Neuroticism and impulsivity: Their hierarchical organization in the personality characterization of drug-dependent patients from a decision tree learning perspective

Sergi Valero a,⁎, Constanza Daigre a,b, Laia Rodríguez-Cintas a,b, Carmen Barra a,b, Montserrat Gomà-i-Freixanet c, Marc Ferrera a,d, Miquel Casas a,d, Carlos Roncero a,b,d

a Department of Psychiatry, Hospital Universitari Vall d'Hebron, CIBERSAM, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain
b Outpatient Drug Clinic (CAS) Vall Hebron, Hospital Vall d'Hebron-Agència de Salut Pública de Barcelona, Barcelona, Catalonia, Spain
c Department of Clinical and Health Psychology, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain
d Department of Psychiatry and Legal Medicine, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain

Abstract

Objective: Neuroticism and impulsivity are the personality variables most consistently associated with drug-dependent patients. To date, no data mining procedures have been applied to explore the differential role of personality variables in this population.

Methods: The personality profile of 336 drug-dependent patients was compared with that of a sample of community participants in the context of a decision tree learning approach using the Alternative Five Factor Model. The resulting discriminant model was cross-validated.

Results: Neuroticism and impulsivity were the most relevant variables in the resulting model, but their association appeared to be hierarchically organized. In the personality characterization of these patients, neuroticism became the main discriminant dimension, whereas impulsivity played a differential role, explained by means of an interaction effect. Decision tree learning models appear to be a heuristic theoretical and empirical approximation to the study of relevant variables, such as personality traits, in drug-dependency research.

© 2014 Elsevier Inc. All rights reserved.

The general statement that individual differences may lead some individuals to show overt signs of a specific disorder can also be applied to addictive disorders [1]. The influence of specific personality traits on addictive disorders has accumulated sufficient scientific background to be worth considering [2–4].

Research assessing normal personality in drug-dependent patients has identified neuroticism as one of the most robust factors characterizing this population [5–8]. The link between impulsivity and substance use disorders (SUDs) has also been well established [9,10], either measuring impulsivity by self-reports, neuropsychological measures, and/or delay tasks. Moreover, longitudinal studies have identified behavioral disinhibition as a relevant predictor of subsequent substance use in specific populations [11]. In fact, impulsivity is considered as a trait that cuts across a great deal of psychopathological categories, but its relevance has appeared to be particularly robust in substance use and other externalizing disorders [12]. Whereas neuroticism and impulsivity have consistently been associated with addictive disorders, little is known about the interaction or cumulative effect of these two factors simultaneously. Other personality dimensions have also been related to substance use, for example, conscientiousness [13], but more research is needed to consolidate its relevance.

Data mining is a non-trivial process of identifying novel, potentially useful, and valid patterns in databases [14]. This area of computer science is characterized by the use of various statistical techniques, databases, artificial intelligence, and pattern recognition [15]. It has commonly been used in several applied fields such as marketing, fraud detection, performance, medicine and scientific research, among others, contributing to the generation of new associations and improving the results obtained with other methodologies [16]. In the field of psychiatry and clinical psychology, its application is still incipient.

⁎ Corresponding author. Tel.: +34 934894295; fax: +34 934894587. E-mail address: svalero@vhebron.net (S. Valero).

http://dx.doi.org/10.1016/j.comppsych.2014.03.021
0010-440X/© 2014 Elsevier Inc. All rights reserved.
Decision tree learning is one of the strategies of the data mining procedure [17,18]. Using pictorial representations, this strategy provides discriminant nodes in which participants are recurrently classified. The data are submitted to a partition process under a supervised learning algorithm. The purpose of this process is to maximize the distance between groups in each partition. This process is repeated descendentely, selecting the next discriminant node in the tree. Compared with other data mining methods such as artificial neural network, support vector machine, or naïve Bayesian classifier, decision tree learning has been considered to be more easily interpreted due to its pictorial representation, which can be translated into a set of if-then-or else rules [19].

The attractiveness of applying decision trees to clinical settings is that they represent rules that can be easily expressed and interpreted by clinicians and patients alike [20]. These rules of thumb are interesting in terms of communicability and dissemination of medical and psychological knowledge. In personality research, data mining can be heuristic providing new data which can help in refining previous results and/or generating new hypotheses. To date, the use of decision tree learning models to explore the role of personality traits substance use is scarce and as far as we know, this technique has only been tested in adolescents with alcohol consumption [21].

To the best of our knowledge, this is the first study that uses a data mining procedure to identify discriminant personality variables in the field of drug dependency in adults. The aim of this research was to determine, by means of decision tree learning, which personality variables would discriminate a sample of adult drug dependents from a control group. In view of the strategy used, our study is exploratory and, consequently, no hypotheses were formulated.

1. Method

1.1. Participants

The sample consisted of 336 drug-dependent participants (75% men, age mean = 36.99, SD = 11.47) seeking treatment at the Drug Unit of the Psychiatry Department at the Hospital Universitari Vall d’Hebron during the period from March 2007 to June 2013. The inclusion criteria were as follows: being ≥18 years old, a diagnosis of drug dependence (alcohol, cocaine, opiates, cannabis, ecstasy and amphetamine) in accordance with the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition [22], and having signed informed consent and completed the assessment process. The exclusion criteria included the following: schizophrenia, bipolar disorder, mental retardation, or mental disorder due to a general medical condition.

The control sample was made up of 486 subjects, extracted randomly from a more comprehensive general population sample pool stratified by age and gender. As a consequence of this stratification, the samples did not differ either in age (t(818) = .33, p = .741) or gender (χ(1) = .07, p = .932).

1.2. Measures

The Semi-structured Clinical Interview for Axis I Disorders [22] was used to diagnose patients with drug dependence.

Personality was assessed using the Spanish version of the Zuckerman–Kuhlman Personality Questionnaire (ZKPQ) [23]. This questionnaire consists of five scales. (1) Neuroticism–Anxiety (N–Anx, 19 items) items describe frequent emotional upset, tension, worry, fearfulness, obsessive indecision, lack of self-confidence and sensitivity to criticism. (2) Activity (Act, 17 items) items describe the need for general activity, an inability to relax and do nothing when the opportunity arises, a preference for hard and challenging work, an active busy life and high energy level. Two facet scores can be obtained from this scale: Need for General Activity, impatience and restlessness (GenAct, 9 items) and need for Work Activity (WorkAct, 8 items). (3) Sociability (Sy, 17 items) items describe the number of friends one has and the amount of time spent with them, outgoingness at parties and a preference for being with others as opposed to being alone and engaging in solitary activities. Two facet scores can also be obtained: Parties and friends (Parties, 9 items) and Isolation Intolerance (Isol, 8 items). (4) Impulsive Sensation-Seeking (ImpSS, 19 items) items involve a lack of planning and the tendency to act without thinking and the seeking of excitement, novel experiences and willingness to take risks for these types of experiences. The ImpSS items are general in content and do not describe specific activities such as risky sports, drinking, having sex, or drug consumption. Two facet scores can be obtained from this scale: Impulsivity (Imp, 8 items) and Sensation Seeking (SS, 11 items). (5) Aggression–Hostility (Agg–Host, 17 items) items describe a readiness to express verbal aggression; rude, thoughtless or antisocial behavior; vengefulness and spitefulness; having a quick temper, and impatience with others. The ZKPQ also includes an Infrequency scale (Infreq, 10 items). Rather than being regarded as a scale in the normative sense, it should only be used to detect inattention to the task or simply as a validity measure for the individual test-taker. More details about the psychometric properties of the instrument in clinical and general populations are reported elsewhere [23,24].

Unlike the NEO Personality Inventory-Revised [25], the neuroticism dimension from the ZKPQ does not include either impulsivity or hostility traits, instead, specific scales for these two traits are included [26]. In addition, the ImpSS scale can be separated into two facets, impulsivity and sensation seeking, providing a more conceptually and empirically refined discrimination of drug-dependent patients.

1.3. Procedure

All referred participants underwent the standard assessment protocol. Patients were evaluated by two psychiatrists and a clinical psychologist experienced in diagnosing drug dependency. The ZKPQ was administered and scored blindly to the clinical evaluation. Patients were assessed at
the Drug Unit of the Psychiatry Department at the Hospital Universitari Vall d’Hebron. After assessment, patients underwent a personalized treatment fitting the specific clinical and personality profile. The average standard treatment lasted around 2 months. The research protocol was approved by the Hospital Ethics Committee. Patients did not receive any monetary compensation for their participation in this study.

Control subjects were a subsample extracted from a much larger one [24]. This community sample was obtained from different sites such as classrooms, leisure associations, yearly health check, etc. Individuals responded anonymously and only age, sex, and educational level were recorded. No economic reward was provided to participants.

1.4. Statistical analysis

The total sample (participants plus controls) was divided into two random subsamples: the training sample (60% of the total subjects) and the test sample (the remaining 40%). Data mining was conducted in the training sample using the decision tree learning model. As mentioned previously, this method approximates the function for the target attribute by learning from a decision tree from the previous examples. It provides recursive partitions of the sample according to a node that represents a split point. In this study, we used the scales and facets of the ZKPQ as discriminant factors, and the dichotomic status of participants (drug dependents vs. controls) as the conditions to be differentiated. When continuous discriminant variables are included, as in this study, the nodes represent discriminant points that maximally differentiate between levels of the two or more categories of the dependent factor (drug-dependents vs. controls). In order to avoid over-fitting of the model, results were limited to a maximum of three cut-offs[27], providing a model analogous to a three-way interaction. Only binary partitions were considered, and a pruning algorithm based on the Minimum Description Length Principle was selected [17]. Finally, a cross-validation procedure was conducted with the test sample. Chi-square and odds ratios are provided to test for the discriminant role of the cut-offs in the training and test samples. Analyses were performed using the free software Konstanz Information Miner (KNIME 2.7.1).

2. Results

Sociodemographic variables of the drug-dependent sample are reported in Table 1.

Table 2 and the Fig. show that, at the first step, data mining analysis identified N–Anx as the most powerful dimension discriminating between drug-dependent and control participants. The majority of drug-dependent participants were located at the upper partition (76.7%), while most of the control group participants were located at the lower partition (59.1%).

2.1. High Neuroticism–Anxiety

Subjects with high scores on Neuroticism–Anxiety were subsequently divided according to their Impulsivity scores, the second strongest discriminant factor between the two categories. The group scoring high on Neuroticism–Anxiety and Impulsivity was mostly made up of drug-dependent participants (69.3% of the total), whereas the group scoring high on Neuroticism–Anxiety and low on Impulsivity was mostly made up of controls (60.8%).

2.2. Low Neuroticism–Anxiety

No discriminant factors between drug-dependent and control participants were found in the subsample scoring low on Neuroticism–Anxiety.

2.3. High Neuroticism–Anxiety and low Impulsivity

In this subsample, Aggression–Hostility was identified as the most differentiating trait. Participants with either low or high scores on this trait were mostly controls (60.8%), but the prevalence of controls was higher at the lower partition of the node (72.2%).

2.4. High Neuroticism–Anxiety and high Impulsivity

When the combination of high Neuroticism–Anxiety and high Impulsivity occurs, Isolation Intolerance appears as the best differentiating variable. Individuals with low scores on Isolation Intolerance were mostly drug-dependent participants (81.1%), whereas those with high scores on Isolation Intolerance were mostly from the control group (59.5%).

2.5. Discriminability and replicability

Cut-off points obtained in the training sample were tested both in the training and test samples (Table 2). In the training sample, chi-square provided an estimation of the
The discriminant capacity of the model, separating drug-dependent versus control participants for every node. In the test sample, the chi-square became a measure of replicability. The cut-offs obtained in the training sample were cross-validated in the test sample. Odds ratios provided an estimation of the magnitude and confidence of the differences between partitions. For both samples (the training and test samples), and for all nodes, chi-square analyses were statistically significant. All odds ratios were >2 (range: 2.69–4.08 in the test sample), indicating that when applying the corresponding cut-off, the probability of switching from one partition to the other was at least twice as high.

Table 2
Summary of the Decision Tree Model (training sample) and its cross-validation (test sample).

<table>
<thead>
<tr>
<th>Step</th>
<th>Group/Subgroup</th>
<th>Discriminant variable (score cut-offs)</th>
<th>Chi-square (Sig)</th>
<th>OR (95% CI)</th>
<th>Chi-square (Sig)</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Primary sample</td>
<td>N–Anx (≤7 vs &gt;7)</td>
<td>56.15 (p &lt; .005)</td>
<td>4.77 (3.14–7.24)</td>
<td>32.56 (p &lt; .005)</td>
<td>3.73 (2.37–5.86)</td>
</tr>
<tr>
<td>2a</td>
<td>Those with low N–Anx</td>
<td>No discriminant variable</td>
<td>22.18 (p &lt; .005)</td>
<td>3.49 (2.08–5.86)</td>
<td>10.36 (p &lt; .005)</td>
<td>2.69 (1.50–4.83)</td>
</tr>
<tr>
<td>2b</td>
<td>Those with high N–Anx</td>
<td>Imp (≤4 vs &gt;4)</td>
<td>4.29 (p = .038)</td>
<td>2.34 (1.11–4.94)</td>
<td>7.78 (p = .005)</td>
<td>4.08 (1.59–10.52)</td>
</tr>
<tr>
<td>3a</td>
<td>Those with high N–Anx and low Imp</td>
<td>Agg–Host (≤7 vs &gt;7)</td>
<td>14.42 (p &lt; .005)</td>
<td>0.16 (0.07–0.37)</td>
<td>5.63 (p = .018)</td>
<td>0.32 (0.14–0.77)</td>
</tr>
<tr>
<td>3b</td>
<td>Those with high N–Anx and high Imp</td>
<td>Isolation (≤4 vs &gt;4)</td>
<td>47.46 (p &lt; .005)</td>
<td>2.34 (1.11–4.94)</td>
<td>7.78 (p = .005)</td>
<td>4.08 (1.59–10.52)</td>
</tr>
</tbody>
</table>

N–Anx = Neuroticism–Anxiety; Imp = Impulsivity; Agg–Host = Aggression–Hostility; Isolation = Isolation Intolerance.

Fig. Decision tree in the training sample. Note: N–Anx = Neuroticism–Anxiety; Imp = Impulsivity; Agg–Host = Aggression–Hostility; Isola = Isolation Intolerance.
3. Discussion

To our knowledge, this study is the first one to use a data mining procedure, decision tree learning, to explore the differential role of personality variables in drug-dependent and control individuals. The results of this discriminant process are replicated using a cross-validation strategy.

The first node obtained in the tree structure was the neuroticism–anxiety dimension, while trait impulsivity was the second one. The combination of these two personality variables and the sequence in which they were obtained is of great importance in the field of substance use disorders. Two considerations are worth commenting on. First, neuroticism and impulsivity appeared to be the most discriminant factors obtained in the decision tree learning analysis, in accordance with the literature on pathological substance use [28]. Second, as a consequence of the intrinsic discriminant process of the decision tree strategy, neuroticism and impulsivity did not play the same discriminant role when comparing drug-dependent and control individuals. Neuroticism, the variable obtained in the first discriminant node, became the personality dimension with the highest capacity to differentiate between the two samples. Impulsivity was located in the second order, but its relevance in the discriminant process was due to the fact that it only applies to individuals with high scores on Neuroticism–Anxiety. The interaction effect observed between neuroticism and impulsivity provides a specific way to delimit the role of personality traits in the conceptualization of drug-dependent individuals.

The association between anxiety, an overt expression of the underlying neuroticism dimension, and impulsivity, has been highlighted within addictive disorders literature. The anxious-impulsivity concept was first conceptualized by Newman and Wallace [29] as a lack of behavioral control in situations where it seemed impossible to escape from aversive consequences. More recently, a similar concept, defined as negative urgency [30], has been considered as a form of an impulsive tendency in response to intense negative affect or stressful conditions [1,31]. According to these considerations, it has been suggested that high levels of anxious-impulsivity could be a consequence of deficits in emotion regulation associated with the frontal–limbic brain system. Impulsivity is multifaceted, and the combination of an impulsive and anxious personality could be considered as an endophenotype of drug dependence [32].

The results obtained in our study are consistent with this point of view. However, as mentioned above, the sequence in which these results were obtained demands a more in-depth consideration. Low levels of neuroticism, operationalized in this study as having a score ≤7 on the N–Anx factor of the ZKPQ, and without considering the scores obtained on impulsivity, appear to be a protective factor against being endorsed in the drug-dependent group. The probability of being included in the drug-dependent group is conditioned by the presence of high levels of neuroticism, and this is the only situation where impulsivity becomes empirically relevant.

Neuroticism is the single most important factor associated with many forms of psychopathology and mental health, particularly with common mental disorders including anxiety, depression and substance use disorders [33]. Hence, it has been suggested to be a crucial personality factor with public health significance [5]. However, its omnipresence in personality research with psychopathological samples sometimes causes it to be considered as a non-informative marker of psychopathology [34]. By adding impulsivity to the process of characterizing of substance use disorders, a more specific and clinically relevant approach is provided. However, when a classical statistical approach is applied (e.g., regression models or mean comparisons) neuroticism and impulsivity, or other traits, are generally reported according to their statistical significance status, but their relevance is not hierarchically organized. As a consequence, every variable is reported as if its relevance was similar, therefore its importance is usually expressed monotonically without considering the potential interaction effects among them [9,13,35,36]. According to our findings, the subordination of impulsivity as a relevant factor only in highly anxious subjects is a step forward in the characterization of drug-dependent individuals. This provides a more specific approach by organizing and furthering our understanding of personality traits in drug-dependent or other psychiatric populations.

In this research, different addictive substances are included and simultaneously analyzed. This strategy could be interpreted as a loss of specificity for the results obtained, but the relevance of neuroticism and impulsivity has been repetitively replicated in different addictive substances [10,28,37]. We hypothesize that if other discriminant personality variables entered in the model, they would probably be located at lower-level nodes in a decision tree learning model. In our study, the variables located at the last explored node were isolation intolerance and aggression–hostility. These variables appeared as factors discriminating between high- versus low-impulsive subjects in the high-anxious subsample. The difficulty of extracting consistent conclusions at the last node explored is obvious. Two different variables were obtained at the third node, and this interaction is an intrinsically complex situation to interpret. Moreover, the relevance of these variables should be interpreted only in terms of discriminability in a small and specific subsample of individuals, those who were filtered in the previous discriminant nodes. The relevance of these variables is probably influenced by the inherent and specific characteristics of this sample (e.g., inclusion and exclusion criteria, etc.). Therefore, this study should be considered as a first approach to the study of personality dimensions in drug-dependent populations using a data mining procedure. More research is needed to explore the relevance of these three variables by analyzing, for example, the distinct personality profiles associated with the specific addictive substances consumed. Another point worth mentioning is the nature of the control group. These participants were extracted from a
more comprehensive general population sample that formed part of a wider study aimed at obtaining normative data for the ZKPQ. As a consequence, these individuals did not have any assessment other than the personality questionnaire itself. It is not possible to describe this sample in terms of drug dependence or substance use. However, this control group, paired with the drug-dependent patients by age and gender following a random stratification selection and recruited from different social contexts as described above, provides a group of comparison endowed with a high ecological validity.

To conclude, the present study implements a data mining procedure, decision tree learning, to obtain the discriminant personality traits in adult drug-dependent patients, contrasting the obtained results by means of a cross-validation procedure. The outstanding results of this study show that neuroticism and impulsivity are the most relevant personality traits to be taken into account in the conceptualization of drug dependence. However, their relevance must take into account their hierarchical relationship. In drug-dependent individuals, impulsivity only becomes a key discriminant factor when high neuroticism is present. The decision tree learning method is an effective and easy-to-understand procedure to be used in the analysis of complex relationships. This method leads to a refinement of previously obtained results not only by providing an operative tool to make useful decisions in clinical settings (i.e., improving diagnosis) but also by delimiting and prioritizing key clinical variables when a specific therapeutic intervention has to be implemented.

References


