Pathways to cannabis abuse: a multi-stage model from cannabis availability, cannabis initiation and progression to abuse

Nathan A. Gillespie¹, Michael C. Neale¹ & Kenneth S. Kendler¹,²
Departments of Psychiatry¹ and Human Genetics,² Virginia Commonwealth University, Richmond, VA, USA

ABSTRACT

Aims Although previous twin studies have modeled the association between drug initiation and abuse, none has included the obvious risk factor of drug availability. Our aim is to determine whether the genetic and environmental risk factors for cannabis availability also generate variation in cannabis initiation and/or progression to DSM-IV symptoms of abuse. Design We used multi-stage modeling, also known as causal-common-contingent (CCC) analysis, to partition the genetic and environmental factors into common and stage-specific components. Participants This report is based on data collected from 1772 adult males from the Mid Atlantic Twin Registry. Measurements The twins participated in two structured interviews which included clinical and non-clinical measures of cannabis abuse as well as retrospective assessments of perceived cannabis availability between ages 8 and 25 years. Findings Cannabis availability explained almost all the shared environmental risks in cannabis initiation and abuse. The influence of availability on the symptoms of abuse was indirect and mediated entirely by cannabis initiation. Conclusion These findings have begun to elucidate the causal processes underlying the liability to drug use and abuse in terms of putative risk factors. Specifically, our results show that the latent shared environmental factors in cannabis initiation and abuse can be explained by measured aspects of the shared environment—those responsible for variation in cannabis availability.

Keywords Abuse, availability, cannabis, initiation, progression, transition.

INTRODUCTION

We know from studies of smoking initiation and nicotine dependence that there is significant genetic and environmental covariance between initiation and dependence [1]. A number of twin studies have revealed that genetic and environmental effects can explain significant proportions of variance in initiation, abuse and dependence for a variety of licit and illicit substances, including cannabis [1–17]. Moreover, progression from initial drug involvement to regular use and abuse appear to be influenced by the same of the same genetic and environmental factors [3,15,17].

Rates of illicit drug use are higher among individuals with substance-using peers and those who live in areas where drugs are obtained easily [18–27]. However, genetically informative studies to date have not examined the association between such environmental risk factors and drug initiation. Contrary to the belief that it is an entirely ‘environmental’ risk factor, variation in cannabis availability can be explained by genetic and environmental factors. Recently, Gillespie and colleagues [28] found that between the ages of 8 and 25 years there is an increase in the additive genetic and decline in the shared environmental effects in self-report measures of cannabis availability. To what extent these same genetic and environmental influences can account for the liability to initiate cannabis and then progress to abuse is unknown.

Any association between availability, initiation and abuse should be modeled as a multi-stage process where availability or exposure (upstream variable) is followed by initiation which, under certain circumstances, can result...
in abuse or dependence (downstream variable). Multistage modeling or causal-common-contingent (CCC) pathway analyses [17] represent an improvement over the standard independent assessments of drug use and abuse because of the ability to partition genetic and environmental factors into (i) those that are common to upstream and downstream variables and (ii) those specific to each stage.

**Aim**

Coffey & Lynskey [22] have shown that the transition from cannabis initiation to harmful use is more likely in males for whom drug availability is a strong determinant. Although several papers have used CCC pathway analyses to model the association between drug initiation and drug abuse or dependence [3,15,17], none has included the obviously critical upstream risk factor of drug availability, nor decomposed the association into genetic and environmental components of variance. Therefore, in order to determine the importance of this upstream risk factor, this study will model the association between cannabis availability, initiation and measures of abuse as a three-stage process.

**METHODS**

**Subjects and measures**

As part of an ongoing study of adult male twins from the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD), this report is based on data collected from a second and third wave of interviews conducted between 1994 and 2004. The VATSPSUD is described in detail elsewhere [29]. Briefly, twins were eligible for participation in this study if one or both twins were matched successfully to birth records, were a member of a multiple birth with at least one male, were Caucasian and were born between 1940 and 1974. Of 9417 eligible individuals for the first wave (1993–96), 6814 (72.4%) completed the initial interviews. At least 1 year later, we contacted those who had completed the initial interview to schedule a second interview.

**Second interview—cannabis initiation and abuse**

The second interview (1994–98) was completed by 5602 individuals, or 82.2% of those who had completed the first interview. This interview included assessments of life-time cannabis drug use, abuse and dependence using an adaptation of the Structured Clinical Interview (SCID) for DSM-III-R [30] based on six abuse and seven dependence items.

Subjects who endorsed having ‘ever tried’ cannabis [marijuana, hashish or 9-tetrahydrocannabinol (THC)] were then asked two questions: ‘How old were you the first time you took cannabis?’ and ‘How old were you when you used cannabis the most?’ A cannabis initiation score (yes/no) was then calculated and scored positive if subjects responded to both items.

Subjects who endorsed using cannabis ‘six or more times in a life-time’ were given the six abuse items. With the percentage of endorsements shown in brackets, these criteria assessed whether or not subjects’ use of cannabis (i) occurred during dangerous situations (36%), (ii) caused legal or resulted in traffic problems (4%), (iii) caused problems with family, friends or colleagues (14%), (iv) caused physical or psychological problems (13%), (v) occurred while doing something important (21%) and (vi) caused subjects to miss school, work, or appointments (8%). Dependence items were not included in our analyses because not all subjects met the more stringent criteria of ‘11 or more times in a month’ before dependence criteria could be administered.

**Third interview—cannabis availability**

The third interview wave, or ‘MM3’ (1998–2004), was completed solely by members of male–male twin pairs. Individuals were eligible for the MM3 if they came from a male–male pair, and if both had been interviewed in wave 2. Excluding triplets, interviews were completed by 1772 twins, aged 24–62 years [μ = 40.3, standard deviation (SD) = 9.0], representing a response rate of 74.7%. This included 459 monozygotic (MZ) and 282 dizygotic (DZ) twin pairs and 290 incomplete twin pairs or singletons (MZ = 154, DZ = 136). This interview included an assessment of perceived drug availability for cannabis for five age periods: 8–11, 12–14, 15–17, 18–21 and 22–25 years. We used a life-history calendar method to increase the validity of our retrospectively collected data [31–33]. This method has shown that although human memory is relatively poor at recall, it can be improved significantly when probed with careful questioning involving specific time-periods and events. For cannabis, subjects were asked, ‘When you were [AGE] how easy would it have been to get CANNABIS if you wanted to use it?’. Responses were scored on a four-point ordinal scale (0 = ‘very difficult or don’t know’, 1 = ‘somewhat difficult’, 2 = ‘somewhat easy’ and 3 = ‘very easy’). We combined ‘very difficult’ and ‘don’t know’ because during the pilot phase, interviewers noted consistently that ‘don’t know’ responses typically meant not knowing how to obtain a drug rather than not knowing anything about the drug. It is true that ‘don’t know’ responses were very rare. Complete cannabis availability data were available from 1788 males. Cronbach’s alpha for the five measures of cannabis availability was 0.80. Retest correlations, based on a
sample of 141 subjects for whom the mean interval was 29 days, averaged 0.68 across five age periods [28].

Among the total of 1772 males with complete cannabis availability, 54% (n = 956) reported having tried cannabis at least once beginning on average at 18.4 years (SD = 4.3) (see Fig. 1a). The mean onset age for any symptom of abuse or dependence was 18.9 years (SD = 4.2) (see Fig. 1b).

Zygosity and interview protocol

Zygosity was diagnosed using a combination of self-report measures, photographs and DNA analysis (see [6]). In both interviews, most subjects (~90%) were interviewed by telephone. A small number were interviewed in person because of subject preference, residence in an institutional setting (usually jail) or not having a telephone. The project was approved by the Virginia Commonwealth University Institutional Review Board. Subjects were informed about the goals of the study and provided informed consent before interviews. Interviewers had a Master’s degree in a mental health-related field or a Bachelor’s degree in this area plus 2 years of clinical experience. The two members of a twin pair were each interviewed by different interviewers who were blind to interview information about the co-twin.

Data preparation

Data were prepared for raw ordinal data analysis in the software Mx [34]. This approach assumes that the observed ordinal categories within each item are an imprecise measure of a latent normal distribution of liability, and that this liability distribution has one or more threshold values which discriminate between the categories. Thresholds can be conceived of as cut-points along a standard normal distribution which classify individuals in terms of a probability or risk of endorsing one of two or more discrete (ordinal) categories. Cannabis availability was summed across the five epochs and recoded onto a five-point ordinal scale ranging from 0 to 4. A total cannabis abuse score based on the six diagnostic criteria was also calculated. This was recoded onto a four-ordinal scale ranging from 0 to 3. All thresholds were adjusted for the linear effects of age.

Statistical analyses

In studies of monozygotic (MZ) and dizygotic (DZ) twins reared together, liability can be divided into additive genetic (A), common environment (C) and random environment (E) variance components. With multivariate data it is possible to decompose the association between availability, initiation and abuse into these same components. Based on the multi-stage methods for CCC pathway analyses [17,35] we fitted four models to test specific hypotheses about the nature of the association between the three measures. The key issue is whether any of the genetic and environmental variance in drug availability is associated with variation in drug initiation and subsequent liability to abuse.

Model 1

In Fig. 2, model 1 specifies that the genetic and environmental factors for drug availability, i.e. A1, C1 and E1, generate variation directly in both drug initiation (via \(a_{2,1}, c_{2,1}\) and \(e_{2,1}\)) and for drug abuse (via the \(a_{3,1}, c_{3,1}\) and \(e_{3,1}\)). The model also predicts that A1, C1 and E1 can act indirectly on drug abuse via drug initiation along the \(\beta_2\) pathway. The genetic and environmental factors for initiation (A2, C2 and E2) may also generate variation in drug abuse indirectly via \(\beta_2\).

Because initiation is not contingent upon reported availability, direct pathways from A1, C1 and E1 can be
estimated. The reason that there are no direct pathways from A2, C2 and E2 to abuse in this model is because initiation is strictly binary. Therefore, there is no information on the cross-trait association between measures within individuals, which would allow a full decomposition of the association between initiation and abuse into genetic, common and specific environmental effects [35].

The contingent nature of abuse on initiation means that individuals who do not initiate cannot, by definition, develop abuse, and so diagnostic data will be missing. However, if we restrict the genetic and environmental covariance between initiation and the outcome measures into a common pathway it is possible to estimate the regression of outcome on initiation.

Model 2

In model 2, the genetic and environmental pathways a3,1, c3,1 and e3,1 in Fig. 2 are removed. In this model, none of the association between availability and abuse is due to latent, correlated liabilities stemming directly from A1, C1 and E1 via a3,1, c3,1 and e3,1 to the abuse phenotype. Rather, any association between availability and the diagnostic measures is mediated entirely via initiation through path \( \beta_2 \) at the phenotypic level.

Models 3 and 4

In model 3 (Fig. 3) none of the associations between availability, initiation and abuse are due to correlated latent, genetic and environment effects. Instead, the association occurs entirely at the phenotypic level either directly via \( \beta_3 \), or indirectly via \( \beta_1 \) through initiation and along the \( \beta_2 \) pathway. \( \beta_1 \) represents any direct influence of availability on abuse that does not influence initiation. Model 4 predicts that \( \beta_3 \) is non-significant, i.e. \( \beta_3 = 0 \).

We fitted all four models adjusted for age at interview to raw ordinal data using Mx (version 1.66b) [34] using the total number of abuse symptoms as our outcome measure. Model 1, which has the most parameters, was compared to the nested models 2, 3 and 4 using a likelihood ratio \( \chi^2 \) statistic in which the degrees of freedom equal the difference between the degrees of freedom between the full and nested submodels. Sample size-adjusted Bayesian information criterion (BIC) was also used as an index of fit and parsimony [36] to judge the best-fitting model.

The best-fitting model will retain all parameters. This decision is based on Sullivan & Eaves’ [37] finding that, in analyses based on discrete traits, estimates from the full ACE model will be more accurate and that attempts at
Table 1 Multivariate analysis of cannabis availability, cannabis initiation and the number of DSM-IV symptoms of abuse.

<table>
<thead>
<tr>
<th>Symptoms of DSM-IV abuse</th>
<th>Model</th>
<th>-2LL</th>
<th>df</th>
<th>D-2LL</th>
<th>Δdf</th>
<th>P</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability</td>
<td>1</td>
<td>7384.69</td>
<td>4049</td>
<td></td>
<td></td>
<td></td>
<td>-3922.20</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7385.77</td>
<td>4052</td>
<td>1.08</td>
<td>3</td>
<td>0.78</td>
<td>-3927.31</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>7464.12</td>
<td>4055</td>
<td>79.43</td>
<td>6</td>
<td>0.00</td>
<td>-3893.78</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>7474.64</td>
<td>4055</td>
<td>89.95</td>
<td>6</td>
<td>0.00</td>
<td>-3888.51</td>
</tr>
</tbody>
</table>

Model 1: see Fig. 2. Model 2: see Fig. 2 in which $a_1$, $c_1$, and $e_1$ = 0. Model 3: see Fig. 1. Model 4: see Fig. 3 in which $β_1 = 0$. Δ-2LL = change in $-2 \times \log$-likelihood (asymptotically distributed as a $χ^2$) with degrees of freedom equal to the difference in degrees of freedom between saturated (model 1) and nested submodels below. BIC: sample size-adjusted Bayesian information criterion (larger negative values imply better fit).

parsimony result in oversimplification of the models rather than a simpler and more accurate representation of the data. This will probably occur in cases such as ours, which involve more complex multivariate modeling, and where the sample is not large enough to make definitive conclusions. Removing all parameters with lower bounds spanning zero, including parameters with small point estimates, i.e. $<0.10$, assumes that the component of variance is known to be zero without any error variance, and if this argument is incorrect then future research might ignore an important source of variance [37].

RESULTS

The correlations [and 95% confidence intervals (CI)] based on the empirical model between (i) cannabis availability and initiation and (ii) between cannabis availability and symptoms of abuse were 0.48 (95% CI: 0.42–0.53) and 0.23 (95% CI: 0.12–0.33), respectively.

As summarized in Table 1, model 2 provided the best fit to the data to the symptom and diagnostic data. This is shown in Fig. 4, which includes the standardized path coefficients and 95% CI. As summarized in Table 2, most of the genetic (97%) and non-shared environmental variance (97%) in initiation is unique to this measure, whereas almost all its shared environment variance (96%) is explained by cannabis availability. Variation in the abuse phenotype comes from several sources. Approximately two-thirds of the genetic variance (62%) comes from initiation, while just over one-third (36%) is unique to it. Almost all the shared environmental variation (92%) in the symptoms can be explained by shared environmental effects for cannabis availability. Most (82%) of the non-shared environmental variance in abuse symptoms is associated with cannabis initiation, whereas only a small proportion (16%) is unique.

DISCUSSION

This is, to our knowledge, the first study to model the genetic and environmental association between cannabis availability, initiation and measures of abuse. In addition to unique genetic and non-shared environmental effects, our modeling suggests that much of the variation in abuse can be explained by a combination of (i) shared environmental risks in cannabis availability which act indirectly via initiation; and (ii) the genetic and non-shared environmental risks for cannabis initiation.

Our genetic and environmental estimates for initiation and abuse are consistent with previous findings [6,9,12,16,38–41]. Our results are also consistent with results based on overlapping twin samples from the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders [3,42]. Portions of the genetic and environmental variation in abuse can be explained by genetic and environmental risks to initiating cannabis. Our modeling is an extension of the two-stage CCC design used by Maes et al. (2004), Neale et al. (2006) and Agrawal and colleagues [3]. It illustrates that the familial risks in initiation and abuse can be explained directly and indirectly by the risk factors for availability.

Our results lend support to the concept of availability proneness, which predicts that the liability to initiate and...
progress to abuse are a function of both hereditary and environmental influences [43]. Moreover, portions of these risks can be explained by availability which influences the risk of initiation as well as progressing to symptoms of abuse. These results raise important considerations such as whether or not reducing availability can lead to a reduction in initiation and subsequent abuse. Indeed, this is the basis for the ‘War on Drugs’.

Rates of cannabis initiation, use and abuse are not necessarily higher in communities which have decriminalized, in part or in full, the possession or use of illicit substances [21,44,45], nor are rates lower in countries such as the United States, which have severe penalties for possession and use [46]. Cannabis use in the Netherlands, where possession of small amounts is legal, is lower than in neighbouring countries with stricter laws. In 2000, the life-time prevalence among young adults was 45% in Denmark, 40% in France, 31% in Germany and 43% in the United Kingdom compared to 32% in the Netherlands [47]. Similar to the US general population [29,48] life-time cannabis use among Virginian twins is 50–56% in males [49] and 46–53% among females [38].

It is unknown if the above prevalence differences are confounded by legal, cultural and economic differences which may impact differentially on the availability of licit and illicit substances. These same differences may also alter the observed association between availability, initiation and abuse. For instance, if the variance in measures of cannabis availability is likely to be more restricted in places such as the Netherlands, then other aspects of the shared environment or genetic risks may prove to be better risk indicators of initiating use and progression to abuse.

Our measure of availability may have indexed perceived rather than actual cannabis availability, and so it remains an empirical question as to which is a better predictor of initiation and progression to abuse. One study has shown that perceived availability was a significant predictor of cannabis, heroin, alcohol and tobacco use but not of lysergic acid diethylamide (LSD) or non-prescribed tranquilizers [50]. The large ‘C’ component indicates a common environmental risk or set of risk factors shared between twins. Because the reporting period mainly spanned adolescence, it is possible that the ‘C’ is capturing aspects of the twin environment provided by the twins’ parents who may be more or less conducive to drug exposure and initiation by way of lack of parental supervision or parental attitudes towards drug use. Alternatively, the ‘C’ component in availability might reflect the general impact of neighbourhood, school or church effects, or the specific effects of correlated exposure to deviant peer groups or level of participation in prosocial activities [51,52].

Limitations

Our findings must be interpreted in the context of at least three potential limitations. First, our data were collected from white males born in Virginia. Previous analyses using the same data [6] suggested that this sample does not differ from the general population in rates of psychopathological conditions, including illicit substance use, and that it is likely to be broadly representative of US men. With regard to sex differences, rates of cannabis and other substance abuse are lower in females. Although this may be attributable to greater male exposure opportunities rather than greater chances of progressing from opportunity to use [53], the genetic and environmental pathways to abuse and dependence across a range of illicit substances still appear to be the same for men and

<table>
<thead>
<tr>
<th>Table 2 Percentage (%) of additive genetic variance, shared and non-shared environmental variance in cannabis availability, cannabis initiation and the number of DSM-IV symptoms of abuse. a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cannabis availability</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>Total genetic effects</td>
</tr>
<tr>
<td>% Associated with cannabis initiation</td>
</tr>
<tr>
<td>% Associated with cannabis availability</td>
</tr>
<tr>
<td>Unique genetic variance</td>
</tr>
<tr>
<td>Total shared environmental effects</td>
</tr>
<tr>
<td>% Associated with cannabis initiation</td>
</tr>
<tr>
<td>% Associated with cannabis availability</td>
</tr>
<tr>
<td>Unique shared environmental variance</td>
</tr>
<tr>
<td>Total non-shared environmental effects</td>
</tr>
<tr>
<td>% Associated with cannabis initiation</td>
</tr>
<tr>
<td>% Associated with cannabis availability</td>
</tr>
<tr>
<td>Unique non-shared environmental variance</td>
</tr>
</tbody>
</table>

| aBased on the best-fitting model 2.
women [3]. Nevertheless, it remains to be seen whether similar results will be obtained for females if sex differences in self-reported drug availability exist. Secondly, our data were collected retrospectively. Twins were asked to recollect drug availability, which introduces the potential for recall bias and telescoping effects [54]. In order to improve the quality of retrospective recall of the cannabis availability, the interview utilized a Life History Calendar (LHC) format as developed by Thornton [32]. By applying the principles of cognitive psychology, the LHC provides multiple cues to improve the subject’s chances of accurate recall by making the task more akin to the accurate and well-retained process of recognition than to the less reliable task of free recall. In addition, LHCs have been shown empirically to improve the accuracy of retrospective reporting [32,55]. As shown elsewhere [28], retest correlations were highest for substances obtained more easily and lower for the more illicit substances. Moreover, the correlations declined with greater intervening time intervals. All our analyses were therefore corrected for the linear effects of age at interview to remove, in part, possible cohort effects. Finally, our model-fitting strategy was not exhaustive. The CCC model is an improvement over previous separate modeling of drug use and abuse [1] because it models for the contingent nature between initiation and abuse. However, the single regression parameter (β) specifies that the causal contributions of A, C and E on initiation are proportional to their effects on abuse. It is possible that a feedback mechanism exists between the third and first stages, such that manifesting symptoms of abuse increases the likelihood of reporting increased drug availability. In other words, diagnosed individuals may be more likely to seek out cannabis actively. To test for this possibility, we began with model 2 and added a direct transmission path from the symptoms of cannabis abuse back to cannabis availability. This feedback model deteriorated significantly when compared to the saturated model 1 as judged by the change in log-likelihood (Δ-2LL = 7.96, Δdf = 2, P = 0.03). This model did not perform as well as model 2 in terms of the Bayesian information criterion (∆-2LL = 3921.985). Moreover, the standardized path coefficient from symptoms to availability was very small (β = 0.03).

CONCLUSION

Despite this paper’s narrow focus on cannabis, our findings are of broad etiological significance; these findings have begun to elucidate the causal processes underlying the genetic and environmental factors on drug use and abuse in terms of particular measured risk factors. The previously estimated shared environmental factors in initiation and abuse are correlated and can be traced back to aspects of the shared environment responsible for variation in cannabis availability. Similarly, we have shown how proportions of the genetic liability to initiate and then progress to abuse can be explained, in part, by genetic effects in availability. We still do not know if the A and C components in drug availability are the causal components or whether, for instance, the C component is an index of other, broader aspects of the shared environment such as peer group deviancy or lack of parental monitoring. Nevertheless, it is a valuable step forward to have decomposed into genetic and environmental components the covariance between a known, yet understudied, risk factor on adverse substance use.

Declarations of interest

None.

Acknowledgements

This project was supported by grants from the US National Institutes of Health (DA-11287, MH/AA/DA-49492, DA-18673, MH-01458 and AA-00236). Funding was also received from the US National Institute on Drug Abuse (1K99DA023549-01A2) and an Australian and NHMRC Sidney Sax Postdoctoral Fellowship. The authors thank Indrani Ray for database assistance. We thank Dr Linda Corey for assistance with the ascertainment of twins from the Virginia Twin Registry, now part of the Mid-Atlantic Twin Registry (MATR), directed by Dr Judy Silberg. The registry has received support from NIH, the Carman Trust and the W. M. Keck, John Templeton and Robert Wood Johnson Foundations.

References

7. Kendler K. S., Bulik C. M., Silberg J., Hettema J. M., Myers J.,

© 2009 The Authors. Journal compilation © 2009 Society for the Study of Addiction

Addiction, 104, 430–418


42. Kendler K. S., Myers J., Prescott C. A. The specificity of genetic and environmental risk factors for symptoms of


52. van den Bree M. B., Pickworth W. B. Risk factors predicting changes in marijuana involvement in teenagers. *Arch Gen Psychiatry* 2005; 62: 311–19.

