Mémoire de Maîtrise en médecine

Exploring the proposed DSM-5 diagnostic threshold and severity scale for the alcohol use disorder.

Etudiant
Danielle Rosenfield

Tuteur
Prof. Martin Preisig
Dpt de psychiatrie, CHUV

Co-tuteur
Mme Caroline Vandeleur
Dpt de psychiatrie, CHUV

Expert
Prof. Jacques Besson
Dpt de psychiatrie, CHUV

Lausanne, le 14 avril 2012

Abstract

Aims : Several studies have questioned the validity of separating the diagnosis of alcohol abuse from that of alcohol dependence, and the DSM-5 task force has proposed combining the criteria from these two diagnoses to assess a single category of alcohol use disorders (AUD). Furthermore, the DSM-5 task force has proposed including a new 2-symptom threshold and a severity scale based on symptom counts for the AUD diagnosis. The current study aimed to examine these modifications in a large population-based sample.

Method : Data stemmed from an adult sample (N=2588 ; mean age 51.3 years (s.d.: 0.2), 44.9% female) of current and lifetime drinkers from the PsyCoLaus study, conducted in the Lausanne area in Switzerland. AUDs and validating variables were assessed using a semi-structured diagnostic interview for the assessment of alcohol and other major psychiatric disorders. First, the adequacy of the proposed 2-symptom threshold was tested by comparing threshold models at each possible cut-off and a linear model, in relation to different validating variables. The model with the smallest Akaike Criterion Information (AIC) value was established as the best model for each validating variable. Second, models with varying subsets of individual AUD symptoms were created to assess the associations between each symptom and the validating variables. The subset of symptoms with the smallest AIC value was established as the best subset for each validator.

Results : 1) For the majority of validating variables, the linear model was found to be the best fitting model. **2)** Among the various subsets of symptoms, the symptoms most frequently associated with the validating variables were : a) drinking despite having knowledge of a physical or psychological problem, b) having had a persistent desire or unsuccessful efforts to cut down or control drinking and c) craving. The least frequent symptoms were : d) drinking in larger amounts or over a longer period than was intended, e) spending a great deal of time in obtaining, using or recovering from alcohol use and f) failing to fulfill major role obligations.

Conclusions : The proposed DSM-5 2-symptom threshold did not receive support in our data. Instead, a linear AUD diagnosis was supported with individuals receiving an increasingly severe AUD diagnosis. Moreover, certain symptoms were more frequently associated with the validating variables, which suggests that these symptoms should be considered as more severe.

Keywords : alcohol use disorders, DSM-5, threshold, severity scale.

1.0 Introduction

The use of excessive alcohol is a major public health issue due to the many consequences involved. The medical consequences, affecting most of the body organs include gastritis and gastrointestinal tumors (Bujanda 2000), liver cirrhosis (Stoffolini et al., 2010), pancreatitis (Yadav and Whitcomb 2010), cardiomyopathy, heart failure, hypertension, arrythmias and stroke (Movva and Figueredo 2012) decreased immunity (Brown et al., 2006), peripheral neuropathy (Mellion et al., 2011), myopathy (Preedy et al., 2003), brain damage (Harper 2009) and decreased fecundity (Hassan and Killick 2004). In addition, the social consequences are vast, such as disruption of familial and social relationships and occupational impairment (Klingemann 2001). Another complication is traffic accidents due to people driving under the influence of alcohol (Centers for disease Control and Prevention 2010). All these problems impose a high cost on society with considerable investments in medical care, social aid, and reparations of material damage as well as considerable loss in salaries of the incapacitated or prematurely deceased (Rehm et al., 2007).

In order to more fully assess the impact, consequences and prevalence of alcohol use disorders (AUDs), research has focused on developing valid diagnoses over the last two decades. More specifically, a large body of research has been gathered on AUDs described in the fourth edition of the Diagnostic and Statistical Manual for Mental Disorders (DSM-IV) (American Psychiatric Association, 2000). According to the DSM-IV, an individual classifies for "substance dependence" if three or more of the following criteria occur at any time in the same 12 month period: 1) tolerance, 2) withdrawal, 3) the substance is taken in larger amounts or over a longer period than was intended ("larger/longer"), 4) there is a persistent desire or unsuccessful efforts to cut down or control substance use ("cut down/control"), 5) a great deal of time is spent in activities necessary to obtain the substance, use the substance or recover from its effects ("time spent"), 6) important social, occupational, or recreational activities are given up or reduced because of substance use ("activities given up"), 7) the substance use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the substance ("physical or psychological problems"). The DSM-IV also contains a separate diagnosis termed "substance abuse", for which an individual classifies if 1 or more of the following symptoms occur within a 12-month period: 1) recurrent

substance use resulting in a failure to fulfill major role obligations at work, school, or home ("failure to fulfil roles"), 2) recurrent substance use in situations in which it is physically hazardous ("hazardous use"), 3) recurrent substance-related legal problems ("legal problems"), 4) continued substance use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of the substance ("social or interpersonal problems").

Relying on the extensive research on AUDs, the Substance Related Disorders Work Group of the DSM-5 task-force (APA, 2010) has proposed four modifications of the AUD diagnosis in view of the fifth edition of the DSM (DSM-5).

The first modification is the combination of alcohol dependence and alcohol abuse criteria. The DSM-IV conceptualizes alcohol dependence and abuse as two distinct yet hierarchical diagnoses. Indeed, a diagnosis of abuse, considered a milder form of AUDs, cannot be made if the criteria for dependence have ever been met for alcohol, except if abuse precedes dependence. The theoretical basis for this distinction originally stems from Edwards and Gross's conceptualization of alcohol use disorders (1976) which distinguished dependence, defined as impaired control over drinking, from abuse, defined by the adverse social and personal consequences alcohol may engender.

Research has questioned the validity of separating abuse and dependence diagnoses into two discrete categories using confirmatory factor analysis (CFA). This method determines whether an observed set of criteria fit onto a single common factor or if they fit better on multiple factors. Applying CFA to observed abuse and dependence criteria in a national sample of young adults, Harford and Muthén (2001) provided some evidence for a 2-factor model with one factor representing dependence and one representing abuse but the factors did not perfectly match the DSM categories. The factor defined as "dependence" contained items of the DSM abuse category (for example "failure to fulfil roles") and vice versa. Also, the 2 factors showed high inter-correlations. Last, when correlated to external validators, the 2 factors had many similar correlations for validators including family history and consumption patterns. Grant et al. (2007) found evidence for a 3-factor model: the factors being labelled "dependence", "abuse" and a smaller factor with one symptom, "tolerance". However, the 3 factors were again highly inter-correlated and shared

some correlations to external validators such as parental father history of antisocial personality disorder and treatment seeking in the last 12 months. Although the authors concluded in supporting the DSM-IV dependence category (with a more reserved opinion on the abuse category), they suggested the possibility of a single underlying dimension. Using the same method as the two previous studies, Proudfout et al. (2006) found both a one and two factor model to show an adequate fit to their data but considering the extremely high correlation between the two factors, they preferred a more parsimonious 1 factor model including all abuse and dependence criteria. Following this study, many others (Langenbucher et al., 2004, Saha et al., 2006, Borges et al. 2010, Dawson et al., 2010, Shmulewitz et al., 2010, Hagman and Cohn 2011) have confirmed a preference for a unidimensional AUD trait combining abuse and depdence criteria into one single factor.

The second modification introduced in the DSM-5 proposal concerns the removal of the "legal problems" criterion (a former abuse criterion) and the addition of a "craving" criterion. The removal of the "legal problems" criterion is based on it's extremely low prevalence and the consensual better model fit when it is removed as shown in different studies using CFA and item response theory (IRT) [1] procedures (Langenbucher et al., 2004, Proudfoot et al., 2005, Saha et al., 2006, Schmulewitz et al., 2010).

Concerning the addition of the craving criterion, one study (Keyes et al., 2010) showed that: 1) when combined to the other dependence and abuse criteria, the craving criterion fit on a single-factor model using CFA, 2) the craving criterion fit on a unidimensional IRT scale along with dependence and abuse criteria, ranking high in severity, 3) model fit values (Bayesian Information Criterion (BIC) and Aikake Information Criterion (AIC))[2] for the IRT models were lower, showing a better fit of the model, when craving was included, and 4) the craving item was highly correlated with different external validator variables such as age of onset, likelihood of prior dependence and current major depression diagnoses. Keyes et al. also pointed out

---

[1] IRT two-parameter models examine the relationship between the observed data to the criteria and the underlying unobserved latent trait. The *a* parameter represents the discrimination power of a criterion, i.e the capacity of a criterion to discriminate individuals on the higher end of the continuum and those on the lower end ; it defines to which extent the criterion is related to the underlying construct. The *b* parameter represents the severity of a criterion, ; the symptom severity is directly related to the symptom prevalence, the higher the severity, the less it is endorsed.

[2] AIC and BIC are model selection statistics where the lowest AIC or BIC value corresponds to the best fitting model.

that since the ICD-10 contains a craving criterion, the addition of such a criterion in the DSM-5 would facilitate comparisons between the two classification systems.

Another study (De Bruijn et al., 2005) compared individuals with AUDs according to 3 different alcohol use disorder classifications: the DSM-IV, the ICD-10 and the Craving Withdrawal Model (CWM). The CWM, a classification model created by the authors themselves[3], attributes a dependence diagnosis if an individual endorses both the craving and the withdrawal items and attributes an abuse diagnosis if an individual endorses 2 or more of the remaining 10 DSM-IV dependence and abuse criteria. In order to compare the models, De Bruijn et al. diagnosed each individual with no diagnosis, an abuse diagnosis (or harmful use diagnosis in the ICD-10 classification) and a dependence diagnosis according to the 3 different classifications and then compared their mean scores on discriminant validators (e.g alcohol intake, psychiatric comorbidity, functional status). They showed that the CWM had the best discriminant validity of the three systems, meaning that it best separated individuals with no diagnosis, those with an abuse diagnosis and those with a dependence diagnosis in relation to the validators. This implied that the craving and withdrawal criteria could be sufficient to diagnose an AUD.

The third modification proposed by the DSM-5 task force consists of defining a new 2-symptom threshold for the combined abuse and dependence criteria together with the craving criterion. The delimitation of a threshold to account for the presence or absence of a psychiatric disorder has often been questioned in regard to the debate over the categorical or dimensional nature of diagnoses. Indeed, categories facilitate communication between clinicians, diagnostic decisions and insurance coverage whereas dimensions better reflect the complexe nature of a diagnosis and offer ordinal scaling for research which increases the overall power of the applied tests (Hasin and Beseler 2009).

The validity of the current 3-symptom threshold for dependence and 1-symptom threshold for abuse set by the DSM-IV has been studied by different researchers. Grove et al. (2010) examined each potential threshold in relation to 7 external validators: mental health, disability, early age of onset, suicidality, psychological distress, days out of role and service use. The aim was to validate a defendable

---

[3] inspired by the Withdrawal-Gate model developped by Langenbucher (Langenbucher et al., 2000)

threshold by indicating significant group differences on validating variables at a given threshold. When examining abuse and dependence separately, the BIC values were lowest at a threshold of 3 for dependence and 1 for abuse in 4 out of 7 validating variables, partially validating the current thresholds However, when the abuse and dependence criteria were combined, no clear threshold could be confirmed.

Hasin and Beseler (2009) used the validating variables family history and early drinking onset to examine the appropriate threshold in a population of lifetime and current drinkers. Again, a discontinuity in the relationship between the number of symptoms and the validating variables was used to identify a potential threshold. A dimensional model (i.e. no threshold) was compared to threshold models for 1) the dependence criteria, 2) the dependence and the abuse criteria combined, 3) the dependence and the abuse criteria and a hypothetical new criterion, binge-drinking. For dependence criteria alone as well as for dependence and abuse criteria combined, a dimensional model provided the best fit for the relationship between number of symptoms and the 2 validating variables suggesting an underlying severity continuum in place of a defined categorical threshold. In this view, a determined cutoff would reflect nothing more than an arbitrary decision. However when binge-drinking was added to the dependence and abuse criteria to examine it's potential for being incorporated as a new diagnostic criterion, the dimensional model did not fit the data indicating departure from linearity. The authors therefore advised caution in it's addition to an otherwise dimensional model composed of abuse and dependence criteria.

The threshold topic has also been examined though the analysis of "diagnostic orphans", who represent individuals who don't meet the DSM-IV diagnostic threshold for either dependence or abuse as they have only one or two dependence symptoms and no abuse symptoms. The course of these individuals and their risk of later meeting a full dependence diagnosis have been established in two studies (McBride and Adamson 2010; Harford et al., 2010). McBride and Adamson (2010) showed adult "diagnostic orphans" to have higher rates of progression to alcohol abuse and even higher rates of progression to dependence than those with no AUD symptoms at a three-year follow-up. Indeed, "diagnostic orphans" with 1 or 2 dependence symptoms were respectively 3 or 4 times more likely to progress to the abuse group and 4 or 8 times more likely to progress to dependence at follow-up. However no pattern of symptom endorsement could be identified at baseline in order to identify

the "diagnostic orphans" who remitted and those who progressed to abuse or dependence.

These results were replicated in another longitudinal study (Harford et al., 2010) with a five-year follow-up which also showed "diagnostic orphans" to be more at risk for progressing to dependence at follow-up. Moreover, Harford et al. showed "diagnostic orphans" to have higher levels of heavy drinking and drug use than those with no AUD symptoms although lower than those with alcohol abuse and dependence. These findings suggest that the AUD diagnoses should include less severe forms of alcohol dependence and that individuals with one or two symptoms deserve clinical attention in order to prevent progression to a full blown diagnosis.

Nevertheless, some contradictory results were found in an earlier study using a community sample to establish the appropriate threshold for the 7 symptom alcohol dependence category (Hasin and Paykin, 1998). This study showed that, although "diagnostic orphans" compared to the "no diagnosis" group exhibited higher percentages of binge drinking, current drug use and a family history of treatment for alcohol, they were not more likely than the "no diagnosis" goup to develop a dependence diagnosis at one year follow-up. In a subsequent study using a national sample, the same authors (Hasin and Paykin, 1999) found similar results and advocated for maintaining a minimum of three symptoms for a diagnosis of alcohol dependence as defined by DSM-IV.

A cross-sectional study (Ray et al., 2008) also found results in favour of maintaining the current thresholds for abuse and dependence in a outpatient treatment-seeking population. When comparing "diagnostic orphans" to patients with dependence, patients with abuse and patients with no diagnosis, the authors found that "diagnostic orphans" differed significantly in regard to certain external validators. Most of all, they had higher rates of lifetime diagnoses of cannabis dependence than patients with no diagnosis and lower rates of lifetime diagnoses of substance use disorders (SUDs) other than alcohol compared to patients with abuse or dependence. The interpretation was that "diagnostic orphans" constituted a separate and distinct category from individuals with abuse or dependence with a lower risk of progression to various SUDs and threrefore should not be incorporated in the revised DSM-IV. However it could argued that the fact that "diagnostic orphans" constitute a separate category from patients with no diagnosis and patients with dependence or abuse does not mean that they could not be viewed as a less severe form of an AUD. A

longitudinal study would have to complement this finding to show that "diagnostic orphans" don't progress to a DSM-IV AUD.

Therefore, the question of the appropriate thresholds for AUDs still remains controversial. Moreover, to our knowledge there are no studies yet which have tested the 2-symptom threshold using the newly suggested combination of alcohol abuse and dependence criteria taken together with the craving item (DSM-5 proposition).

The final proposition for the upcoming DSM-5 consists of the addition of a severity scale for the AUD diagnosis. Thus, the AUD diagnosis would be comprised of both a categorical and a dimensional aspect. The proposed scale uses symptom counts to scale severity: an individual with 2-3 criteria receives a "moderate AUD diagnosis" and an individual with 4 or more criteria receives a "severe AUD diagnosis".

However, there is a debate as to whether the severity indicator should be a simple symptom count, as proposed by the DSM-5 task force, or whether it should include a hierarchy in the symptoms with certain symptoms having more weight in the diagnosis than others, i.e. being predictive of worse outcome or risk factors.

Several IRT studies were in favour of symptom weighting, arguing that AUD criteria which are most discriminating and severe deserve more weight as they represent a more advanced stage of the disorder. The first IRT study on alcohol, cannabis and cocaine by Langenbucher et al., (2004) showed that in contradiction to former assumptions that abuse precedes dependence, making it a milder diagnosis, dependence criteria were actually often less severe than abuse criteria. For alcohol use disorders, the most severe criteria were "failure to fulfill roles" (DSM-IV abuse criterion) and "activities given up" (DSM-IV dependence criterion) and the least severe criterion were "physical/psychological problems" (DSM-IV dependence criterion) and "larger/longer" (DSM-IV dependence criterion). These results questioned the hierarchy between abuse and dependence. The authors also found all the criteria to be clustered in a moderate range of severity conferring little distinction possibilities between severe and mild forms of a disorder. They proposed an potential reduction of the number of criteria and their replacement by criteria reflecting severe and mild forms of the disorder.

Using an American population-based sample, Saha et al., (2006) also showed that the abuse criteria did not reflect a milder form of the AUD diagnosis than the dependence criteria. The least severe criteria identified were "larger/longer" (DSM-IV

dependence criterion), "withdrawal" (DSM-IV dependence criterion), "tolerance" (DSM-IV dependence criterion) and "cut down/control" (DSM-IV dependence criterion), whereas the most severe were "failure to fulfill roles" (DSM-IV abuse criterion) and "activities given up" (DSM-IV dependence criterion).

Hagman and Cohn (2011) showed partially similar results in a sample comprised of college students. Indeed, the criteria "tolerance" (DSM-IV dependence criterion) and "larger/longer" (DSM-IV dependence criterion) tapped the least severe end of the diagnostic continuum and the criterion "failure to fulfill roles" (DSM-IV abuse criterion) tapped the severe end. However, the criterion "activities given up" (DSM-IV dependence criterion) was not as severe as in the Saha et al. study.

In summary, these IRT studies tended to show a severity ranking with certain symptoms having more weight then others. They also reached the conclusion that alcohol abuse and dependence items lie on a unidimensional scale with all the symptoms intermixed on a continuum of severity.

However, severity ranking based on symptom weighting was questioned in a recent study (Dawson et al., 2010). This study used IRT to rank each abuse and dependence criterion on a severity scale and assessed mean values for each criterion on other severity indicators (i.e validating variables) including consumption patterns, psychological functioning, family history, antisocial behaviour and early onset of drinking in order to validate IRT ranking. Results showed similar IRT severity ranking as that in the Saha et al. and Harford et al. studies. Also, the mean values of the validating variables for each criterion tended to increase with increases in severity ranking. However, the authors also adjusted the results for age, sex and other criteria endorsed. After these adjustments, the relationship between the severity ranking and the mean values on the outcome variables was not maintained. The interpretation was that individuals endorsing high severity criteria, for example "failure to fulfil roles" actually endorsed many other criteria. In this sense a scalar measure of the AUD could do equally well by relying on simple symptom counts of criteria as by using weighted scales with varying criteria.

In another study (Moss et al., 2008), the authors used latent class analysis to create clusters of endorsement patterns and correlated these clusters to severity indicators (i.e. external validators) to examine the predictive values of their clusters. The authors found evidence for 6 clusters of symptoms following a Guttman like scaling (cumulative addition of 1 criterion to the previous cluster) but these clusters were not

linearly associated to severity indicators except for the 2 extreme clusters (individuals with 6 and 7 symptoms respectively). Therefore the authors therefore did not support a severity ranking based on symptom counts.

In light of these studies, the severity scaling remains controversial with no consensus as to how to determine the severity of an AUD.

To date, few studies have yet assessed the proposed DSM-5 changes. To our knowledge, 3 studies (Hagman and Cohn 2010, Agrawal et al 2011, Mewton et al 2011) have examined the change in prevalence of the AUD diagnosis between the actual DSM-IV diagnosis and the newly proposed DSM-5 diagnosis. Prevalences incresased respectively by 14.1%, 11.3% and 61% in indivuduals receiving an AUD diagnosis with the proposed DSM-5 revision as compared to DSM-IV which was mainly due to the inclusion of former "diagnostic orphans" with 2 symptoms.

The Hagman study correlated to external validators groups of individuals with a "moderate" or "severe" AUD according to the proposed revison and showed that individuals with a "severe" AUD had the worst values on the external validators, supporting the validity of the new diagnostic scale.

The Mewton study examined the unidimensionality of the diagnosis with the DSM-5 proposed criteria, concluding in the persistence of unidimensionality with the addition of the craving criterion and with the removal of the legal problems criterion. The same study also approached the threshold problem by concluding in maximisation of agreement between DSM-IV and DSM-5 if a 3 symptom threshold was maintained. However it should be noted that the validation of a threshold by comparing old diagnostic criteria to new diagnostic criteria seems questionnable due precisely to the obsolescence of the old criteria.

Based on the controversial results of earlier studies, the aims of this study are to examine
1) the best threshold for the AUD diagnosis.
Indeed no study has yet confirmed the proposed DSM-5 2-symptom threshold. By correlating different external validators to different thresholds, we will examine whether the AUD diagnosis should be considered categorical and, if so which threshold is the most relevant, or if the diagnosis reflects more of a dimensional nature.

2) the severity of the individual criteria for the AUD diagnosis. We will correlate each individual symptom to external validators to identify if certain symptoms lead to worse outcome or comorbid states than others.

Additional evidence on these issues will contribute to defining the AUD for the DSM-5 which will be a valuable guide for clinicians and researchers alike.

2.0 Methods

2.1 Sample and procedures

The data was obtained from the PsyCoLaus study (Preisig et al., 2009), the psychiatric arm of the population-based cohort study, CoLaus (see Firmann et al., 2008 for a description). The CoLaus survey assessed cardiovascular risk factors and genetic variants and biomarkers associated with these risk factors in a sample of 6738 subjects (92% caucasians), aged 35-75 years in 2003, from the general population of Lausanne, Switzerland, randomly selected from the complete list of residents of the city provided by the population registry.

All 35 to 66 year-old subjects of the CoLaus sample (n=5,535) were invited to participate in the psychiatric evaluation of the study (PsyColaus). Sixty-seven percent of the participants agreed, which resulted in a final sample of 3717 individuals who underwent both somatic/cardiovascular and psychiatric exams. The gender distribution of the PsyColaus sample approximated that of the general population in the same age range although the youngest 5-year band of the sample was underrepresented and the oldest 5-year band overrepresented (Preisig et al., 2009). Participants of PsyCoLaus and individuals who refused to participate did not differ on scores of the General Health Questionnaire (GHQ-12 : Goldberg 1972 ; French translation : Bettschart et al., 1996), a self-rating instrument which assessed psychiatric symptoms at the time of the physical exam (Preisig et al., 2009). The aims of the PsyCoLaus study were to: 1) establish the lifetime and 12-month prevalence of threshold (DSM-IV) and subthreshold psychiatric syndromes and migraine; 2) test the validity of postulated definitions for subthreshold psychiatric disorders using comorbidity patterns, risk of suicidal attempts, health service use, social functioning (Global Assessment of Function scores, GAF) and family history as

validator variables; 3) determine the association between psychiatric disorders and risk factors for cardiovascular disease; and 4) identify genetic variants and biomarkers that can modify the risk for psychiatric disorders and for comorbid cardiovascular disease and psychiatric disorders (Preisig et al., 2009).

The Institutional Ethic's Committee of the University of Lausanne approved the CoLaus and subsequently the PsyCoLaus study. All participants signed a written informed consent after having received a detailed description of the goal and funding of the study (Preisig et al., 2009).

The present paper includes 2588 individuals of the PsyCoLaus study. To be included in the study, participants had to be current drinkers or lifetime drinkers, defined as those having consumed at least one alcoholic drink per week for 6 months at any time, until presently (current drinkers) or in their lifetime. The mean age of the sample was 51.3 years (s.d.: 0.2), 44.9% were women and the mean level of socioeconomic status (SES) was 3.5 (s.d.: 0.0) according to the Hollingshead scale (Hollingshead, 1975). Almost three quarters of the sample were Swiss citizens (73.0%). In addition, data from all interviewed and non interviewed first-degree family members of probands were available for family aggregation analyses.

2.2 Measures

Probands and participating family members underwent face-to-face interviews. The interviewers were required to be masters-level psychologists, who were trained over a two-month period. During data collection, each interview was reviewed by an experienced senior psychologist.

The presence of psychiatric disorders was assessed using the semi-structured Diagnostic Interview for Genetic Studies (DIGS : Nurnberer et al., 1994). The DIGS was developed by the National Institute of Mental Health Molecular Genetics Initiative for the precise assessment of phenotypes through a large spectrum of DSM-IV Axis I criteria. The French translation of the DIGS (Leboyer et al, 1995) revealed excellent inter-rater reliability for major mood and psychotic disorders (Preisig et al., 1999) as well as for SUDs (Berney et al., 2002), although the 6-week test-retest reliability was slightly lower (Preisig et al., 1999; Berney et al., 2002).

Using the alcohol section of the DIGS, interviewers collected information on the 11 proposed criteria for the DSM-5 AUD diagnosis: D1- ("tolerance"), D2- ("withdrawal"), D3- ("larger/longer"), D4- ("cut down/control"), D5- ("time spent"), D6- ("activities given up"), D7- ("physical or psychological problems"), A8- ("failure to fulfill roles"), A9- ("hazardous use"), A10- ("social or interpersonal problems"). The new proposed criterion was added: N11- ("craving") and the "legal problems" criterion was removed.

The DIGS also provides questions on age of onset of regular consumption, the typical quantity of alcohol consumption, the largest quantity of alcohol consumption within 24 hours, treatment seeking behavior including discussion with a professional, help from AA (alcoholics anonymous) or other professionals, hospitalization or outpatient treatment for alcohol problems, current, lifetime and worst scores on the Global Assessment of Functioning scale (GAF), suicide attempts and comorbid disorders including mood and anxiety disorders, post-traumatic stress disorder (PTSD), antisocial personality disorder, attention-deficit hyperactivity disorder (ADHD) and marijuana, cocaine ad narcotics use disorders. These variables were used as external validator variables.

Family history information was collected from probands and participating relatives on all participants and non participants using the Family History-Research Diagnostic Criteria (FH-RDC) (Andreasen et al., 1977). The validity of the French version of the FH-RDC has been established though the assessment of agreement between diagnoses derived from direct interviews and family history reports for a series of diagnoses in adults (Rougemont-Buecking et al., Vandeleur et al., 2008) and children (Rothen et al., 2009).

2.3 External validator variables

The external validator concept has been used in psychiatry since its definition by Robins and Guze in 1970 (Robins and Guze 1970). It consists of testing a hypothesis using external correlates which are either risk factors, comorbidities or behavioral patterns related to a psychiatric diagnosis but not included in the diagnosis (Kendler 1990). Some validators are common to all psychiatric disorders such as treatment seeking behaviors or familial aggregation and some are more specific to certain diagnoses such as, for example, age of onset as a risk factor for AUD.

In view of the many validators used in this study, we have divided them into four categories: course variables, treatment-seeking variables, comorbidity variables and a family history of DSM-IV alcohol abuse / dependence.

The external validators used in this study will be used to address the 2 research questions exposed in the introduction:

1) Concerning the threshold model versus the linear model, the question is whether there is a significant jump in the odds ratio or beta coefficients of the validating variables at a certain threshold or whether the values increase linearly.

2) Concerning the weighting of the symptoms, the question is which symptoms are associated with the most validating variables and are certain symptoms associated more frequently with certain categories of validating variables than others (for example, are certain symptoms more predictive of treatment-seeking variables than others)?

2.4 Statistical analyses

The statistical analyses were conducted in three stages. First, the subjects were divided into 10 groups: those with 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9-11 DSM-5 (proposed revision) AUD symptoms. In order to test the associations between these groups and the categorical validator variables, overall χ^2 tests were applied (using Fisher's exact tests when the expected frequency was smaller than 5), and in order to test the associations between the groups and continuous validator variables, overall analyses of variance tests (ANOVA) were applied. These tests were conducted in order to assess the adequacy of chosen validators. The validators with insignificant results were excluded due to their overall inability to distinguish individuals with different numbers of symptoms.

The second stage of the analysis consisted of creating 10 different models: model 1 had a threshold at 1 symptom, model 2 a threshold at 2 symptoms, model 3 one at 3 symptoms, etc. and the last model, model 10 was a linear model showing increasing or decreasing parameter estimates of validator variables with increasing number of symptoms and no threshold. The validation of model 10 over the other models would support a dimensional nature of the AUD diagnosis whereas the validation of the threshold models would validate the categorical nature of the AUD diagnosis. In

order to test for the best fitting model, the Akaike Information Criterion (AIC) (Akaike 1978) was applied. The AIC measures the best fit between the predicted model and the observed values. The model with the lowest AIC value is indicative of the best fitting model.

The third stage of the analysis was conducted to evaluate the effect of each individual symptom on the AUD diagnosis. To do so, linear regression models with all possible subsets of symptoms were fitted for each continuous validator variable and logistic regression models with all possible subsets of symptoms were fitted for dichotomous validator variables. For each validator variable, the model with the best subset of symptoms (i.e the model that offers the best fit to observed data) was identified by choosing the model with the smallest AIC value. It is to be noted that some of the best fitting models may contain non significant beta coefficients or odds ratios. These parameters are not to be removed from the model because they increase the overall model's fit to data but are not to be considered predictive of the validator.

The third analysis allowed for the identification of the most prominent symptoms explaining each respective validator variable. Symptoms that were present in the highest number of models, thus predicting the highest number of validator variables were considered to be the most severe symptoms for an AUD diagnosis.

3.0 Results

3.1. Demographic characteristics and discriminant values of the validators

Table 1 shows that the subjects, grouped according to the number of AUD symptoms, did not differ by age and Swiss citizenship although they differed by sex and SES.

Concerning the discriminant values of the validators, overall distribution differences were found for the following validator variables: onset of regular consumption, average alcohol consumption, maximum consumption in 24 hours, the three types of GAF scores, the presence of suicide attempts, treatment-seeking behaviors, comorbidity including bipolar disorder, social phobia, post-traumatic stress syndrome (PTSD), antisocial personality, attention deficit and hyperactivity disorder (ADHD),

marijuana, cocaine, sedatives and narcotics abuse / dependence and finally a family history of abuse / dependence. These validators could be considered as pertinent variables for further analyses as they showed discriminant values when associated to the number of AUD symptoms (Table 1). The remaining validators, not showing discriminant values, were the following: major depressive disorder (MDD), dysthymia, general anxiety disorder (GAD), panic attacks, agoraphobia and obsessive-compulsive disorder (OCD). These validators were excluded from further analyses.

According to Table 1, general trends can be observed in the expected direction including worse outcomes on the course variables and more treatment seeking behaviours in the groups with a higher versus a lower number of symptoms.