes 5, 9, 18 and 19, with one SNP associated with cannabis initiation was found on the Zinc finger protein, *ZNF181*. All participants were of European descent and were selected from the Netherlands Twin Registry. Cannabis initiation was defined as ever/never having used cannabis while age of onset was determined by asking participants an open-ended question [29].

Pasman et al. conducted three independent GWASs in three separate cohorts, all of which included European participants: ICC, UKBiobank, and 23andMe. While results from 23andMe were unable to be shared due to privacy policies, the lead author kindly provided SNPs reaching borderline significance threshold with lifetime cannabis use for GWAS conducted in the ICC and UKBiobank cohort. One SNP in the ICC cohort and 18 SNPs in the UKBiobank were associated with lifetime cannabis use, with no genes specified in either. Lifetime cannabis use was defined as any cannabis use during lifetime [9].

Sherva et al. identified 42 SNPs associated with DSM-IV cannabis dependence criteria count across 27 different genes/regions including *INTS7, SNORA26, RPS20P10, PI4K2B, CSMD1, PSMB7, HABP2, MEFV, CST7, APMAP, ACSS1, snoU13, TPST2, SCN9A, CTA-445C9.15, CTA-445C9.14-CTA-4, SCN9A-SCN7A, ARL2BPP5-RP11-541P9.3, RP11-755E23.3-CCDC67, SNORD11-RNU6-1014P, RP5-860P4.2-CST7, RNU6-1257P, APMAP-ACSS1, C9.15, RPS20P10-CYP26B1, PI4K2B-ZCCHC4*, and *CST7-APMAP*. European and African American participants were selected from the Yale-Penn Study, the SAGE study and the ICGHD cohorts [4].

Table 2 Individual study characteristics

First Author Last Name, Year	Title of Publication	Study Design	Cohort used	Sample Size	% Male	Mean age	Ethnicity	Outcome of interest
Agrawal, 2011	A Genome-wide Association Study of DSM-IV Cannabis Dependence	Case–Control	SAGE	3054	NR	39.00	2019 European-Americans, 1035 African Americans	Life-time history of DSM-IV cannabis dependence, modified to included cannabis withdrawal
Agrawal, 2014	DSM-5 Cannabis Use Disorder: A Phenotypic and Genomic Perspective	Case–control	SAGE	3053	49%	38.10	2018 European-Americans, 1035 African Americans	DSM-5 cannabis use disorder factor scores
Demontis, 2019	Genome-wide association study implicates CHRNA2 in cannabis use disorder	Case–cohort	iPSYCH	CUD: 2387 Control: 48,985 Total: 51,372	NR	CUD: 24.77 Control: 22.67	European	International Statistical Classification of Disease and Related Health Problems, 10th revision (ICD-10) diagnosis reflecting problematic and persistent use of cannabis
Minica, 2015	Heritability, SNP- and Gene-Based Analyses of Cannabis Use Initiation and Age at Onset	Cohort	Netherlands twin registry	6744	39.1%	39.09	European	Self-reported use of cannabis ever in lifetime and self-reported age of onset of cannabis use
Pasman, 2018	GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal effect of schizophrenia liability	Cohort	ICC study UK Biobank 23andMe	35,297 126,785 22,683	44.3% 43.7% NR	35.7 55 NR	European	Any cannabis use within lifetime
Sherva, 2016	Genome-wide Association Study of Cannabis Dependence Severity, Novel Risk Variants, and Shared Genetic Risks	Cohort/Case-cohort	Yale-Penn, SAGE, ICGHD	14,754	53.4%	39.24	8754 European-American, 6000 African American	Criterion count for DSM-IV Cannabis Dependence

Table 3 Q-genie scores

First Author Last Name, Year	Reported conflicts of interest	Reported study limitations	Q-Genie Score	Quality Assessment
Agrawal, 2011	Drs. LJ Bierut, J. Rice, A. Goate and S Saccone are listed as inventors on the patent "Markers for Addiction" (US 20,070,258,898); covering the use of certain SNPs in determining the diagnosis, prognosis, and treatment of addiction. Dr. Bierut has acted as a consultant for Pfizer, Inc. in 2008. All other authors report no competing interests	SAGE study was ascertained for alcohol dependence led to a high level of comorbidity in the cannabis dependent cases and exposed controls Use of controls with other forms of substance dependence but not cannabis dependence protected against signals that may have been less specific Power computations revealed that minor allele frequencies ranging from 15–40% association signals with odds ration exceeding 1.45 were able to be detected	53	Good Quality
Agrawal, 2014	Laura J. Bierut is listed as an inventor on Issued U.S. Patent 8,080,371,"Markers for Addiction" covering the use of certain SNPs in determining the diagnosis, prognosis, and treatment of addiction	The sample was ascertained from three family studies of substance use disorders for the express purpose of identifying genetic variants for alcoholism, nicotine and cocaine dependence and related pathology Differences in DSM-IV and DSM-5 criteria leading to different assessments of withdrawal and diagnosis of CUD across study populations	55	Good quality
Demontis, 2019	T. Werge has been a lecturer and advisor to H. Lundbeck A/S. T.E. Thorgeirsson, D.F. Gudbjartsson, G.W. Reginsson, H. Stefansson and K. Stefansson are employees of deCODE genetics/Amgen	None reported	57	Good quality
Minica, 2015	None	No statistically significant GWAS findings that pass the threshold $p < 1.0 \times 10^{-8}$	49	Moderate quality
Pasman, 2018	P.F., S.L.E. and members of the 23andMe Research Team are employees of 23andMe Inc. J.A.R.-Q. was on the speakers' bureau and/or acted as consultant for Eli Lilly, Janssen- Cilag, Novartis, Shire, Lundbeck, Almirall, BRAINGAZE, Sincrolab and Rubió in the last 5 years. He also received travel awards (air tickets and hotel) for taking part in psychiatric meetings from Janssen-Cilag, Rubió, Shire and Eli Lilly. The Department of Psychiatry chaired by him received unrestricted educational and research support from the following pharmaceutical companies in the last 5 years: Eli Lilly, Lundbeck, Janssen- Cilag, Actelion, Shire, Ferrer and Rubió	Lifetime cannabis was analyzed as a single dichotomous measure combining experimental and regular users in a single group The power of some analyses may have been limited	52	Good quality

Table 3 (continued)

First Author Last Name, Year	Reported conflicts of interest	Reported study limitations	Q-Genie Score	Quality Assessment
Sherva, 2016	Dr Kranzler reports being a consultant or an advisory board member for Alkermes, Indivior, Lundbeck, and Otsuka (unrelated to the present study) and being a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative, which is supported by AbbVie, Ethypharm, Lilly, Lundbeck, and Pfizer. No other disclosures were reported	One of the significant SNPs identified (rs143244591 on chromosome 3) has little supportive evidence for association from other SNPs in the region None of the GWAS SNPs identified in the full GWAS analysis are rare Lack of evidence of associations in both the European American and African American participants The Yale-Penn samples who underwent genotyping on the HumanOmni1-Quad and Human Core Exome chips showed more consistent results than the corresponding SAGE population The cohorts used have higher rates of polysubstance dependence than the general population and may not be generalizable to individuals who only use cannabis	52	Good quality

Table 4 SNPs reaching borderline significance threshold

First Author Last Name, Year	Outcome associated with SNP	Build	SNP IDs	Chr:Pos	Alleles	Minor Allele	Gene or Locus	MAF	N	Measure of Association type	Measure of Association value	Measure of Variability	Measure of Variability value	*p* value	Ethnicity
Agrawal, 2011	DSM-IV cannabis dependence	hg18	rs1019238	17			ANKFN1	EA = 0.4, AA = 0.1	3054	OR	1.453	95% CI	1.254–1.682	6.1E−07	AA & EA
			rs1431318	17			ANKFN1	EA = 0.44, AA = 0.25	3054	OR	0.708	95% CI	0.616–0.812	9.14E−07	AA & EA
Agrawal, 2014	Cannabis use disorder factor scores from 12 DSM IV + 5 criteria	hg18	rs4364205	3		T		0.41	1035	β	−0.19	95% CI	−0.26–(−0.12)	1.3E−07	AA
Demontis, 2019	Cannabis use disorder	hg19	rs56372821	8	A/G	A		0.163	51,372	OR	0.728			**9.31E−12**	EA
			rs4732724	8	C/G	C		0.324	51,372	OR	0.82			**1.34E−08**	EA
			rs6558008	8	A/C	A		0.755	51,372	OR	1.237			**2.45E−08**	EA
			rs73229090	8	A/C	A		0.11	51,372	OR	0.702			**7.41E−10**	EA
			rs937220	8	A/G	A		0.259	51,372	OR	0.8258			2.46E−07	EA
Minica, 2015	Cannabis initiation	GRCh37	rs35487050	19:35,221,228	C/T		ZNF181		6744	β	0.81	SE	0.16	1.68E−07	E
			rs35917943	19:35,147,183	C/A			< 0.05	6744	β	0.77	SE	0.15	1.62E−07	E
			rs35760174	19:35,221,582	C/G				6744	β	0.76	SE	0.15	7.04E−07	E
	Age of onset of cannabis use		rs142324060	5:95,425,757	G/A			< 0.05	5148	β	0.68	SE	0.11	7.66E−08	E
			rs78505392	5:95,422,966	C/G				5148	β	0.58	SE	0.1	2.16E−07	E
			rs12003072	9:86,771,161	A/C				5148	β	0.52	SE	0.09	3.04E−07	E
			rs77097806	5:95,456,735	A/G				5148	β	0.56	SE	0.1	3.54E−07	E
			rs6879646	5:95,450,187	A/G				5148	β	0.57	SE	0.1	3.61E−07	E
			rs4613744	5:95,451,494	C/T				5148	β	0.55	SE	0.1	5.07E−07	E
			rs60218730	5:95,492,765	G/T				5148	β	0.59	SE	0.11	5.98E−07	E
			rs74305417	9:86,779,774	C/G				5148	β	0.52	SE	0.09	6.2E−07	E
			rs142981069	18:58,826,022	G/A				5148	β	0.47	SE	0.09	7.25E−07	E
			rs12386084	18:58,827,145	C/G				5148	β	0.47	SE	0.09	7.25E−07	E

Table 4 (continued)

First Author Last Name, Year	Outcome associated with SNP	Build	SNP IDs	Chr:Pos	Alleles	Minor Allele	Gene or Locus	MAF	N	Measure of Association type	Measure of Association value	Measure of Variability	Measure of Variability value	*p* value	Ethnicity
			rs117918936	18:58,828,323	G/A				5148	β	0.47	SE	0.09	7.25E−07	E
			rs2160801	18:58,829,024	T/A				5148	β	0.47	SE	0.09	7.25E−07	E
			rs145424173	18:58,829,597	T/C				5148	β	0.47	SE	0.09	7.25E−07	E
			rs117538409	18:58,830,942	G/C				5148	β	0.47	SE	0.09	7.25E−07	E
			rs17817245	18:58,832,135	A/G				5148	β	0.47	SE	0.09	7.25E−07	E
			rs140206809	18:58,833,215	A/G				5148	β	0.47	SE	0.09	7.25E−07	E
			rs117692712	18:58,834,506	T/G				5148	β	0.47	SE	0.09	7.25E−07	E
			rs17817423	18:58,835,462	C/T				5148	β	0.47	SE	0.09	7.25E−07	E
			rs9916935	18:58,835,931	T/C				5148	β	0.47	SE	0.09	7.25E−07	E
			rs192013604	18:58,838,324	T/C				5148	β	0.47	SE	0.09	7.25E−07	E
			rs117471640	18:58,838,402	A/G				5148	β	0.47	SE	0.09	7.25E−07	E
			rs78456402	9:86,781,900	C/A				5148	β	0.5	SE	0.09	9.09E−07	E
			rs11998981	9:86,783,107	T/C				5148	β	0.5	SE	0.09	9.09E−07	E
			rs79236058	5:95,478,830	G/A				5148	β	0.57	SE	0.1	9.59E−07	E
Pasman, 2018	Lifetime cannabis use	hg19	rs7609594	3:85,482,595	G/A	A	CADM2	0.38	126,785	β	0.068	SE	0.01	**5.86E−11**	E
			rs4099556	4:37,067,936	G/A	A	MIR4801	0.176	126,785	β	−0.079	SE	0.013	**3.56E−09**	E
			rs78680891	3:51,028,954	C/G	C		0.002299	126,574	β	−0.52473	SE	0.106473	8.30E−07	E
			rs2071704	4:3,240,159	C/T	C		0.419929	125,953	β	−0.04974	SE	0.010162	9.83E−07	E
			rs76565656	7:148,048,453	G/A	G		0.122539	124,918	β	−0.07974	SE	0.015258	1.73E−07	E
			rs4308708	8:76,702,058	A/C	A		0.041699	126,193	β	−0.13628	SE	0.025107	5.70E−08	E
			rs10883796	10:104,655,315	G/A	G		0.296441	126,444	β	−0.05638	SE	0.010951	2.62E−07	E
			rs11186071	10:91,969,337	G/A	G		0.004168	126,711	β	−0.90907	SE	0.18524	9.22E−07	E
			rs11214441	11:112,846,713	T/A	T		0.394257	126,761	β	−0.05354	SE	0.010263	1.82E−07	E
			rs5442	12:6,954,864	G/A	G		0.066921	126,785	β	−0.09622	SE	0.019181	5.27E−07	E
			rs75448266	12:6,798,632	A/C	A		0.084049	125,672	β	−0.0944	SE	0.017746	1.04E−07	E
			rs17083392	13:69,026,559	A/G	A		0.004286	126,745	β	−1.60924	SE	0.3177	4.08E−07	E
			rs2019135	13:69,034,323	T/C	T		0.004267	126,772	β	−1.76061	SE	0.323766	5.39E−08	E
			rs2019150	13:69,034,188	C/A	C		0.004262	126,771	β	−1.76061	SE	0.323766	5.39E−08	E
			rs677755	13:69,041,438	T/C	T		0.00486	126,760	β	1.426	SE	0.290601	9.24E−07	E
			rs74632168	13:69,048,954	C/T	C		0.004056	126,772	β	−1.72062	SE	0.3190914	6.91E−08	E
			rs2650494	16:28,318,440	A/G	A		0.414312	124,286	β	−0.05259	SE	0.010223	2.68E−07	E
			rs325363	18:38,331,295	G/A	G		0.229965	126,624	β	0.059212	SE	0.011689	4.07E−07	E

Table 4 (continued)

First Author Last Name, Year	Outcome associated with SNP	Build	SNP IDs	Chr:Pos	Alleles	Minor Allele	Gene or Locus	MAF	N	Measure of Association type	Measure of Association value	Measure of Variability	Measure of Variability value	*p* value	Ethnicity
			rs4492854	11:112,983,534	C/T	C		0.4363	30,366	β	0.1113	SE	0.021251	1.63E−07	E
Sherva, 2016	DSM-IV cannabis dependence criteria count	GRCh37	-	1:212,183,114:I			INTS7		1250					2.05E−07	AA
			rs141482228				INTS7		1250					**8.84E−10**	AA
			rs77349458				SNORA26,INTS7		1250					1.57E−07	AA
			rs77448142				RPS20P10,RPS20P10-CYP26B1		1250					2.13E−07	AA
			-	2:167,239,818:D			SCN9A-SCN7A		4750					7.43E−07	AA
			rs313542				PI4K2B		2640					**7.12E−10**	EA
			rs76897780				PI4K2B,PI4K2B-ZCCHC4		2640					4.63E−07	EA
			rs73323306				ARL2BPP5-RP11-541P9.3		4750					**7.61E−09**	AA
			rs7832545				CSMD1		2640					**3.61E−10**	EA
														3.02E−07*	
			rs77378271				CSMD1		2640					**5.30E−10**	EA
														6.69E−07*	
			rs75721860				CSMD1		2640					**2.75E−09**	EA
			rs7853028				PSMB7		1250					**8.94E−09**	AA
			rs4643011				HABP2		1250					7.21E−07	AA
			rs116474042				RP11-755E23.3-CCDC67		4750					3.81E−07	AA
			rs189167038				SNORD11-RNU6-1014P		1250					**2.64E−09**	AA
			rs142162415				SNORD11-RNU6-1014P		1250					**1.88E−08**	AA
			rs117047810				SNORD11-RNU6-1014P		1250					5.68E−08	AA
			rs11466037				MEFV		1250					8.08E−07	AA
			rs183568578				MEFV		1250					5.62E−07	AA
			rs114269992				RP5-860P4.2-CST7		1250					2.13E−07	AA

Table 4 (continued)

First Author Last Name, Year	Outcome associated with SNP	Build	SNP IDs	Chr:Pos	Alleles	Minor Allele	Gene or Locus	MAF	N	Measure of Association type	Measure of Association value	Measure of Variability	Measure of Variability value	*p* value	Ethnicity
														1.63E−07*	
			rs114620529				RP5-860P4.2-CST7		1250					2.15E−07	AA
														1.61E−07	
			rs116629084				RP5-860P4.2-CST7		1250					1.70E−07	AA
														1.27E−07*	
			rs115384512				RP5-860P4.2-CST7		1250					8.91E−08	AA
														7.19E−08*	
			rs147641662				CST7;RP5-860P4.2-CST7		1250					8.87E−08	AA
														7.13E−08*	
			rs191783144				CST7;RP5-860P4.2-CST7		1250					8.81E−08	AA
														7.07E−08*	
			rs146806338				CST7;RP5-860P4.2-CST7		1250					8.80E−08	AA
														7.03E−08*	
			rs114828727				CST7;APMAP;CST7-APMAP		1250					9.01E−08	AA
														6.92E−08*	
			rs114128002				APMAP		1250					8.06E−08	AA
														6.33E−08*	
			rs115342711				RNU6-1257P;APMAP		1250					2.54E−07	AA
														1.15E−07*	
			-	20:24,955,422:I			RNU6-1257P;APMAP		1250					2.54E−07	AA
														1.15E−07*	
			rs115764440				APMAP		1250					2.33E−07	AA
														9.03E−08*	
			rs116068335				APMAP-ACSS1		1250					1.66E−07	AA
														5.62E−08*	
			rs114637142				ACSS1;APMAP-ACSS1		1250					9.15E−08	AA
			rs114071901				ACSS1		1250					1.06E−07	AA
			rs113232742				ACSS1		1250					8.82E−08	AA

Table 4 (continued)

First Author Last Name, Year	Outcome associated with SNP	Build	SNP IDs	Chr:Pos	Alleles	Minor Allele	Gene or Locus	MAF	N	Measure of Association type	Measure of Association value	Measure of Variability	Measure of Variability value	*p* value	Ethnicity
			rs114199928				ACSS1		1250					8.42E−08	AA
			rs114836364				ACSS1		1250					8.23E−08	AA
			rs116669368				ACSS1		1250					5.08E−08	AA
			rs145379934				ACSS1		1250					**3.96E−08**	AA
			rs143020225	2:167,214,714	G/A	G	snoU13:SCN9A	0.95	4750					5.85E−07	AA
			-	chr2:2:26,917,069:I			C9.15.TPST2.CTA-445C9.14-CTA-4		2640					**3.15E−10**	EA
			rs4149485				TPST2.CTA-445C9.15		2640					**2.24E−10**	EA

[*] Adjusted for the DSM-IV criteria counts for alcohol, cocaine, and opioid dependence

AA = African American, EA = European American, E = European

Bold type indicates variants for which the p value reached genome-wide significance ($p \leq 5 \times 10^{-8}$)

While no SNPs were reported within the same region not allowing further quantitative analysis, several phenotypic similarities exist across studies. Interestingly, two studies found that educational attainment was negatively associated with CUD [26, 27] and a third found positive genetic correlations with educational attainment [9]. Two studies found that cannabis dependence was significantly related to alcohol, nicotine, and cocaine dependence [4, 26] with a third reporting a positive genetic correlation between lifetime cannabis use and smoking and alcohol use and dependence [9].

Risk of bias across studies

Outcomes assessed for GRADE include lifetime cannabis use, diagnosis of CUD, criterion count for CUD and age of onset of cannabis use. All outcomes included two studies except for age of onset of cannabis use which only included one study. The full GRADE table can be found in Table 5. All outcomes were rates as important, and no outcome was rated as having a "very serious" concern pertaining to any certainty criteria. Only the outcomes of diagnosis of CUD and criterion count for CUD had a serious rating, both of which were in the category of indirectness. Both of these outcomes were downgraded due to the use of different diagnostic criteria. More specifically, for the outcome of diagnosis of CUD Agrawal et al. [26] utilized the DSM-IV and Demontis et al. (2019) utilized the ICD-10 and for the outcome of criterion count of CUD Sherva et al. [4] utilized the DSM-IV criteria and Agrawal et al. [28] utilized a combination of DSM-IV and DSM-5 criteria.

Discussion

Summary of evidence

In this review we identified 96 genetic variants to be associated with different measures of cannabis. Of these genetic variants, 18 reached the genome-wide significance threshold of $p \leq 5 \times 10^{-8}$, all of which are available in Table 4. As no genetic variants included in this review were reported in more than one study, meta-analyses were not possible. However, of the genetic variants identified in this review, several are located on genes in which previous studies have reported associations with mental health, namely *ANKFN1*, *INTS7*, *PI4K2B*, *CSMD1*, *CST7*, *ACSS1*, and *SCN9A*.

With cannabis being a legal substance, research on the benefits and harms of cannabis has been on the rise. However, a limited number of GWASs have been conducted on cannabis use to determine any genetic associations. This systematic review was able to qualitatively summarize findings from GWASs reporting borderline genome-wide significance to aid in identifying SNPs that may be replicable in future studies. We have identified six eligible studies that reported independent GWAS results, one of which primarily focused on a GWAS meta-analysis. Of the included studies, only participants from European or African American ethnicities were included, suggesting a need for genetic studies being conducted in more diverse ethnic populations. All six studies reported at least one borderline significant SNP; however, no two studies identified the same SNP. SNPs were found to be associated with CUD, cannabis initiation, age of onset of cannabis use, DSM-IV cannabis dependence criteria count, or lifetime cannabis use on various gene regions. According to assessment using the Q-genie tool and GRADE tool, no study or outcome was deemed to be of poor quality. Additionally, with GWAS requiring a sample size of thousands of participants for adequate power, all studies met this threshold [30].

While the majority of genes identified in the included studies had either no known function or biological plausibility, and none had any additional associations with cannabis use, as mentioned above, several did have associations with mental health conditions and are discussed briefly, namely *ANKFN1*, *INTS7*, *PI4K2B*, *CSMD1*, *CST7*, *ACSS1*, and *SCN9A*. *ANKFN1* is a protein coding gene which has been associated with smoking cessation and nicotine dependence [31]. *INTS7* is a component of the integrator complex, which is involved in the small nuclear RNA U1 and U2 transcriptions [32] and has been associated with bipolar temperament [31, 33]. *PI4K2B* contributes to the overall PI4-kinase activity of the cell [32] and is associated with attention deficit hyperactivity disorder (ADHD), logical memory and abnormality of neuronal migration [33]. *CSMD1* has been associated with behavioural disinhibition, schizophrenia, cognitive tests, chronic bronchitis, and bipolar disorder [31, 33]. *CST7* is associated with alcohol consumption and myocardial infarction [31, 33]. *ACSS1* catalyzes the synthesis of acetyl-CoA and has been associated with performance on standardized cognitive tests and bitter alcoholic beverage consumption [31–33]. *SCN9A* medicates the voltage-dependent sodium ion permeability of excitable membranes and plays a role in pain mechanisms, especially in the development of inflammatory pain [31]. As it is known that cannabis can have a negative impact on learning, memory and chronic bronchitis, known relation to mental illness and suggested role in pain management, these regions may have implications in cannabis use despite having no clear known biological relevance [3, 19].

Additionally, it is also important to highlight that genes identified in this review associated with cannabis use or CUD have also been associated with other neuropsychiatric disorders namely nicotine dependence, ADHD, bipolar disorder, schizophrenia, and alcohol

Table 5 GRADE Assessment

No of studies	Certainty assessment						Effect			Certainty	Importance
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	No of events	No of individuals	Rate (95% CI)		
Cannabis use in lifetime											
2	Observational studies	Not serious	Not serious	Not serious	Not serious					–	IMPORTANT
Diagnosis of cannabis use disorder											
2	Observational studies	Not serious	Not serious	Serious [a]	Not serious					–	IMPORTANT
Criterion count of cannabis use disorder											
2	Observational studies	Not serious	Not serious	Serious [a]	Not serious		–			–	IMPORTANT
Age of onset of cannabis use											
1	Observational studies	Not serious	Not serious	Not serious	Not serious		–			–	IMPORTANT

[a] Different diagnostic classification

consumption suggesting that the genetic risk for the development of these disorders may not be independent. Previously genetic associations have been found amongst schizophrenia, bipolar disorder, ADHD, depression, and autism spectrum disorder, with a high genetic correlation between schizophrenia and bipolar disorder and a moderate correlation between ADHD and depression, ADHD and autism spectrum disorder, and ADHD and depression [34]. A recent GWAS meta-analysis added to the evidence on shared genetic associations amongst neuropsychiatric disorders by identifying that an increased risk of cannabis use disorder is genetically correlated with increased liability for smoking initiation, alcohol use, nicotine dependence, and psychiatric disorders (e.g. ADHD, schizophrenia, major depression) [35]. These genetic correlations among neuropsychiatric disorders, including cannabis, could reflect genuine pleiotropy or could indicate these psychiatric disorders, including CUD, are not completely independent [34, 35]. As such, it is important to discuss the biological and individual factors that influence the development of neuropsychiatric disorders.

Neuropsychiatric disorders are influenced by a range of factors, including genetics, personality/mood characteristics, psychological status, behaviour, neurocognitive functioning, and demographic characteristics [36, 37]. To begin, non-specific to CUD, the prenatal environment, including prenatal nutrition, maternal stress, and maternal substance abuse, can impact brain development and therefore the behavioural outcome of children. Potential mechanisms through which the prenatal environment can impact brain development occurs on multiple levels including genetic selection, epigenetic modification, mediation of brain-immune communications, abnormal metabolism pathways, synthetic mediation of hormones and the hypothalamic–pituitary–adrenal axis, and mediation of the microbiota-gut brain axis [37]. Furthermore, nutritional deficiency during critical stages of pregnancy has been linked to emotional and behavioural problems in children including decreased attention, decreased IQ, language delay, and neurodevelopment and related neuropsychiatric disorders [37–39]. More specifically, prenatal malnutrition has been linked to an increased risk of schizophrenia during the 1944–1945 Dutch Hunger Winter and the 1959–1961 Chinese famine. Additionally, a "U" relationship between serum 12(OH)D concentration and emotion, behaviour and attention has been found [38, 40]. Interestingly, the hippocampus, which plays an important role in learning and memory, has been suggested to be sensitive to the exposure of prenatal nutrition deficiency [39, 41]. The hippocampus has also been proven to be crucial in the pathophysiology of many neuropsychiatric disorders, in which the changes result from alerted brain development [41]. Maternal stress has been associated with poor offspring outcomes including cognition, health and educational attainment, however methodological challenges exist leading to potential misattribution of socially mediated (i.e. postnatal parenting) mechanisms to biological ones (i.e. alterations to developing fetal brain) [42, 43]. Finally, prenatal exposure to alcohol and other substances has been increasingly common and the consequence of the exposure differs depending on the substance used. Alcohol, tobacco, cannabis and opioids are among the most frequent used substances during pregnancy and offspring outcomes may include birth defects, developmental disability, fetal alcohol syndrome, childhood obesity, decreased birth weight, poor inhibitory control and other organ deficits [44]. Thus, many neuropsychiatric disorders appear to result from interactions among genetic background, the prenatal environment and postnatal lifestyle choices [45, 46]. Given the known association between deficits within the prenatal environment and other neuropsychiatric disorders it is plausible to suggest that the prenatal environment and subsequential gene expression may play a role in future cannabis use and/or CUD.

As previously mentioned, a variety of factors contribute to the complex etiology of neuropsychiatric disorders such as epigenetic modification. Epigenetic modifications that can regulate gene expression include DNA methylation, nucleosomal structure and positioning, post-translational modification of nucleosome histones, histone replacement and small RNA molecules that influence protein production [47]. The most studied form of epigenetic modification is DNA methylation, which can be influenced by a range of factors including genetic factors, disease, environmental exposures, and lifestyle. DNA methylation changes can be either persistent or reversible once the exposure is no longer present, adding value for biomarker development [48]. How cannabis, THC and other exogenous cannabis receptor modulators alter epigenetic mechanisms have been previously reviewed [47]. Relatively little is known about the molecular pathways influenced by cannabis, however, one study identified 13 proteins, 3 metabolites and 2 lipids significantly associated with a metabolite of THC and another found acute effects of cannabis or THC on the central nervous system and heart rate [49, 50]. In addition to DNA methylation, post-transcriptional chemical medication of RNA is rapidly emerging as a key role in regulating gene expression, known as epitranscriptomics [51]. Of growing interest within this felid is N4-acetylcytidine (ac4C), a key role in the transcriptional translation process. ac4C has been implicated in the occurrence of various disease such as inflammation, metabolic diseases, autoimmune diseases, and cancer [52]. While the role ac4C may play in neuropsychiatric disorders remains unknown, it is

important to consider the role epitransciptomics plays in the gene expression with normal development.

Current knowledge on cannabis has demonstrated that cannabis can induce structural changes to brain regions including the hippocampus, amygdala, cerebellum, prefrontal cortex and striatum as well as grey matter volume [53–55]. Potential pre-existing neurobiological factors may exist in cannabis use as well as gene x drug interactions. For instance, in young teens, reduced orbitofrontal cortex volume has been found to predict initiation of cannabis use in later adolescence. The G allele of rs2023239 of *CNR1* is linked with higher cortical CBR1 and is associated with smaller hippocampal volume in chronic cannabis users, but not healthy controls and findings that suggest only individuals with a high genetic risk of schizophrenia experience a negative impact on cortical maturation during early adolescence thus suggestive of gene × drug interactions [56–58]. In addition, functional MRI evidence suggest specific brain activity signatures with cannabis use such as increased functional connectivity associated with the default node network and insula networks and hippocampal and parahippocampal atrophy have been associated with chronic cannabis use [59, 60]. However, neuroimaging studies of cannabis users have yielded inconsistent findings and may reflect individual differences that preceded cannabis use. The inconsistent findings in the literature highlight the need for large longitudinal studies utilizing before-and-after cannabis use neuroimaging [61]. Taken together, it is plausible that structural differences in brain regions could be influenced by genetic differences between individuals, explaining the mixed evidence within neuroimaging. Further research is required to determine the complex interactions amongst individual genetic predispositions, prenatal environment, and postnatal environment contributing to individual cannabis use behaviour and/or the development of CUD. Understanding the genetic predispositions is one piece of the puzzle in understanding the complex development of cannabis use and CUD.

Finally, it is important to consider the shared genetic basis of other substance use disorders. Heritability estimates across substance use disorders vary, with heritability lowest for hallucinogens (0.39) and highest for cocaine use (0.72) [62, 63]. Additionally, substance use disorders are the result of gene x environment interactions, with partial risk inborn and another part determined by environmental experiences [62]. Previous reviews have summarized the literature on GWASs for various substance use disorders including alcohol use disorder, nicotine use disorder, CUD, OUD, and cocaine use disorder. However, genetic studies within specific substance use disorders have had varying success in replicating previously identified associations, limiting evidence for shared genetic basis across substance use disorders [63, 64]. The complexity of substance use disorder make genetic prediction efforts difficult, and while currently only alcohol use disorder have been genetically correlated with CUD, continued advancements in molecular genetic studies and substance use disorder at larges further our understanding of the biological pathways underlying substance use disorders [9, 63, 65]. For instance, *CNR1* and *CNR2*, components of the endocannabinoid system, are major targets of investigation for their impact in neuropsychiatry and addiction phenotypes suggested shared genetic risk factors [66, 67]. In regards to neuropsychiatric disorders, Mendelian randomization studies have found mixed evidence on the causal effect of cannabis initiation and schizophrenia, finding weak evidence that cannabis initiation increases schizophrenia risk and strong evidence that schizophrenia liability increases the odds of cannabis initiation, and causal evidence of ADHD on cannabis initiation [68–72]. Through continued advances, it is hoped that the underlying genetic basis for CUD, or a shared genetic basis for all substance use disorders, will be identified to provide preventative measures and treatment for substance use disorders in the future.

Limitations

While this systematic review was rigorous and involved a peer-reviewed protocol, it is not without limitations. First, our inclusion criteria limited our review to only GWASs, meaning any GWAS meta-analyses and candidate gene studies were excluded. GWAS meta-analyses and candidate gene studies are often more powered due to their larger sample sizes and minimal genetic variants tested, respectfully [11]. However, including only GWASs was decided a priori to capture novel genetic variants associated with cannabis use and avoid the inclusion of multiple studies which could use the same genetic dataset. Second, it is important to note that this review is susceptible to publication bias, as studies that do not achieve genome-wide significance may be less likely to be published, and thus, not included in this review. Unpublished GWAS findings may exist with SNPs reaching the borderline significance threshold. While we cannot eliminate publication bias entirely, we searched abstracts, GWAS catalogs, and databases for any near significant findings that were not published. If a relevant abstract was identified, without the full study published, the first author was contacted to determine whether the full GWAS had been published or was going to be submitted to a journal. Finally, if a study met our inclusion criteria but did not report any SNPs that fell below the genome-wide significance threshold, study authors were contacted to confirm if any SNPs had reached the borderline significant threshold

set for this review. Third, due to the heterogeneity of the reported findings, it was not possible to conduct a meta-analysis or sex and ethnicity subgroup analyses. Although we could not conduct a meta-analysis, we qualitatively summarized the studies and reported a comprehensive list of all SNPs reaching the significance threshold for this study.

Conclusion

This systematic review was able to summarize GWAS findings within the field of cannabis use. The results can inform future candidate gene studies and GWASs of possible replicable SNPs that require further investigation. We were able to identify all GWASs conducted on cannabis use, highlighting the need for further research as no two GWASs reported the same SNP or gene associated with cannabis use. Further, included GWASs had limited ethnic diversity, with only European or African American participants. Recommendations are made for future research to replicate reported associations and include diverse ethnic populations to test whether SNPs associated with cannabis use reported are generalizable across study populations and if associations differ by ethnicity.

Abbreviations

GWAS: Genome-wide Association Study; CUD: Cannabis Use Disorder; GWASs: Genome-wide Association Studies; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; HuGeNET: The Human Genome Epidemiology Network; PROSPERO: International Prospective Register of Systematic Reviews; DSM: Diagnostic and Statistical Manual; ICD-10: International Statistical Classification of Diseases and Related Health Problems -10; Q-Genie: The quality of genetic association studies; SAGE: Study of Addiction: Genetics and Environment; ICC: International Cannabis Consortium; ICGHD: International Consortium on the Genetics of Heroin Dependence; iPSYCH: Integrative Psychiatric Research; ADHD: Attention-Deficit Hyperactivity Disorder.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12920-021-01035-5.

Additional File 1. PRISMA 2009 checklist.

Acknowledgements

Not applicable.

Authors' contributions

ZS is the guarantor. AH, CC and ZS conceptualized the systematic review. AH implemented the design of the review and search strategy with the aid of CC and SS. AH, CC, AD, MB and RK screened studies, extracted data and assess the quality of the studies. AH prepared the first draft, and the final manuscript was reviewed and revised by CC, AD, MB, RK, FK, LT and ZS. All authors read and approved the final manuscript.

Funding

This systematic review was partially funded by CIHR (Grant number PJT-156306). The funding agency had no role in the study design, interpretation of data, or dissemination of results.

Availability of data and materials

All data generated or analysed during this study are included in this published article. Primary articles included in this systematic review can be found at the following links: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1369-1600.2010.00255.x. https://www.sciencedirect.com/science/article/abs/pii/S0376871613004766. https://www.nature.com/articles/s41593-019-0416-1. https://link.springer.com/article/10.1007/s10519-015-9723-9. https://www.nature.com/articles/s41593-018-0206-1?fbclid=IwAR20cNM7ZDBKTGp7QNtm521tkEECrKyDTFLQ5pQV9emvsIbrsM_TMeAUTSg. https://jamanetwork.com/journals/jamapsychiatry/article-abstract/2504223.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare they have no competing interests.

Author details

[1]Neuroscience Graduate Program, Department of Psychiatry and Behavioural Neurosciences, McMaster University, 100 West 5th St., Hamilton, ON L8N 3K7, Canada. [2]Health Science Library, McMaster University, 1280 Main St. W., Hamilton, ON L8S 4L8, Canada. [3]Integrated Science Program, McMaster University, 1280 Main St. W., Hamilton, ON L8S 4L8, Canada. [4]Michael G. DeGroote School of Medicine, McMaster University, 1280 Main St. W., Hamilton, ON L8S 4L8, Canada. [5]Department of Health Research Method, Evidence and Impact, 1280 Main St. W., Hamilton, ON L8S 4L8, Canada.

Received: 31 March 2021 Accepted: 15 July 2021
Published online: 12 August 2021

References

1. Degenhardt L, Charlson F, Ferrari A, Santomauro D, Erskine H, Mantilla-Herrara A, et al. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Psychiatry. 2018;5(12):987–1012.
2. Statistics Canada. National Cannabis Survey, forth quarter 2019. 2019.
3. Volkow ND, Baler RD, Compton WM, Weiss SRB. Adverse health effects of Marijuana use. N Engl J Med. 2014;370(23):2219–27.
4. Sherva R, Wang Q, Kranzler H, Zhao H, Koesterer R, Herman A, et al. Genome-wide association study of cannabis dependence severity, novel risk variants, and shared genetic risks. JAMA Psychiat. 2016;73(5):472–80.
5. Bahji A, Stephenson C. International perspectives on the implications of cannabis legalization: a systematic review & thematic analysis. Int J Environ Res Public Health. 2019;16(17):3095.
6. Verweij KJH, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG, et al. Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. Addiction. 2010;105(3):417–30.
7. Agrawal A, Chou YL, Carey CE, Baranger DAA, Zhang B, Sherva R, et al. Genome-wide association study identifies a novel locus for cannabis dependence. Mol Psychiatry. 2018;23(5):1293–302.
8. Minică CC, Verweij KJH, Most PJ, Mbarek H, Bernard M, Eijk KR, et al. Genome-wide association meta-analysis of age at first cannabis use. Addiction. 2018;113(11):2073–86.
9. Pasman JA, Verweij KJH, Gerring Z, Stringer S, Sanchez-Roige S, Treur JL, et al. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. Nat Neurosci. 2018;21(9):1161–70. https://doi.org/10.1038/s41593-018-0206-1.
10. Stringer S, Minică CC, Verweij KJH, Mbarek H, Bernard M, Derringer J, et al. Genome-wide association study of lifetime cannabis use based on a large

meta-analytic sample of 32 330 subjects from the International Cannabis Consortium. Transl Psychiatry. 2016;6(3):e769.
11. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019;20(8):467–84.
12. The Government of Canada. Cannabis Health Effects. 2019.
13. Hall W. What has research over the past two decades revealed about the adverse health effects of recreational cannabis use? Addiction. 2015;110(1):19–35.
14. Gobbi G, Atkin T, Zytynski T, Wang S, Askari S, Boruff J, et al. Association of cannabis use in adolescence and risk of depression, anxiety, and suicidality in young adulthood: a systematic review and meta-analysis. JAMA Psychiat. 2019;76(4):426–34.
15. Moher D, Liberati A, Tetzlaff J, Altman D, Group P, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. PLoS Med. 2009;6(7):e1000097.
16. Little J, Higgins JPT, Bray M, Ioannidis J, Khoury M, Manolio T, et al. The HuGENet™ HuGE review handbook, version 1.0. Ottawa, Ontario, Canada HuGENet Canada Coord Cent. 2006;
17. Hillmer A, Chawar C, D'Elia A, Kapoor R, Butt M, Samaan Z. A systematic review and meta-anlaysis of genome-wide association studies and cannabis use. PROSPERO: International prospective register of systematic reivews. 2020.
18. Hillmer A, Chawar C, Sanger S, D'Elia A, Butt M, Kapoor R, et al. Genetic determinants of cannabis use: a systematic review protocol. Syst Rev. 2020;9(1):1–6.
19. National Academies of Sciences Engineering and Medicine. The health effects of cannabis and cannabinoids: The current state of evidence and recommendations for research [Internet]. Washington, DC: National Academies Press; 2017. Available from: https://www.ncbi.nlm.nih.gov/books/NBK423845/. https://doi.org/10.17226/24625
20. Veritas Health Innovation. Covidence systematic review software [Internet]. Melbourne, Australia; Available from: https://www.covidence.org/
21. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. Genet Epidemiol Off Publ Int Genet Epidemiol Soc. 2008;32(3):227–34.
22. Sohani ZN, Meyre D, de Souza RJ, Joseph PG, Gandhi M, Dennis BB, et al. Assessing the quality of published genetic association studies in meta-analyses: The quality of genetic studies (Q-Genie) tool. BMC Genet. 2015;16.
23. Guyatt GH, Oxman AD, Schünemann H, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol. 2011;64(4):380–2.
24. Foroutan F, Guyatt G, Zuk V, Vandvik PO, Alba AC, Mustafa R, et al. GRADE Guidelines 28: Use of GRADE for the assessment of evidence about prognostic factors: rating certainty in identification of groups of patients with different absolute risks. J Clin Epidemiol. 2020;121:62–70.
25. Center for History New Media. Zotero: The Next-Generation Research Tool [Internet]. George Mason University; 2009. Available from: http://www.zotero.org/
26. Agrawal A, Lynskey MT, Hinrichs A, Grucza R, Saccone SF, Krueger R, et al. A genome-wide association study of DSM-IV cannabis dependence. Addict Biol. 2011;16(3):514–8.
27. Demontis D, Rajagopal VM, Thorgeirsson TE, Als TD, Grove J, Leppälä K, et al. Genome-wide association study implicates CHRNA2 in cannabis use disorder. Nat Neurosci. 2019;22(7):1066.
28. Agrawal A, Lynskey MT, Bucholz KK, Kapoor M, Almasy L, Dick DM, et al. DSM-5 cannabis use disorder: a phenotypic and genomic perspective. Drug Alcohol Depend. 2014;134(1):362–9. https://doi.org/10.1016/j.drugalcdep.2013.11.008.
29. Minică CC, Dolan CV, Hottenga J-J, Pool R, Fedko IO, Mbarek H, et al. Heritability, SNP-and gene-based analyses of cannabis use initiation and age at onset. Behav Genet. 2015;45(5):503–13.
30. Ziyatdinov A, Kim J, Prokopenko D, Privé F, Laporte F, Loh P-R, et al. Estimating the effective sample size in association studies of quantitative traits. bioRxiv. 2019;
31. Bethesda (MD): National Library of Medicine (US) National Centre for Biotechnology Infromation. dbGene [Internet]. 2004. Available from: https://www.ncbi.nlm.nih.gov/gene/
32. UniProt Consortium. UniProt: a worldwide hub of protien knowledge. Nucleic Acids Res. 2019;47(D1):D506–15.
33. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. Ensembl 2020. Nucleic Acids Res. 2020;48(D1):D682–8.
34. Wu Y, Cao H, Baranova A, Huang H, Li S, Cai L, et al. Multi-trait analysis for genome-wide association study of five psychiatric disorders. Transl Psychiatry. 2020;10(1):1–11.
35. Johnson EC, Demontis D, Thorgeirsson TE, Walters RK, Polimanti R, Hatoum AS, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. Lancet Psychiatry. 2020;7(12):1032–45.
36. Taylor M, Cousijn J, Filbey F. Determining risks for cannabis use disorder in the face of changing legal policies. Curr Addict Rep. 2019;6(4):466–77.
37. Yan X, Zhao X, Li J, He L, Xu M. Effects of early-life malnutrition on neurodevelopment and neuropsychiatric disorders and the potential mechanisms. Prog Neuro-Psychopharmacol Biol Psychiatry. 2018;83:64–75.
38. Wang X, Jiao X, Xu M, Wang B, Li J, Yang F, et al. Effects of circulating vitamin D concentrations on emotion, behavior and attention: a cross-sectional study in preschool children with follow-up behavior experiments in juvenile mice. J Affect Disord. 2020;275:290–8.
39. Chen J, Zhao X, Cui L, He G, Wang X, Wang F, et al. Genetic regulatory subnetworks and key regulating genes in rat hippocampus perturbed by prenatal malnutrition: implications for major brain disorders. Aging (Albany NY). 2020;12(9):8434.
40. Xu M-Q, Sun W-S, Liu B-X, Feng G-Y, Yu L, Yang L, et al. Prenatal malnutrition and adult schizophrenia: further evidence from the 1959–1961 Chinese famine. Schizophr Bull. 2009;35(3):568–76.
41. Oberlander VC, Xu X, Chini M, Hanganu-Opatz IL. Developmental dysfunction of prefrontal–hippocampal networks in mouse models of mental illness. Eur J Neurosci. 2019;50(6):3072–84.
42. Aizer A, Stroud L, Buka S. Maternal stress and child outcomes: Evidence from siblings. J Hum Resour. 2016;51(3):523–55.
43. DiPietro JA. Maternal stress in pregnancy: considerations for fetal development. J Adolesc Heal. 2012;51(2):S3-8.
44. Chang G. Maternal substance use: consequences, identification, and interventions. Alcohol Res. 2020;40(2):6.
45. Newnham JP, Moss TJM, Nitsos I, Sloboda DM, Challis JRG. Nutrition and the early origins of adult disease. Asia Pac J Clin Nutr. 2002;11:S537–42.
46. Taylor PD, Poston L. Developmental programming of obesity in mammals. Exp Physiol. 2007;92(2):287–98.
47. Szutorisz H, Hurd YL. Epigenetic effects of Cannabis exposure. Biol Psychiatry. 2016;79(7):586–94.
48. Markunas CA, Hancock DB, Xu Z, Quach BC, Fang F, Sandler DP, et al. Epigenome-wide analysis uncovers a blood-based DNA methylation biomarker of lifetime cannabis use. Am J Med Genet Part B Neuropsychiatr Genet. 2021;186(3):173–82.
49. Hinckley JD, Saba L, Raymond K, Bartels K, Klawitter J, Christians U, et al. An approach to biomarker discovery of cannabis use utilizing proteomic, metabolomic, and lipidomic analyses. Cannabis Cannabinoid Res. 2020;
50. Zuurman L, Ippel AE, Moin E, Van Gerven JMA. Biomarkers for the effects of cannabis and THC in healthy volunteers. Br J Clin Pharmacol. 2009;67(1):5–21.
51. Karthiya R, Wasil SM, Khandelia P. Emerging role of N4-acetylcytidine modification of RNA in gene regulation and cellular functions. Mol Biol Rep. 2020;47:1–11.
52. Jin G, Xu M, Zou M, Duan S. The processing, gene regulation, biological functions, and clinical relevance of N4-acetylcytidine on RNA: a systematic review. Mol Ther Acids. 2020;20:13–24.
53. Battistella G, Fornari E, Annoni J-M, Chtioui H, Dao K, Fabritius M, et al. Long-term effects of cannabis on brain structure. Neuropsychopharmacology. 2014;39(9):2041–8.
54. Scheffler F, Du Plessis S, Asmal L, Kilian S, Phahladira L, Luckhoff HK, et al. Cannabis use and hippocampal subfield volumes in males with a first episode of a schizophrenia spectrum disorder and healthy controls. Schizophr Res. 2021;231:13–21.
55. Batalla A, Bhattacharyya S, Yucel M, Fusar-Poli P, Crippa JA, Nogue S, et al. Structural and functional imaging studies in chronic cannabis users: a systematic review of adolescent and adult findings. PLoS ONE. 2013;8(2):e55821.
56. Schacht JP, Hutchison KE, Filbey FM. Associations between cannabinoid receptor-1 (CNR1) variation and hippocampus and amygdala volumes in heavy cannabis users. Neuropsychopharmacology. 2012;37(11):2368–76.

57. French L, Gray C, Leonard G, Perron M, Pike GB, Richer L, et al. Early cannabis use, polygenic risk score for schizophrenia and brain maturation in adolescence. JAMA Psychiat. 2015;72(10):1002–11.
58. Cheetham A, Allen NB, Whittle S, Simmons JG, Yücel M, Lubman DI. Orbitofrontal volumes in early adolescence predict initiation of cannabis use: a 4-year longitudinal and prospective study. Biol Psychiatry. 2012;71(8):684–92.
59. Bloomfield MAP, Hindocha C, Green SF, Wall MB, Lees R, Petrilli K, et al. The neuropsychopharmacology of cannabis: a review of human imaging studies. Pharmacol Ther. 2019;195:132–61.
60. Ferland J-MN, Hurd YL. Deconstructing the neurobiology of cannabis use disorder. Nat Neurosci. 2020;23(5):600–10.
61. Volkow ND, Swanson JM, Evins AE, DeLisi LE, Meier MH, Gonzalez R, et al. Effects of cannabis use on human behavior, including cognition, motivation, and psychosis: a review. JAMA Psychiat. 2016;73(3):292–7.
62. Ducci F, Goldman D. The genetic basis of addictive disorders. Psychiatr Clin. 2012;35(2):495–519.
63. Deak JD, Johnson EC. Genetics of substance use disorders: a review. Psychol Med. 2021;1–12.
64. Prom-Wormley EC, Ebejer J, Dick DM, Bowers MS. The genetic epidemiology of substance use disorder: a review. Drug Alcohol Depend. 2017;180:241–59.
65. Zhou H, Sealock JM, Sanchez-Roige S, Clarke T-K, Levey DF, Cheng Z, et al. Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. Nat Neurosci. 2020;23(7):809–18.
66. Onaivi ES, Ishiguro H, Sgro S, Leonard CM. Cannabinoid receptor gene variations in drug addiction and neuropsychiatric disorders. J Drug Alcohol Res. 2013;2(1):1–11.
67. Marco EM, García-Gutiérrez MS, Bermúdez-Silva F-J, Moreira F, Guimarães F, Manzanares J, et al. Endocannabinoid system and psychiatry: in search of a neurobiological basis for detrimental and potential therapeutic effects. Front Behav Neurosci. 2011;5:63.
68 Gage SH, Jones HJ, Burgess S, Bowden J, Davey Smith G, Zammit S, et al. Assessing causality in associations between cannabis use and schizophrenia risk: a two-sample Mendelian randomization study. Psychol Med. 2017;47(5):971–80.
69. Vaucher J, Keating BJ, Lasserre AM, Gan W, Lyall DM, Ward J, et al. Cannabis use and risk of schizophrenia: a Mendelian randomization study. Mol Psychiatry. 2018;23(5):1287–92.
70. Johnson EC, Hatoum AS, Deak JD, Polimanti R, Murray RM, Edenberg HJ, et al. The relationship between cannabis and schizophrenia: a genetically informed perspective. Addiction. 2021;
71. Treur JL, Demontis D, Smith GD, Sallis H, Richardson TG, Wiers RW, et al. Investigating causality between liability to ADHD and substance use, and liability to substance use and ADHD risk, using Mendelian randomization. Addict Biol. 2021;26(1):e12849.
72. Artigas MS, Sánchez-Mora C, Rovira P, Richarte V, Garcia-Martínez I, Pagerols M, et al. Attention-deficit/hyperactivity disorder and lifetime cannabis use: genetic overlap and causality. Mol Psychiatry. 2020;25(10):2493–503.

Publisher's Note

8.6 CHAPTER 4 Supplementary File 1

Reporting checklist for genetic association study.

Based on the STREGA guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STREGAreporting guidelines, and cite them as:

Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, Khoury MJ, Cohen B, Davey-Smith G, Grimshaw J, Scheet P, Gwinn M, Williamson RE, Zou GY, Hutchings K, Johnson CY, Tait V, Wiens M, Golding J, van Duijn C, McLaughlin J, Paterson A, Wells G, Fortier I, Freedman M, Zecevic M, King R, Infante-Rivard C, Stewart A, Birkett N; STrengthening the REporting of Genetic Association Studies. STrengthening the REporting of Genetic Association Studies (STREGA): An Extension of the STROBE Statement.

		Reporting Item	Page Number
Title and abstract			1-2
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	

Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	
Background/rationale			5-6
	#2	Explain the scientific background and rationale for the investigation being reported	
Objectives			7
	#3	State specific objectives, including any prespecified hypotheses. State if the study is the first report of a genetic association, a replication effort, or both.	
Study design			7-8
	#4	Present key elements of study design early in the paper	
Setting			7-8
	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	

Eligibility criteria			8
	#6a	Cohort study – Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up. Case-control study – Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls. Cross-sectional study – Give the eligibility criteria, and the sources and methods of selection of participants. Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant.	
	#6b	Cohort study – For matched studies, give matching criteria and number of exposed and unexposed. Case-control study – For matched studies, give matching criteria and the number of controls per case.	

Variables			9
	#7a	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	
	#7b	Clearly define genetic exposures (genetic variants) using a widely-used nomenclature system. Identify variables likely to be associated with population stratification (confounding by ethnic origin).	
Data sources/measurement			9-11 Supplementary File 2
	#8a	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	

	#8b	Describe laboratory methods, including source and storage of DNA, genotyping methods and platforms (including the allele calling algorithm used, and its version), error rates and call rates. State the laboratory / centre where genotyping was done. Describe comparability of laboratory methods if there is more than one group. Specify whether genotypes were assigned using all of the data from the study simultaneously or in smaller batches.	
Bias			11-12
	#9a	Describe any efforts to address potential sources of bias	
	#9b	Describe any efforts to address potential sources of bias	
Study size			12
	#10	Explain how the study size was arrived at	

Quantitative variables			9-10
	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why. If applicable, describe how effects of treatment were dealt with.	
Statistical methods			13 Supplementary File 2
	#12a	Describe all statistical methods, including those used to control for confounding. State software version used and options (or settings) chosen.	
	#12b	Describe any methods used to examine subgroups and interactions	
	#12c	Explain how missing data were addressed	
	#12d	If applicable, explain how loss to follow-up was addressed	

	#12e	Describe any sensitivity analyses	
	#12f	State whether Hardy-Weinberg equilibrium was considered and, if so, how.	
	#12g	Describe any methods used for inferring genotypes or haplotypes	
	#12h	Describe any methods used to assess or address population stratification.	
	#12i	Describe any methods used to address multiple comparisons or to control risk of false positive findings.	
	#12j	Describe any methods used to address and correct for relatedness among subjects	
Participants			14
	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed.	

		Give information separately for for exposed and unexposed groups if applicable. Report numbers of individuals in whom genotyping was attempted and numbers of individuals in whom genotyping was successful.	
	#13b	Give reasons for non-participation at each stage	
	#13c	Consider use of a flow diagram	
Descriptive data			14-16
	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable. Consider giving information by genotype	
	#14b	Indicate number of participants with missing data for each variable of interest	

	#14c	Cohort study – Summarize follow-up time, e.g. average and total amount.	
Outcome data			14-16
	#15	Cohort study Report numbers of outcome events or summary measures over time.Give information separately for exposed and unexposed groups if applicable. Report outcomes (phenotypes) for each genotype category over time Case-control study – Report numbers in each exposure category, or summary measures of exposure.Give information separately for cases and controls . Report numbers in each genotype category. Cross-sectional study – Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable. Report outcomes (phenotypes) for each genotype category	
Main results			16-21

	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
	#16b	Report category boundaries when continuous variables were categorized	
	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
	#16d	Report results of any adjustments for multiple comparisons	
Other analyses			18-19, 21-22
	#17a	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	
	#17b	Report other analyses done—e.g., analyses of	

		subgroups and interactions, and sensitivity analyses	
	#17c	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	
Key results			23
	#18	Summarise key results with reference to study objectives	
Limitations			25-26
	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	
Interpretation			23-25
	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	
Generalisability			25-26

	#21	Discuss the generalisability (external validity) of the study results	
Funding			26
	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	

8.7 CHAPTER 4 Supplementary File 2

Genetics of cannabis use in opioid use disorder: A Polygenic Risk Score and Genome-Wide Association Study

Supplementary File 2

Data clean-up and quality control
Study Descriptions
The following clean-up and quality control steps were conducted on the Pilot GENOA (n=182), GENOA (n=1,314) and POST data (n=3,125) genetic samples. The Pilot GENOA, GENOA and POST were all prospective cohort studies designed to identify factors associated with opioid use and treatment outcomes including genetic risk factors in patients diagnosed with Opioid Use Disorder and receiving treatment. The Pilot GENOA and GENOA data were merged into and analyzed as one dataset, and are referred to as the GENOA data (1,2). The POST data included 775 samples from a collaborating site, all of which were excluded at the post-imputation stage as none were of European ancestry. While analyzed separately, the quality control steps taken for the GENOA datasets and the POST dataset pre-imputation were identical. The GENOA datasets and POST data set were merged and analyzed together post-imputation. All analyses were performed on PLINK 1.90 and the RStudio interface of R Version 1.1.453 (3–5).

Collection of DNA and genotyping
As part of the GENOA study, whole blood samples were collected. Blood samples were centrifuged, separated and frozen in -20°C within 2 hours of collection at the clinics and then transferred to -80°C freezers located at McMaster University within 1 month of collection. As part of the POST study, approximately 2ml of saliva samples were collected at the baseline using DNAgenotek all-in-one system for the collection, stabilization and transportation of DNA from saliva (OGR-500) (6). DNA was extracted from blood or saliva samples (7) and genotyped by Genomé Quebec using GenomeStudio (v 2.0.4) and the Infinium Global Screening Array – 24 v1.0 (8–10).

Log-transformed Phenotype data
For the outcome of heaviness of use and cannabis dependence, the log of the raw data was used to approach a normal distribution. The data was transformed sing the RStudio interface of R Version 1.1.453 (3,4) using the log10 of the raw phenotype plus 1. A histogram of the transformed data for each outcome is below.

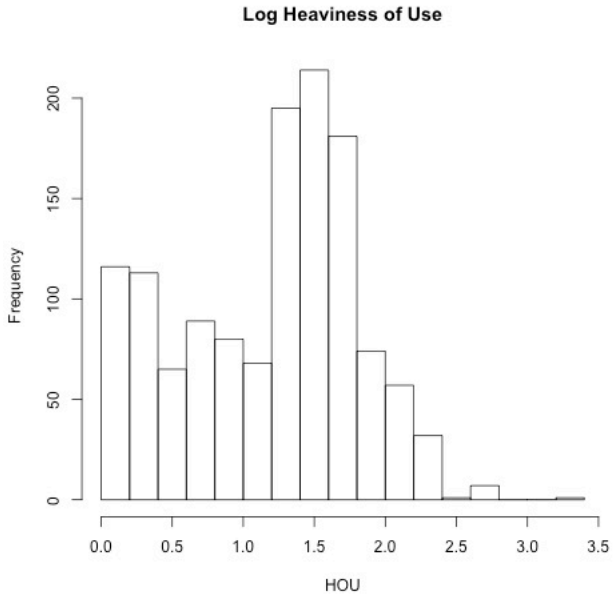


Figure S1. Histogram of the Log of heaviness of use

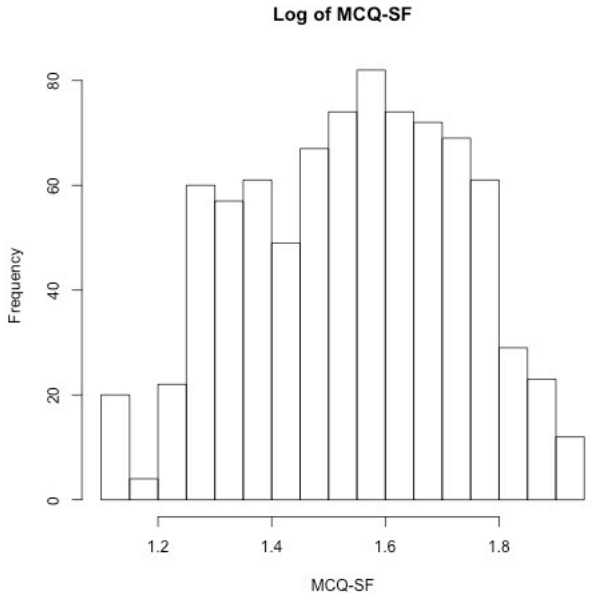


Figure S2. Histogram of the Log of cannabis dependence using the Marijuana Cravings Questionnaire – Short Form

Quality Control

The genotyped files were converted into .bed, .bim, and .fam files and merged into one dataset and chromosomes that failed genotyping were removed. All samples that were genotyped were cross-referenced with the sample shipment documents to ensure that there were no missing samples.
Missingness per sample and per SNP was estimated using PLINK's --missing flag. Samples and SNPs with more than 10% (at early QC stages), and 5%, (later in the QC process) were removed.
Four sets of samples were genotyped twice across the Pilot and GENOA data accidently. For each duo of duplicates, the samples with the lowest missingness rates were kept.
Samples with discordant sex information were identified using the --check-sex flag in PLINK. Chromosome X's inbreeding coefficient was graphed for males and females separately. Males with a coefficient ≥ 0.8 were kept; females with a coefficient ≤ 0.4 were kept.
The sample heterozygosity rates were checked. The resultant values of the heterozygosity rate were calculated using the equation "(N(NM)-O(Hom))/N(NM)" (Number of autosomal genotype observations minus observed number of homozygotes divided by the number of autosomal genotype observations). A histogram was graphed of the heterozygosity rate, and the threshold was determined to be 0.22. Samples with a calculated rate of less than or equal to 0.17 were checked to be of Native American ancestry. Heterozygous haploids and nonmale Y chromosome genotype calls were set as missing.

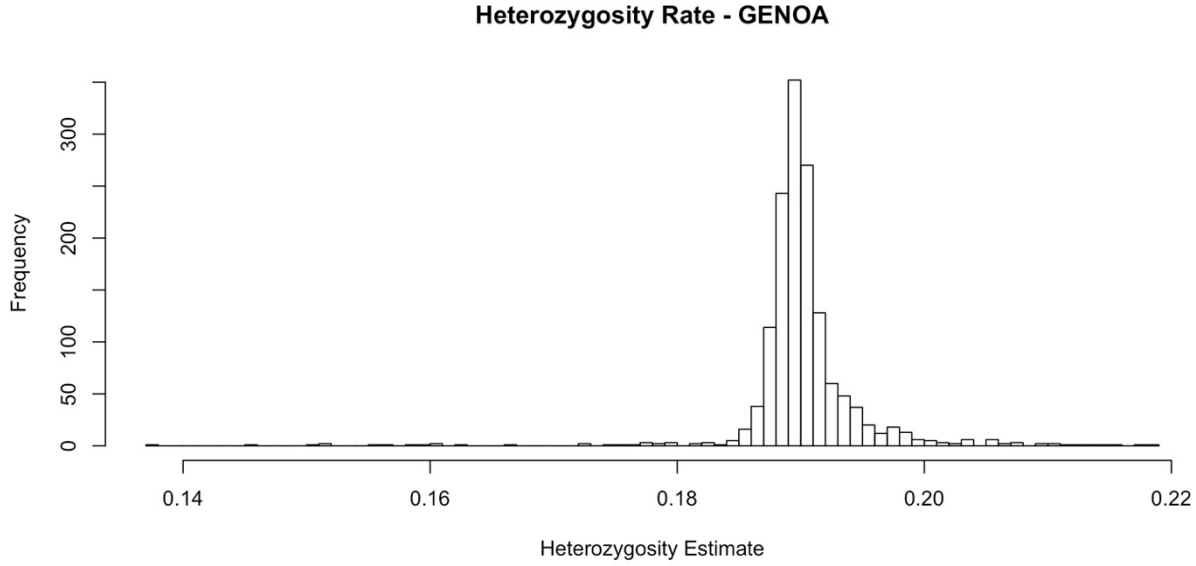


Figure S3. GENOA Heterozygosity Plot

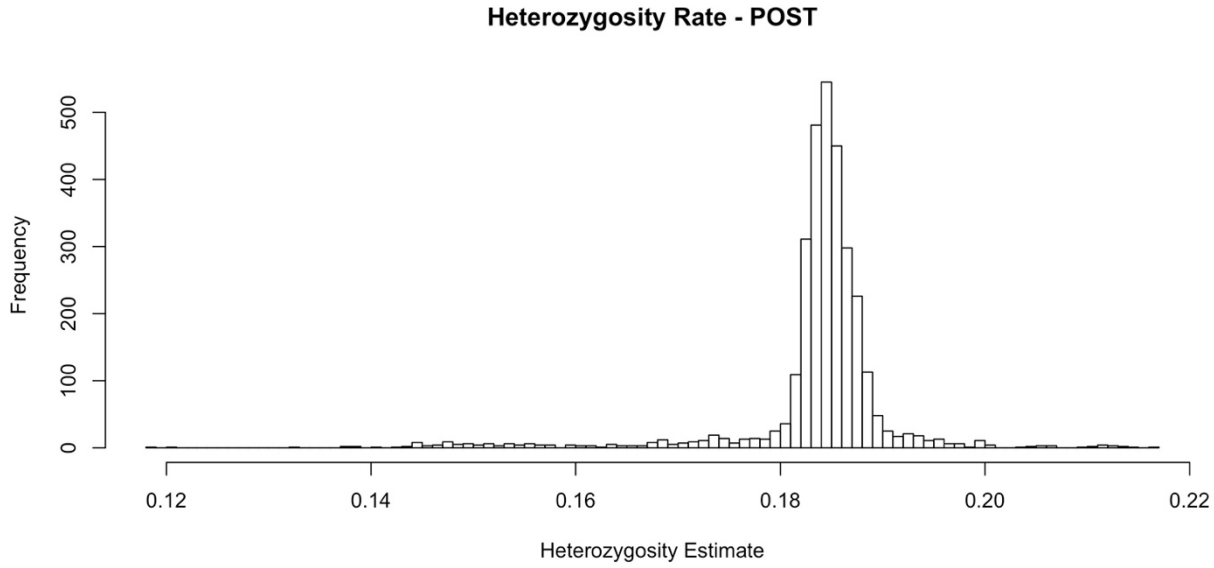


Figure S4. POST Heterozygosity Plot

A principal component analysis was conducted on all samples as part of the ethnicity checks, following the pruning of areas of high range LD (within a 50kb range and r^2 threshold of 0.2). This was conducted in GENOA and POST samples separately, without the use of a reference panel. The self-reported ancestries of the samples were plotted against the genetically determined ancestries (through principal component vectors) to visually highlight any outliers. Outliers were defined as samples whose genetically determined ancestries fall too far from the self-reported ancestries. Samples whose ancestries were corrected were those that were determined to possibly partially belong to the genetically determined ancestry group (ex. self-reported as 'European' but is 'mixed European and Native North American'). Samples that failed the ethnicity check were removed.
Samples with high relatedness values (PLINK's --genome output PI_HAT>=0.2) were identified. Along with samples that were believed to be duplicates, have failed the sex check, ethnicity check, and/or genotyping, they were visualized on their respective plate positions to see if any unusual patterns could be observed. Any newly identified duplicates were checked against the case report forms to verify their duplicate status. All verified duplicates were then removed.

Pre-imputation and imputation
To prepare the data for imputation by the TOPMed Imputation Server (11), the following steps were run on a Linux operating system. Since at that stage only the European ancestry subset had a sample size large enough for the purpose of our analysis (other ancestries of less than 100 samples would not be powered enough for ancestry-stratified analysis), only samples of European ancestry were submitted for imputation and later analyzed.

The reference alleles for the European ancestry subset were set up to match those from HRC reference panel, follow the steps on the McCarthy Group Tools site (V4.2.11) (12). The frequency file used for the 1000 Genomes Phase 3 match was taken from the McCarthy Group tools (https://www.well.ox.ac.uk/~wrayner/tools/) (V4.2.11) (12). SNPs with high MAF (MAF>0.4), differing alleles, not in the reference panel or with an allele frequency difference of >0.2 were removed.
Phasing was done using Eagle2, using TOPMed (13,14).

Post-imputation filtering and quality control
The following steps were performed using a virtual machine instance and cloud storage supported by the Google Cloud Platform (https://console.cloud.google.com/) (15). Imputed individual chromosome files were recoded from .vcf to .ped/.map files, and then to .bed/.bim/.fam files before being merged into one file on PLINK for easy handling. The Rsq values were used for filtering. SNPs with equal to or less than 0.3 Rsq were identified to be of low quality and removed. Further, SNPs with MAF<0.05 were removed.
As the GENOA datasets and the POST dataset were merged at this time, duplicates and first and second-degree related individuals were removed, with a prior decision to keep POST samples over GENOA samples due to robustness of the phenotype data within the POST study.

Principle Component Analysis
To account for population stratification a Principal Component Analysis (PCA) was conducted with data prior to data imputation. Data from the GENOA datasets and POST dataset were combined, and duplicates/related individuals identified post-imputation were removed. Data from the GENOA datasets and POST dataset were merged with the 1000Genome dataset to check for ethnic outliers. The GENOA, POST and 1000Genome datasets were checked for strand flips and corrected. Areas of high linkage disequilibrium and regions of long-range of long-range LD, as reported in the UKBiobank supplementary information table S13 (16), were removed prior to conducting the PCA. Results from the cleaned PCA can be found below in Figure S5-7

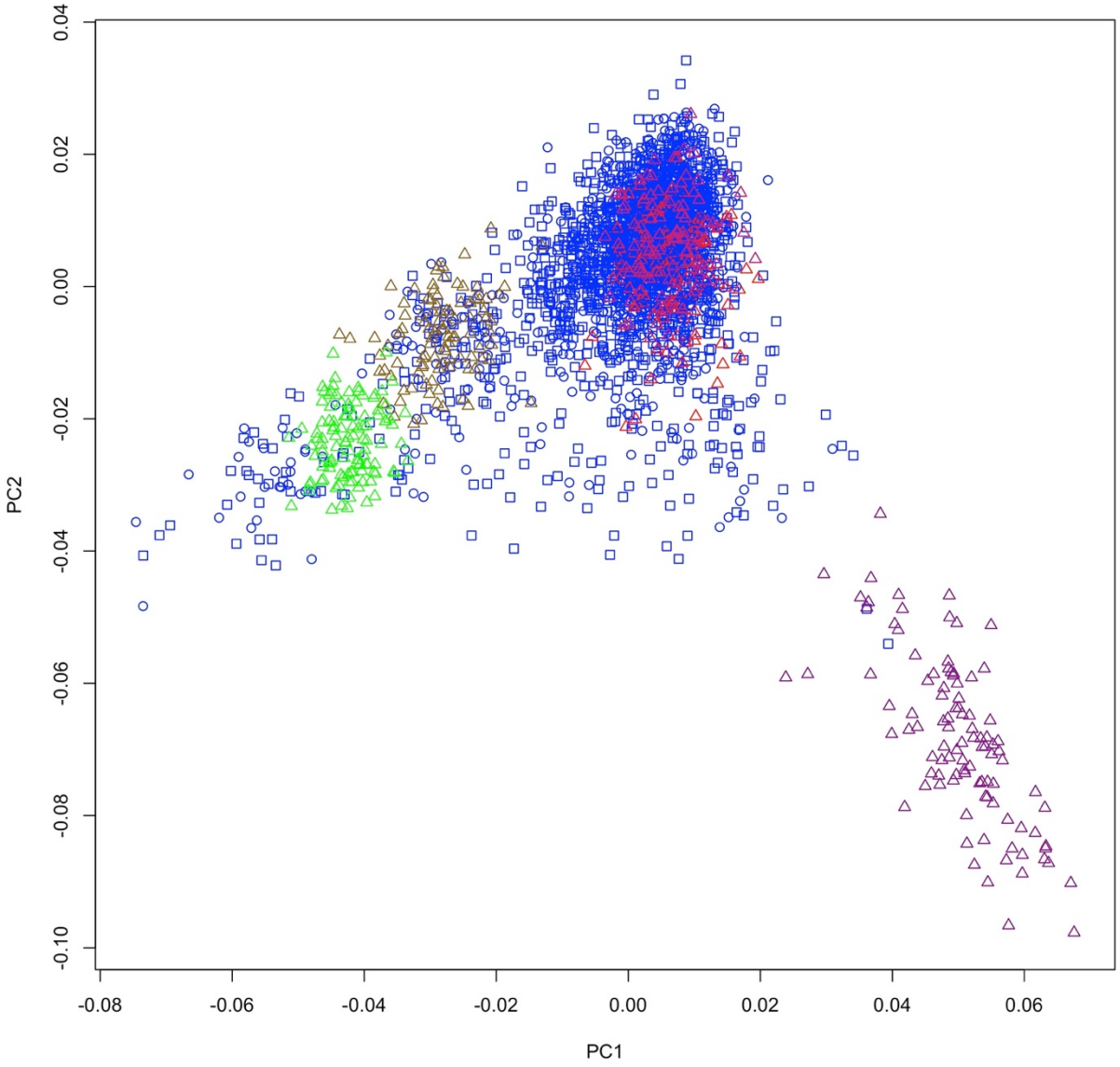


Figure S5. PC1 and PC2 for GENOA, POST and 1000 Genome Data
Legend: Circles = GENOA data, Squares = POST data, Triangles = 1000 Genome data, Blue = European Ancestry, Yellow = African American Ancestry, Red = Admixed American, Green = East Asian Ancestry, Purple = South Asian Ancestry

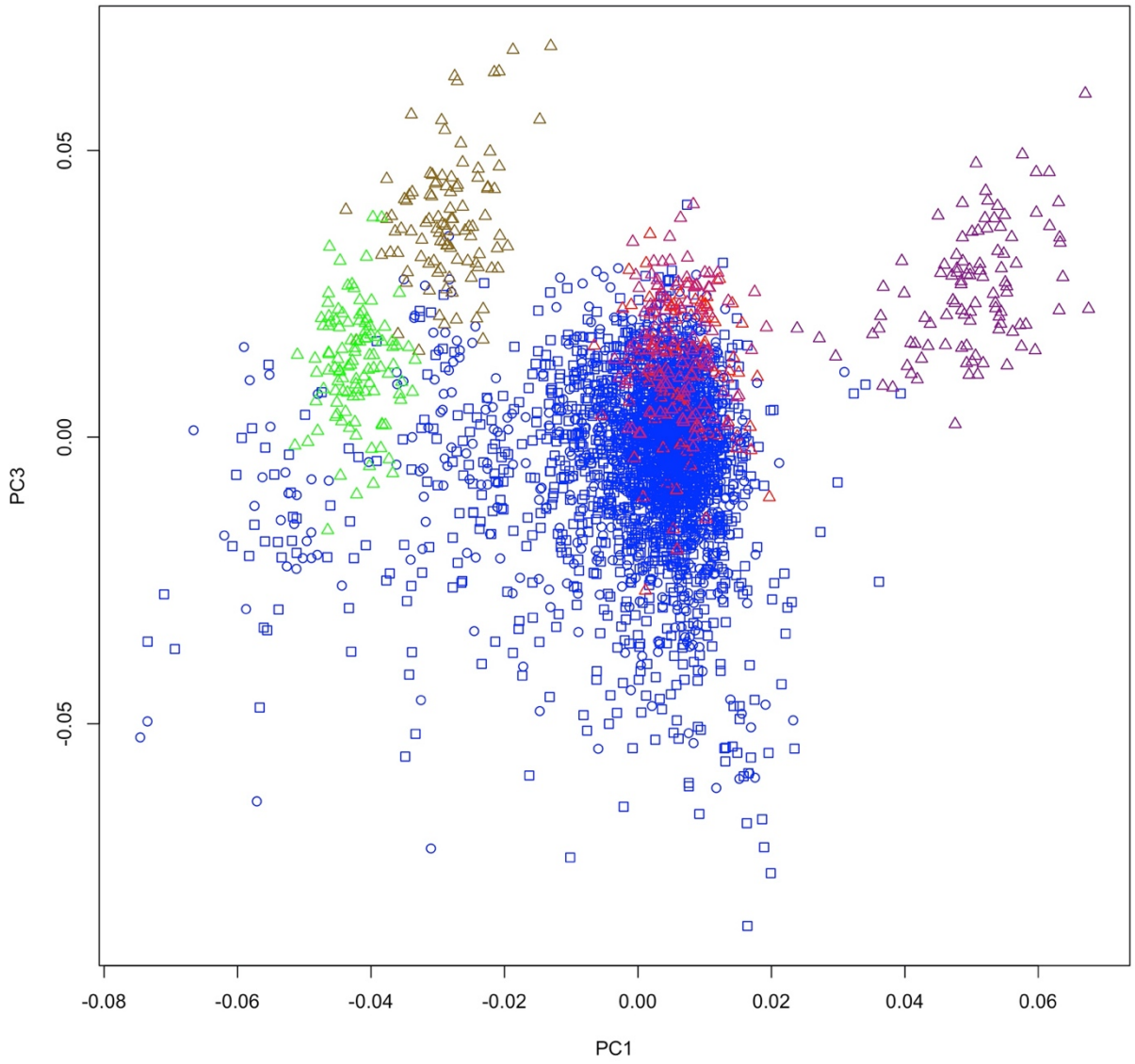


Figure S6. PC1 and PC3 for GENOA, POST and 1000 Genome Data
Legend: Circles = GENOA data, Squares = POST data, Triangles = 1000 Genome data, Blue = European Ancestry, Yellow = African American Ancestry, Red = Admixed American, Green = East Asian Ancestry, Purple = South Asian Ancestry

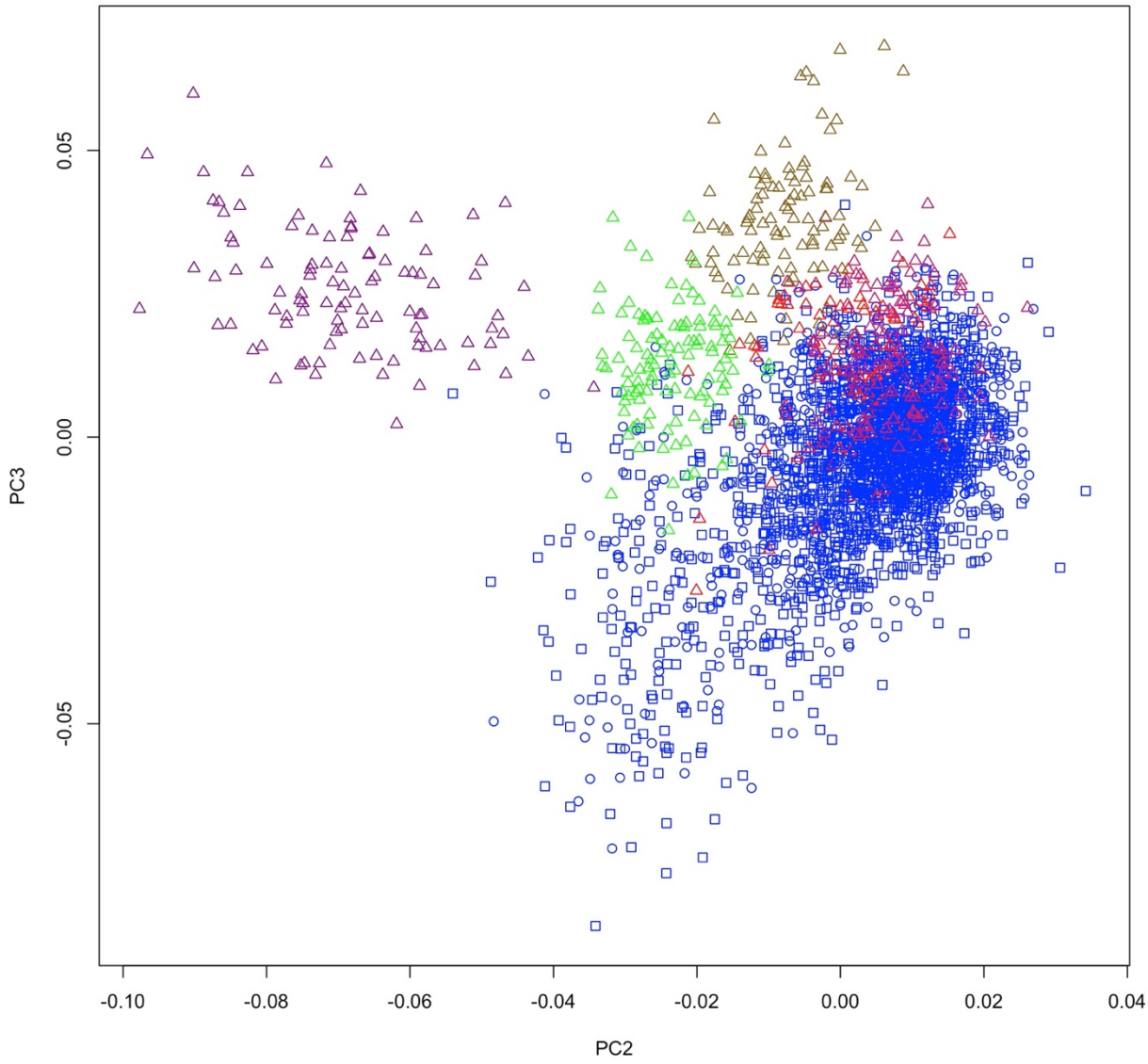


Figure S7. PC2 and PC3 for GENOA, POST and 1000 Genome Data
Legend: Circles = GENOA data, Squares = POST data, Triangles = 1000 Genome data, Blue = European Ancestry, Yellow = African American Ancestry, Red = Admixed American, Green = East Asian Ancestry, Purple = South Asian Ancestry

Polygenic Risk Score Analysis

We used genome-wide summary statistics from the Johnson et al. (2020) GWAS-meta-analysis to construct genome-wide polygenic scores in the GENOA and POST datasets (17). PRSice was used in the construction of the PRS (18,19). Clumping was done within a 250kb physical distance and an LD threshold of $r^2>=0.5$, thus SNPs included in the PRS

were of relatively independent. A series of PRSs were calculated at decreasingly liberal thresholds P-value thresholds (1, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.001, 0.0001, 1^{-05} and $5^{-0.8}$). Results of the PRS can be found in Tables 4 and 5.

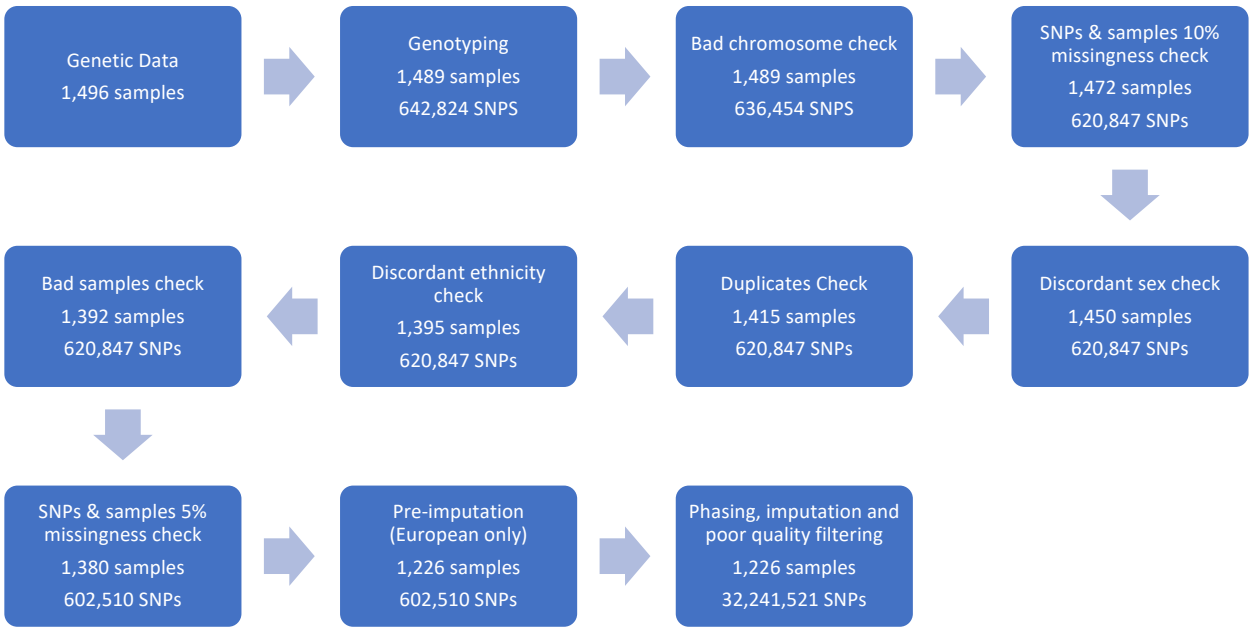


Figure S8. Flow chart of GENOA datasets pre-imputation

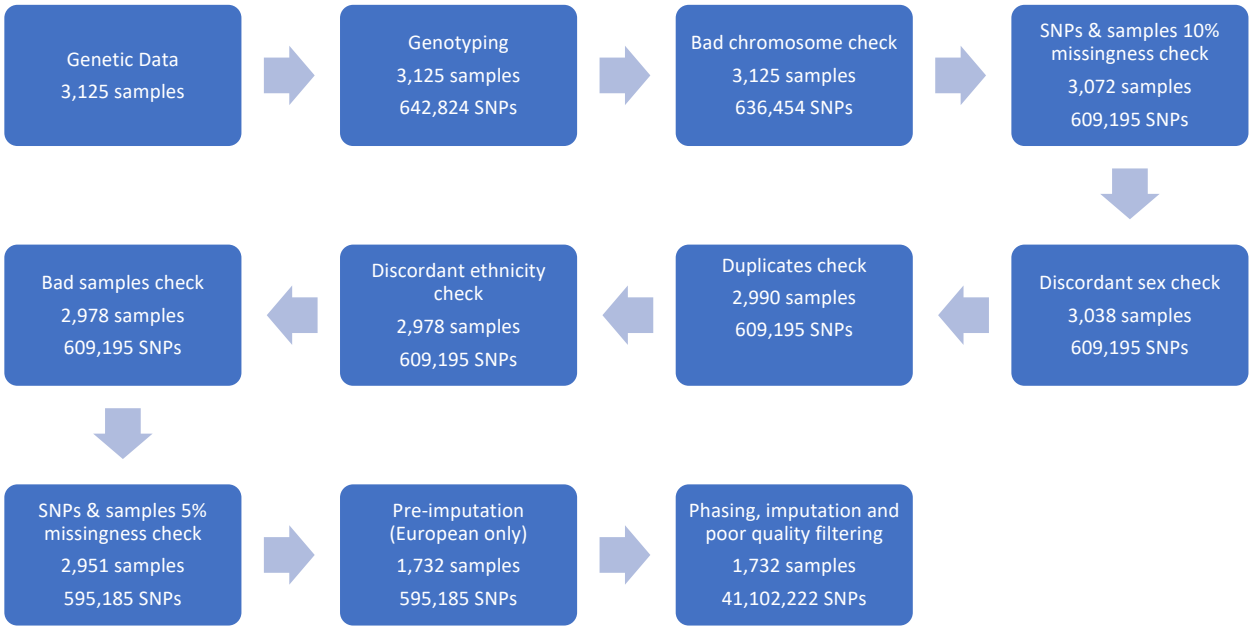


Figure S9. Flow chart of POST data pre-imputation

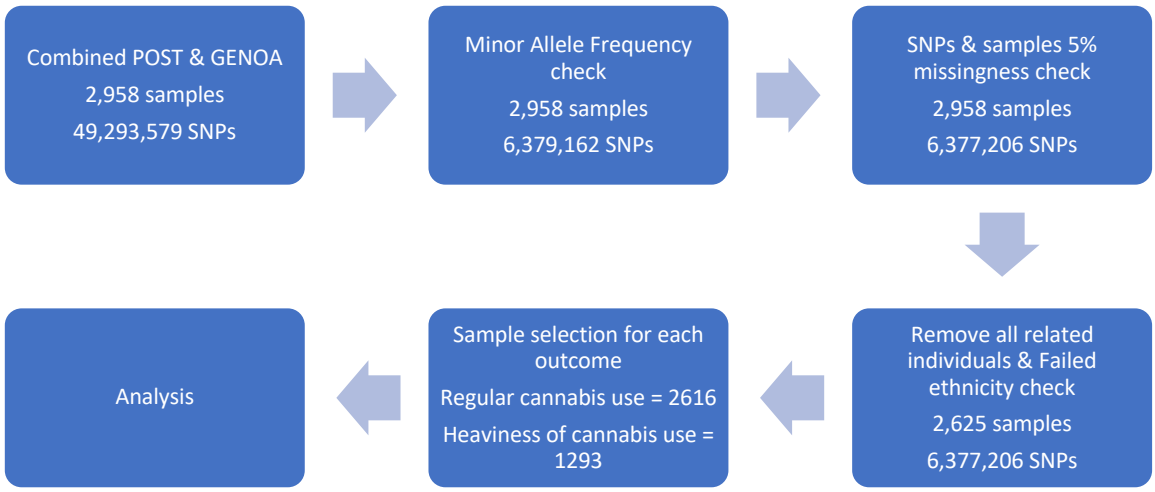


Figure S10. Flow chart of GENOA and POST datasets post-imputation

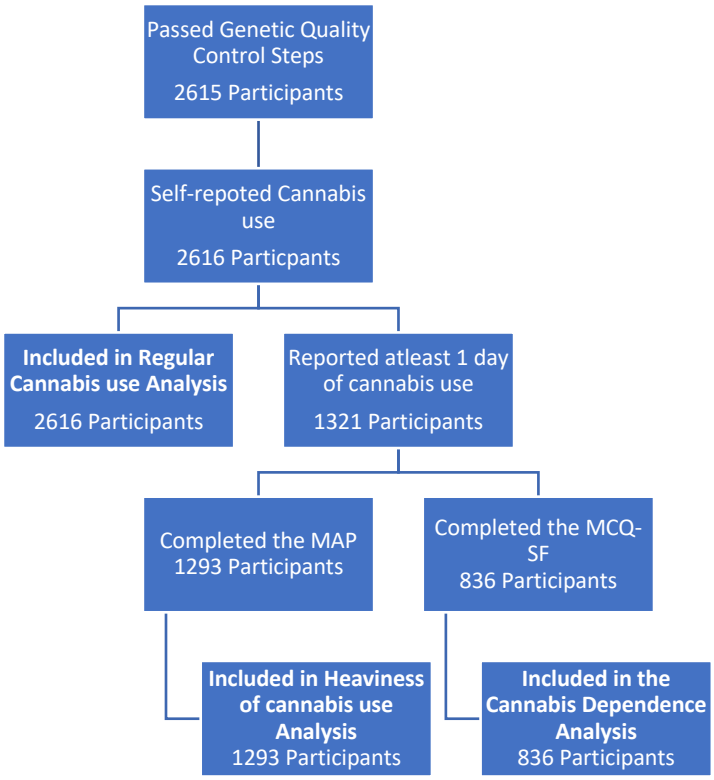


Figure S11. Flow chart of Participant inclusion for each outcome

References

1. Samaan Z, Bawor M, Dennis B, Plater C, Varenbut M, Daiter J, et al. Genetic influence on methadone treatment outcomes in patients undergoing methadone maintenance treatment for opioid addiction: a pilot study. Neuropsychiatr Dis Treat. 2014;1503.
2. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
3. RStudio Team. RStudio: Integrated Development for R. Boston, MA: RStudio; 2020.
4. R Development Core Team. R: A Language and Environment for Statistical Computing. 2011.
5. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284–7.
6. DANgenotek. ORAgene-DISCOVER - OGR-500 [Internet]. 2022. Available from: https://www.dnagenotek.com/us/products/collection-human/oragene-discover/500-series/OGR-500.html
7. Centre d'expertise et de services Génome Québec. DNA Extraction Servics [Internet]. 2020. Available from: https://static1.squarespace.com/static/5a53cd127131a5135e2456ba/t/5f63f155343fba7cda512035/1600385365852/UserGuide_DNAExtraction_en.pdf
8. Centre d'expertise et de services Génome Québec [Internet]. Available from: https://cesgq.com/
9. illumina. Infinium Global Screening Array - 24 v1.0 [Internet]. 2017. Available from: https://grcf.jhmi.edu/wp-content/uploads/2017/12/infinium-commercial-gsa-data-sheet-370-2016-016.pdf
10. illumina. GenomeStudio [Internet]. 2022. Available from: https://support.illumina.com/array/array_software/genomestudio/downloads.html
11. Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature. 2021;590(7845):290–9.
12. Rayner W. McCarthy Tools.
13. Loh P-R, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48(11):1443.
14. TOPMed. TOPMed Imputation Server - Reference Panels [Internet]. Available from: https://topmedimpute.readthedocs.io/en/latest/reference-panels.html
15. Google Cloud. Google Cloud. 2022.
16. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203–9.
17. Johnson EC, Demontis D, Thorgeirsson TE, Walters RK, Polimanti R, Hatoum

AS, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. The Lancet Psychiatry. 2020;7(12):1032–45.
18. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience. 2019;8(7):giz082.
19. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. 2020;15(9):2759–72.

8.8 CHAPTER 5 Supplementary File 1

Reporting checklist for genetic association study.

Based on the STREGA guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STREGAreporting guidelines, and cite them as:

Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, Khoury MJ, Cohen B, Davey-Smith G, Grimshaw J, Scheet P, Gwinn M, Williamson RE, Zou GY, Hutchings K, Johnson CY, Tait V, Wiens M, Golding J, van Duijn C, McLaughlin J, Paterson A, Wells G, Fortier I, Freedman M, Zecevic M, King R, Infante-Rivard C, Stewart A, Birkett N; STrengthening the REporting of Genetic Association Studies. STrengthening the REporting of Genetic Association Studies (STREGA): An Extension of the STROBE Statement.

		Reporting Item	Page Number
Title and abstract			1-4
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	

Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	
Background/rationale			5-6
	#2	Explain the scientific background and rationale for the investigation being reported	
Objectives			6-7
	#3	State specific objectives, including any prespecified hypotheses. State if the study is the first report of a genetic association, a replication effort, or both.	
Study design			7-8
	#4	Present key elements of study design early in the paper	
Setting			7-8
	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	
Eligibility criteria			8
	#6a	Cohort study – Give the eligibility criteria, and the sources and methods of selection of	

		participants. Describe methods of follow-up. Case-control study – Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls. Cross-sectional study – Give the eligibility criteria, and the sources and methods of selection of participants. Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant.	
	#6b	Cohort study – For matched studies, give matching criteria and number of exposed and unexposed. Case-control study – For matched studies, give matching criteria and the number of controls per case.	
Variables			9
	#7a	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	
	#7b	Clearly define genetic exposures (genetic variants) using a widely-used nomenclature system. Identify	

		variables likely to be associated with population stratification (confounding by ethnic origin).	
Data sources/measurement			9-10
	#8a	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	
	#8b	Describe laboratory methods, including source and storage of DNA, genotyping methods and platforms (including the allele calling algorithm used, and its version), error rates and call rates. State the laboratory / centre where genotyping was done. Describe comparability of laboratory methods if there is more than one group. Specify whether genotypes were assigned using all of the data from the study simultaneously or in smaller batches.	
Bias			10-11

	#9a	Describe any efforts to address potential sources of bias	
	#9b	Describe any efforts to address potential sources of bias	
Study size			11
	#10	Explain how the study size was arrived at	
Quantitative variables			9-10
	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why. If applicable, describe how effects of treatment were dealt with.	
Statistical methods			11-12
	#12a	Describe all statistical methods, including those used to control for confounding. State software version used and options (or settings) chosen.	
	#12b	Describe any methods used to examine subgroups and interactions	
	#12c	Explain how missing data were addressed	

	#12d	If applicable, explain how loss to follow-up was addressed
	#12e	Describe any sensitivity analyses
	#12f	State whether Hardy-Weinberg equilibrium was considered and, if so, how.
	#12g	Describe any methods used for inferring genotypes or haplotypes
	#12h	Describe any methods used to assess or address population stratification.
	#12i	Describe any methods used to address multiple comparisons or to control risk of false positive findings.
	#12j	Describe any methods used to address and correct for relatedness among subjects
Participants		12-13
	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if

		applicable. Report numbers of individuals in whom genotyping was attempted and numbers of individuals in whom genotyping was successful.	
	#13b	Give reasons for non-participation at each stage	
	#13c	Consider use of a flow diagram	
Descriptive data			13-15
	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable. Consider giving information by genotype	
	#14b	Indicate number of participants with missing data for each variable of interest	
	#14c	Cohort study – Summarize follow-up time, e.g. average and total amount.	
Outcome data			13-15
	#15	Cohort study Report numbers of outcome events or summary measures over time.Give	

		information separately for exposed and unexposed groups if applicable. Report outcomes (phenotypes) for each genotype category over time Case-control study – Report numbers in each exposure category, or summary measures of exposure.Give information separately for cases and controls . Report numbers in each genotype category. Cross-sectional study – Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable. Report outcomes (phenotypes) for each genotype category	
Main results			15-21
	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	
	#16b	Report category boundaries when continuous variables were categorized	

	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
	#16d	Report results of any adjustments for multiple comparisons	
Other analyses			19-20
	#17a	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	
	#17b	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	
	#17c	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	
Key results			23
	#18	Summarise key results with reference to study objectives	
Limitations			25-26
	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision.	

		Discuss both direction and magnitude of any potential bias.	
Interpretation			23-25
	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	
Generalisability			25-26
	#21	Discuss the generalisability (external validity) of the study results	
Funding			26
	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	