Short communication: Genetic association between schizophrenia and cannabis use

Karin J.H. Verweij a,b,c,*, Abdel Abdellaoui a, Michel G. Nivard a, Alberto Sainz Cort a, Lennie Ligthart a, Harmen H.M. Draisma a,b, Camelia C. Minică a, International Cannabis Consortium 1, Nathan A. Gillespie d, Gonneke Willemsen a, Jouke-Jan Hottenga a, Dorret I. Boomsma a, Jacqueline M. Vink a,c

a Department of Biological Psychology/Netherlands Twin Register, VU University, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands
b Neuroscience Campus Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands
c Behavioural Science Institute, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands
1 International Cannabis Consortium

ABSTRACT

Background and aim: Previous studies have shown a relationship between schizophrenia and cannabis use. As both traits are substantially heritable, a shared genetic liability could explain the association. We use two recently developed genomics methods to investigate the genetic overlap between schizophrenia and cannabis use.

Methods: Firstly, polygenic risk scores for schizophrenia were created based on summary statistics from the largest schizophrenia genome-wide association (GWA) meta-analysis to date. We analysed the association between these schizophrenia polygenic scores and multiple cannabis use phenotypes (lifetime use, regular use, age at initiation, and quantity and frequency of use) in a sample of 6,931 individuals. Secondly, we applied LD-score regression to the GWA summary statistics of schizophrenia and lifetime cannabis use to calculate the genome-wide genetic correlation.

Results: Polygenic risk scores for schizophrenia were significantly (p < 0.05) associated with five of the eight cannabis use phenotypes, including lifetime use, regular use, and quantity of use, with risk scores explaining up to 0.5% of the variance. Associations were not significant for age at initiation of use and two measures of frequency of use analyzed in lifetime users only, potentially because of reduced power due to a smaller sample size. The LD-score regression revealed a significant genetic correlation of rG = 0.22 (SE = 0.07, p = 0.003) between schizophrenia and lifetime cannabis use.

Conclusions: Common genetic variants underlying schizophrenia and lifetime cannabis use are partly overlapping. Individuals with a stronger genetic predisposition to schizophrenia are more likely to initiate cannabis use, use cannabis more regularly, and consume more cannabis over their lifetime.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Numerous studies have observed an association between schizophrenia and cannabis use (McGrath et al., 2010; Moore et al., 2007; Morrison et al., 2009; van Os et al., 2002), although the direction of causation is still under debate. As both traits are substantially heritable, with heritability estimates of around 45% for lifetime cannabis use (Verweij et al., 2010) and 80% for schizophrenia (Sullivan et al., 2003), a shared genetic liability could explain the relationship. With methodological advances in molecular genetics and increased sample sizes in genome-wide association (GWA) studies it has become viable to use measured genetic variation among individuals to examine this relationship. Power et al. (2014) used results from the 2013 schizophrenia GWA study that included 13,833 schizophrenia cases and 18,310 controls (Ripke et al., 2013) to predict cannabis use in a target sample of 2082 Australian individuals. They found that the genetic variants underlying schizophrenia significantly predicted lifetime cannabis use,
and quantity of use within users (maximum variance explained of $R^2 = 0.47\%$ and $0.85\%$, respectively), but not age at initiation of cannabis use.

Here, we used two recently developed methods to further investigate the genetic correlation between both traits. Firstly, we performed a polygenic risk analysis to determine the extent to which common genetic variants that affect the risk of schizophrenia predict various measures of cannabis use. Based on the results from the latest and largest schizophrenia GWAS study (with up to 36,989 schizophrenia cases and 113,075 controls; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) we generated genetic risk scores for schizophrenia in a large independent target sample from the Netherlands. We used LDpred to create the risk score, a method shown to have greater prediction accuracy than the conventional risk prediction approach involving linkage disequilibrium (LD) pruning followed by P-value thresholding (Vijhalmsson et al., 2015). We then determined the association of the genetic risk score for schizophrenia with various cannabis use phenotypes, including lifetime (ever versus never) use, regular use, age at initiation of use, and quantity and frequency of use.

Secondly, we applied LD-score regression to the summary statistics of the largest GWA meta-analyses of schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) and lifetime cannabis use (Stringer et al., 2016) to estimate the genome-wide genetic correlation. To our knowledge, this is the first study to determine the strength of the genetic correlation between schizophrenia and cannabis use; the information can provide useful etiological insights into the relationships between the two traits.

2. Methods

2.1. Polygenic risk prediction analysis

The target sample comprised 6,931 participants (64.1% females) from 1,244 twin families registered at the Netherlands Twin Registry (Boomsma et al., 2006). Participants were between 18 and 94 years old ($M = 43.0, SD = 15.7$). Data on cannabis use were obtained from five waves of self-report questionnaires sent out between 1993 and 2013 (see Willemsen et al., 2013). For subjects who participated in more than one wave we used data from their last questionnaire.

We analysed various cannabis use variables:

1) Lifetime cannabis use was analysed as a dichotomous variable with never users marked as 0 and ever users as 1.
2) Lifetime regular cannabis use (2a) was analysed as a dichotomous variable. Individuals were asked whether they had ever used cannabis regularly during their lifetime. Those who had used regularly were marked as 1 and those that had never used regularly as 0. We also repeated this analysis excluding experimental users, i.e., those that had used regularly were marked as 1 and those that had never used as 0 (2b).
3) Age at onset of cannabis use in years.
4) Lifetime frequency of use was analysed as a continuous variable (4a). Participants were asked about their lifetime frequency of cannabis use. Response categories were: 0 = never; 1 = 1-to-2 times; 2 = 3-to-5 times; 3 = 6-to-10 times; 4 = 11-to-19 times; 6 = 40 times or more. We also repeated this analysis in lifetime users only (4b).
5) Frequency of cannabis use was analysed as a continuous variable (5a). Participants were asked about their frequency of cannabis use during the period of heaviest use. Response categories were: 0 = never used; 1 = once a month or less often; 2 = 2-to-4 times a month; 3 = 2-to-3 times per week; 4 = 4-to-5 times a week; 6 = six days a week or daily. We also repeated this analysis in lifetime users only (5b).

Details of the variables, descriptive statistics and sample sizes are provided in Table 1.

Polygenic risk scores were created using LDpred (Vijhalmsson et al., 2015), which takes into account LD among the SNPs in creating the polygenic risk scores. Risk scores in the target sample were generated by calculating the additive effect size of each marker using the SNP effect sizes from the latest schizophrenia GWA meta-analysis results (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) and the LD structure from the European populations in the 1000 Genomes reference set. We considered all schizophrenia polygenic scores with the expected fraction of causal genetic variants (the fraction of markers with non-zero effects) set at 30%, because in a previous paper the risk score based on this fraction was shown to optimally predict schizophrenia (see Vijhalmsson et al., 2015).

Generalized estimating equation (GEE) modelling was applied to test whether the polygenic risk scores for schizophrenia predicted cannabis use phenotypes in the target cohort. A covariance matrix was used to account for family relatedness and tests were based on robust (sandwich-corrected) standard error (see Minica et al., 2014). Age, sex, birth cohort (three dummy variables representing four 20-year periods), and ten genetic PCs (three ancestry-informative PCs and seven PCs accounting for genotyping batch effects) were included as covariates in the model. Analyses also accounted for two environmental factors (level of urbanisation and socio-economic status) that have previously been found to be associated with schizophrenia and cannabis use (Daniel et al., 2009; Kuepper et al., 2011; Martino et al., 2008; Vassos et al., 2012; Werner et al., 2007). Based on postal code, we determined the level of urbanisation of participants’ residence (ranging from 1 (very high, >2500 addresses per km²) to 5 (very low, <500 addresses per km²) and socio-economic status as measured by the average income of their residential area.

Variance explained by the polygenic risk scores was calculated in regression analyses as the R² for dichotomous variables in the model including polygenic risk scores and covariates minus the R² of the model including only covariates.

2.2. LD-score regression analysis

With LD-score regression (Bulik-Sullivan et al., 2015a, 2015b) the variance in a trait that can be explained by all SNPs based on GWA summary statistics is estimated. The method is based on the fact that an estimated SNP effect-size incorporates effects of all SNPs in LD with that SNP. For polygenic traits, SNPs with high LD have on average higher $\chi^2$ statistics than SNPs with low LD (the more genetic variation a SNP taps, the higher the probability that the SNP will tag a causal variant). When regressing the $\chi^2$ statistics from a GWA study against the LD score for each SNP, the slope of the resulting regression line provides an estimate of the proportion of trait variance accounted for by all genotyped SNPs (Bulik-Sullivan et al., 2015b). Cross-trait LD-score regression is an extension to estimate the genetic covariance between traits using GWA summary statistics of multiple traits (Bulik-Sullivan et al., 2015a). The genetic covariance is estimated using the slope from the regression of the product of $z$-scores from two GWA studies on the LD score. The estimate obtained from this method represents the genetic correlation between the two traits based on all polygenic effects captured by SNPs.

Here, we estimated the genetic correlation between schizophrenia and lifetime cannabis use based on the summary statistics from the largest GWA meta-analyses for schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) and cannabis use (International Cannabis Consortium, Stringer et al., 2016) to date. From the cannabis use GWA study we included only SNPs present in all contributing cohorts were included. Furthermore, we only included HapMap-3 SNPs (as suggested by Bulik-Sullivan et al., 2015a), resulting in 1,125,108 and 433,918 SNPs for schizophrenia and cannabis use, respectively. Standard LD scores were used as provided by Bulik-Sullivan and colleagues based on the 1000 genomes reference set, restricted to European populations.

3. Results

Individuals from more urbanised residential areas were more likely to have used cannabis (Nagelkerke $R^2 = 3.3\%$, $p < 0.001$), whereas socio-economic status was not associated with lifetime cannabis use ($p = 0.35$). Polygenic risk scores for schizophrenia were significantly ($\alpha < 0.05$) associated with five of the eight tested cannabis use phenotypes in the target sample, including lifetime use, regular use, and quantity of use, with risk scores explaining up to 0.5% of the variance (see Table 1). This indicates that individuals with a genetic predisposition to schizophrenia were significantly more likely to use cannabis. Associations were not significant for Age at initiation of use, Lifetime frequency of cannabis use in users, and Frequency of use during period of most use in users, potentially because of reduced power due to a smaller sample size for these variables.

The LD-score regression analysis revealed a significant genetic correlation of $r_g = 0.22$ ($SE = 0.07$, $p = 0.003$) between schizophrenia
Table 1
Genetic risk prediction analysis. Cannabis variables in the target sample, sample size, descriptive statistics, and results of the polygenic risk score analyses for schizophrenia predicting the various cannabis use phenotypes.

<table>
<thead>
<tr>
<th>Dichotomous variables</th>
<th>N</th>
<th>Prevalence</th>
<th>Mean (SD)</th>
<th>B</th>
<th>R²</th>
<th>P-value</th>
<th>P_FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6931</td>
<td>24.8%</td>
<td>–</td>
<td>0.101</td>
<td>0.24%</td>
<td>0.002</td>
<td>0.020</td>
</tr>
<tr>
<td>2a</td>
<td>6603</td>
<td>3.8%</td>
<td>–</td>
<td>0.189</td>
<td>0.43%</td>
<td>0.009</td>
<td>0.025</td>
</tr>
<tr>
<td>2b</td>
<td>5457</td>
<td>4.5%</td>
<td>–</td>
<td>0.204</td>
<td>0.49%</td>
<td>0.007</td>
<td>0.025</td>
</tr>
</tbody>
</table>

N = number of individuals in the target sample, B = effect size, R² = variance explained, for dichotomous traits R² = Nagelkerke’s pseudo R², P_FDR = P-value corrected for multiple testing by false discovery rate correction (FDR); note that this correction is relatively conservative given that most outcome variables are strongly interrelated.

and lifetime cannabis use, implying that part of the common genetic variants overlap between schizophrenia and lifetime cannabis use.

4. Discussion

Overall, results from our two analyses support the hypothesis that the association between schizophrenia and cannabis use is partly due to a shared genetic etiology. Polygenic risk scores based on the SNP effect sizes from a schizophrenia GWA meta-analysis significantly predicted most of the cannabis use variables in our target sample. Individuals with an increased genetic predisposition to schizophrenia were more likely to have initiated cannabis use, to have used cannabis regularly, and to have used more cannabis over their lifetime. These results are in line with the study by Power et al. (2014) who found an association between individuals’ burden of schizophrenia risk alleles and lifetime cannabis use and quantity of use, but not age at initiation. The overall variance explained by the combined SNPs was low (up to 0.5%), but consistent with many previous cross-disorder predictions. This is not surprising because summing estimates of SNP effects also sums the error component of those estimates, and thus does not yield an accurate and unbiased estimate of the variance explained by the aggregate of all SNP effects. This method relies heavily on the accuracy of the SNP effect estimates in the discovery sample, which will increase when larger samples sizes become available. More accurate predictions may also be obtained by better phenotyping; our measures of cannabis use were limited by the potential for inaccuracy and bias in retrospective self-reports. Additionally, the genetic predictions may be stronger for more extreme forms of cannabis use (e.g., problematic use).

With the LD-score regression analysis we quantified for the first time the genetic correlation between schizophrenia and cannabis use. The genetic correlation was estimated to be rather high, r² = 0.22, suggesting that a substantial part of the well-known association between schizophrenia and cannabis use can be explained by shared genetic influences (although we are unable to differentiate between pleiotropy and causality). The estimated genetic correlation is expectedly weaker than the genetic correlation between schizophrenia and bipolar disorder or schizophrenia and depression, but is similar in magnitude to the genetic correlation of schizophrenia with ADHD and anorexia, and stronger than the genetic correlation of schizophrenia with all other diseases and traits investigated by Bulik-Sullivan et al. (2015a). Bioinformatics work and future studies with advanced technologies, novel statistical approaches (such as Mendelian randomisation), and larger sample sizes, should aim to determine the nature of the genetic association between schizophrenia and cannabis use and identify common genes and biological mechanisms that can explain the genetic association.

Contributors

KJHV was responsible for the study concept and the design of the study. GW, DIB and JMV contributed to the data acquisition. KJHV performed the data analyses with assistance from AA, MGN, ASC, LL and JH. KJHV drafted the manuscript. AA, MGN, ASC, LL, HHMD, CCM, NAG, GW, JH, DIB and JMV provided critical revision of the manuscript for important intellectual content. All authors approved the final version for publication. No conflict of interest.

Role of funding source

Nothing declare.

Appendix A.

International Cannabis Consortium.

Affiliations
1. Department of Complex Trait Genetics, VU Amsterdam, Center for Neurogenomics and Cognitive Research, Amsterdam, The Netherlands.
2. Department of Psychiatry, Academic Medical Centre, Amsterdam, The Netherlands.
3. Department of Biological Psychology/Netherlands Twin Register, VU University, Amsterdam, The Netherlands.
5. VU University, Department of Developmental Psychology and EMGO Institute for Health and Care Research, Amsterdam, The Netherlands.
6. The Hospital for Sick Children Research Institute, Toronto, Canada.
7. Department of Psychology, University of Illinois Urbana-Champaign, Champaign, Illinois, USA.
8. Department of Human Neurogenetics, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands.
9. Department of Psychology, University of Minnesota, Minneapolis, Minnesota, USA.
10. Department of Public Health, University of Helsinki, Hjelt Institute, Helsinki, Finland.
11. Estonian Genome Center, University of Tartu, Tartu, Estonia.
12. Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.
13. Psychiatric Genetics Unit, Vall d’Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain.
14. Department of Psychiatry, Hospital Universitari Vall d’Hebron, Barcelona, Spain.
15. Biomedical Research Network Centre on Mental Health (CIBERSAM), Barcelona, Spain.
16. Twin Research and Genetic Epidemiology, King’s College London, London, United Kingdom.
17. Biomedical Genetics Department, Boston University School of Medicine, Boston, Massachusetts, USA.
18. Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA.
19. Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.
20. Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA.
21. School of Social and Community Medicine, University of Bristol, Bristol, UK.
22. MRC Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, UK.
23. Department of Psychiatry, Virginia Institute for Psychiatric and Behavior Genetics, Virginia Commonwealth University, Richmond, Virginia, USA.
24. Research Centre Adolescent Development, Utrecht University, Utrecht, the Netherlands.
25. Department of Psychology and Psychiatry, University of California San Diego, La Jolla, California, USA.
27. Department of Psychiatry and Legal Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain.
28. Genetic Epidemiology, Molecular Epidemiology and Neurogenetics Laboratories, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.
29. Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.
30. Department of Psychiatry, Washington University School of Medicine, St Louis, Missouri, USA.
31. Brain & Mind Research Institute, University of Sydney, Sydney, NSW, Australia.
32. Department of Psychiatry, University of Colorado Denver, Aurora, Colorado, USA.
33. University of Eastern Finland, Institute of Public Health & Clinical Nutrition, Kuopio, Finland.
34. Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, Helsinki, Finland.
35. Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, USA.
36. Department of Molecular, Cellular and Developmental Biology, University of Colorado Boulder, Boulder, Colorado, USA.
37. Department of Psychology, University of Notre Dame, Notre Dame, Indiana, USA.
39. Interdisciplinary Center for Pathology and Emotion Regulation, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.
40. Physiology and Nutritional Sciences, University of Toronto, Toronto, Canada.
41. Department of Psychological & Brain Sciences, Indiana University Bloomington, Bloomington, Indiana, USA.
42. Institute for Behavioral Genetics, Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, Colorado, USA.
43. Department of Psychiatry, University of California San Diego, La Jolla, California, USA.
44. Rotman Research Institute, Baycrest, Toronto, Canada.
45. Psychology and Psychiatry, University of Toronto, Toronto, Canada.
46. Center for the Developing Brain, Child Mind Institute, New York, USA.
47. Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland.
48. UK Centre for Tobacco and Alcohol Studies and School of Experimental Psychology, University of Bristol, Bristol, UK.
49. Psychiatry, Genetics, & Neurobiology, Yale University School of Medicine & VA CT, West Haven, Connecticut, USA.
50. Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands.

References

