HHS Public Access

Author manuscript

Drug Alcohol Depend. Author manuscript; available in PMC 2016 March 01.

Published in final edited form as:

Drug Alcohol Depend. 2015 March 01; 148: 40–46. doi:10.1016/j.drugalcdep.2014.12.011.

The Alcohol Use Disorder and Associated Disabilities Interview Schedule-5 (AUDADIS-5): procedural validity of substance use disorders modules through clinical re-appraisal in a general population sample

Deborah S. Hasin[1,2,3,*], **Eliana Greenstein**[3], **Christina Aivadyan**[3], **Malka Stohl**[3], **Efrat Aharonovich**[1,3], **Tulshi Saha**[4], **Rise Goldstein**[4], **Edward V. Nunes**[1,3], **Jeesun Jung**[4], **Haitao Zhang**[4], and **Bridget F. Grant**[4]

[1]Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA

[2]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York 10032, USA

[3]New York State Psychiatric Institute, New York, New York 10032, USA

[4]Laboratory of Epidemiology and Biometry, National Institute on Alcohol Abuse and Alcoholism, Bethesda, Maryland 20892, USA

Abstract

Background—The purpose of this study was to assess the procedural validity of the substance disorder modules of the lay-administered Alcohol Use Disorder and Associated Disabilities Interview Schedule, Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) Version (AUDADIS-5) through clinician re-appraisal re-interviews.

[*]**Correspondence**, Deborah S. Hasin, Department of Psychiatry, Columbia University Medical Center, 1051 Riverside Drive #123, New York, NY 10032. Phone: 1-646-774-7909, Fax: 1-646-774-7920; deborah.hasin@gmail.com.

Author Disclosures
No authors have any relevant financial interests.

Contributors
Drs. Saha, Goldstein, Jung, Zhang and Grant collected, cleaned and analyzed the data and critically reviewed drafts of the manuscript. Dr. Hasin collected the data, wrote, and revised drafts of the manuscript. Ms. Greenstein, and Ms. Aivadyan collected, cleaned and analyzed the data and critically reviewed drafts of the manuscript. Ms. Stohl analyzed the data and critically reviewed drafts of the manuscript. Drs. Aharonovich and Nunes consulted to the data collection and critically reviewed drafts of the manuscript. All authors have read and approved of submission of this version of the manuscript.

Conflict of Interest
No conflict declared.

Methods—The study employed a test-retest design among 712 respondents from the National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III). A clinician-administered, semi-structured interview, the Psychiatric Research Interview for Substance and Mental Disorders, DSM-5 version (PRISM-5) was used as the re-appraisal. Kappa coeffients indicated concordance of the AUDADIS-5 and PRISM-5 for DSM-5 substance use disorder diagnoses, while intraclass correlation coefficients (ICC) indicated concordance on dimensional scales indicating the DSM-5 criteria count for each disorder.

Results—With few exceptions, concordance of the AUDADIS-5 and the PRISM-5 for DSM-5 diagnoses of substance use disorders ranged from fair to good (κ=0.40–0.72). Concordance on dimensional scales was excellent (ICC≥0.75) for the majority of DSM-5 SUD diagnoses, and fair to good (ICC=0.43–0.72) for most of the rest.

Conclusions—As indicated by concordance with a semi-structured clinician-administered re-appraisal, the procedural validity of the AUDADIS-5 DSM-5 substance use disorder diagnoses found in this study indicates that these AUDADIS-5 diagnoses are useful tools in epidemiologic studies. The considerably stronger concordance of the AUDADIS-5 and PRISM-5 dimensional DSM-5 SUD measures supports a current movement to place more emphasis on dimensional measures of psychopathology, and suggests that such measures may be more informative than binary diagnoses for research, and possibly for clinical purposes as well.

Keywords

alcohol use disorder; substance use disorder; reliability; validity; Alcohol Use Disorder and Associated Disabilities Interview Schedule-5; Psychiatric Research Interview for Substance and Mental Disorders-5

1. INTRODUCTION

Relatively little is known about the procedural validity of structured diagnostic interviews used to produce epidemiologic data on substance use disorders (SUD). The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS) is a fully structured, computer-assisted diagnostic interview designed for trained lay interviewers (Grant et al., 2001). The AUDADIS-IV, which assessed disorders according to DSM-IV, was the measurement instrument in two U.S. national surveys sponsored by the National Institute on Alcohol Abuse and Alcoholism (NIAAA), the National Longitudinal Alcohol Epidemiologic Survey (NLAES; 1991–1992) and the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC, Wave 1 (2001–2002) and Wave 2 (2004–2005) (Compton et al., 2004; Grant et al., 2004a, 2009, 2004b). These surveys produced a rich epidemiologic literature on alcohol, drugs and related conditions (Hasin, In Press). The AUDADIS-IV was also used in clinical (Hasin et al., 2013a), high-risk (Hasin et al., 2007), and genetic studies (Meyers et al., 2013).

In 2012–2013, NIAAA fielded the National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III; Grant, 2014). The AUDADIS-5 (Grant et al., 2011), which assesses diagnoses defined by DSM-5 criteria, is the NESARC-III instrument.

In test-retest reliability studies of AUDADIS-IV SUD diagnoses in untreated (Grant et al., 2003, 1995; Ruan et al., 2008) and treated (Canino et al., 1999; Hasin et al., 1997) U.S. samples and in an international study (Chatterji et al., 1997; Vrasti et al., 1998), most reliability coefficients (*kappa or* κ) for SUDs were good to excellent (κ≥.60), and reliability of dimensional SUD measures (counts of diagnostic criteria) generally exceeded binary diagnoses (Grant et al., 1995; Hasin et al., 1997). Recently, AUDADIS-5 DSM-5 SUDs were shown to have fair to excellent test-retest reliability (κ=.41–.87) (Grant et al., In Press).

Procedural validity through clinician re-appraisal constitutes an important method of validating lay-administered diagnostic interviews. In the international study, AUDADIS-IV substance dependence generally had fair to very good concordance (Cottler et al., 1997) with the clinician-administered Structured Clinical Assessment for Neuropsychiatry (κ=.41–.67) (Easton et al., 1997). In Puerto Rican primary care patients, AUDADIS-IV dependence had fair to very good concordance with psychiatrist re-interviews (Canino et al., 1999).

Given the changes in DSM-5 SUD (Hasin et al., 2013b), information is needed on the procedural validity of AUDADIS-5 SUD measures. Procedural validity is important to the interpretation of NESARC-III and other studies using the AUDADIS-5 because it reflects the extent to which diagnoses from the AUDADIS-5 correspond to diagnoses based on evaluations by experienced clinicians. A subset of NESARC-III participants underwent a clinical re-appraisal using the Psychiatric Research Interview for Substance and Mental Disorders, DSM-5 version, conducted by mental health clinicians experienced in the evaluation and treatment of substance use and commonly co-occurring disorders (Hasin et al., 2011). The PRISM is a semi-structured diagnostic interview designed to address measurement issues in heavy drinkers and drug users (Hasin et al., 2006). Here, we examine the procedural validity of AUDADIS-5 DSM-5 diagnoses of alcohol, drug and tobacco disorders. Given the increasing attention to dimensional measures of psychopathology (Hasin et al., 2013b; Regier et al., 2012), we also examine the procedural validity of DSM-5 dimensional alcohol, drug and tobacco disorder measures. To our knowledge, procedural validation of such dimensional measures through clinical reappraisal has not been done previously.

2. MATERIAL AND METHODS

2.1 Sample and procedures

Participants were selected from the NESARC-III sample. NESARC-III included the non-institutionalized U.S. civilian population 18 years or older, including persons in households or group quarters, e.g., group homes; worker dormitories randomly selected via multistage probability sampling (Grant, 2014), with higher selection probabilities for Hispanics, Blacks, and Asians. The NESARC-III response rate was 60.1%, comparable to most current U.S. national health surveys (Division of Health Interview Statistics),(Centers for Disease Control and Prevention)).

Each NESARC-III participant completed a face-to-face AUDADIS-5 interview (N=36,309); of these, 25,769 consented to participate in a second interview. From these, potential procedural validity participants were selected using an algorithm designed to increase

selection of respondents with greater psychopathy. Screening questions used to route respondents into modules covering alcohol, drug and tobacco use disorders and each mood, anxiety and trauma and stress-related disorder was coded as "positive" or "negative." Respondents with the greatest number of positive screening questions and therefore completed modules were given higher probabilities of selection into the validity study. Using this algorithm, a validity sample was then selected each week from June, 2012 to July, 2013 in order to manage the caseload of validation interviews. Potential procedural validity participants were selected within the Eastern Time Zone to facilitate the telephone re-appraisal interviews that were conducted in New York City. Institutional Review Boards at the National Institutes of Health and New York State Psychiatric Institute approved all procedure validity study procedures; all respondents gave informed consent to participate.

For the procedural validity study, 777 respondents were selected. After excluding individuals unable to participate (4 too ill; 3 dislocated due to Hurricane Sandy), the response rate was 92.5% (712 participants out of 770).

2.2 Procedural validity re-interview design

PRISM-5 clinical re-appraisals were conducted on the telephone. Telephone administration is advantageous because it permits a centralized, closely supervised team of clinical interviewers to conduct interviews throughout a large geographic area, which is not feasible with face-to-face clinical reappraisal (Kessler et al., 2009). The mean test-retest interval between AUDADIS and PRISM interviews was 10.54 days (s.e.=0.27). With consent, PRISM-5 interviews were recorded for quality assurance purposes; 700 participants (98.3%) consented to recording the interview.

All members of the PRISM-5 team were blind to the initial AUDADIS-5 interviews, including values in the selection algorithm. Prior to starting the PRISM-5 interview, the interviewer told respondents that the purpose of the re-interview was to help understand the quality of the initial survey questions, and that he/she did not have information from the first interview. Participants were instructed to respond with "whatever answer seems right to you today. Don't try to make your answers the same as last time, or different - just give the answer that seems right to you now."

2.3 Diagnostic assessment

2.3.1 AUDADIS-5 modules—The AUDADIS-5 DSM-5 SUD diagnoses assessed in the present study included alcohol, cannabis, cocaine, hallucinogens, heroin, other opioids, sedatives/tranquilizers, other stimulants, and tobacco. While DSM-5 was finalized in 2012, the SUD criteria were known in 2011, and incorporated into the AUDADIS-5.

2.3.2 PRISM-5: the validation procedure—The clinician-administered re-interview was the Psychiatric Research Interview for Substance and Mental Disorders (PRISM), DSM-5 version. The PRISM is a semi-structured diagnostic interview initially designed to assess disorders according to DSM-IV criteria (Hasin et al., 2006, 1996). "Semi"-structured indicates that all initial PRISM probes for symptoms and criteria are asked as written. However, in contrast to fully structured interviews, e.g., AUDADIS-5, PRISM interviewers

add their own unstructured follow-up probing, informed by their clinical expertise, in order to obtain more information and clarify unclear responses. The DSM-IV version of the PRISM has shown excellent test-retest reliability and validity (Hasin et al., 2006; Torrens et al., 2004). The PRISM-5 is a computer-assisted interview.

The DSM-IV version of the PRISM covered all criteria for DSM-5 SUDs (Hasin et al., 2012). Therefore, adaptation of PRISM SUD modules for DSM-5 was straightforward, conducted under the supervision of D.S.H. and B.F.G., both DSM-5 SUD Workgroup members (Hasin et al., 2013b). DSM-5 diagnoses were produced from interview data via computer algorithms that operationalized the DSM-5 criteria.

All procedural validity participants received the PRISM-5 alcohol and drug modules. To reduce participant burden, remaining modules were allocated to two shortened versions of the PRISM-5, only one of which included tobacco use disorder. The two versions were randomly assigned to procedural validity respondents. Of the 712 procedural validity participants, 355 (49.9%) were assessed for tobacco use disorders.

2.3.3 PRISM-5 interviewers—Ten interviewers administered the PRISM-5. They had at least a master's degree in a clinical field and clinical experience with substance abuse and/or psychiatric patients (mean years of experience, 4.15, range, 2–14).

2.3.4 PRISM-5 training—PRISM-5 training began with self-study of a training manual with review questions, followed by 4 days of in-class training on clinical and technical aspects of the PRISM-5 and study procedures, security/confidentiality issues, and ethical standards. This combined didactic teaching and role-playing with trainers. Afterwards, trainees conducted recorded PRISM-5 interviews with volunteers role-playing study participants. A PRISM-5 trainer/supervisor rated trainee performance in the recorded interviews on structured quality assurance forms. PRISM-5 interviewers were certified after five recording were rated as satisfactory (allowing one replacement if necessary). All interviewers who collected data for the study successfully completed this process.

2.3.5 PRISM-5 supervision and quality assurance—Ongoing PRISM-5 supervision during the procedural validity study consisted of regular supervision meetings with the two PRISM-5 trainer/supervisors (E.G. and C.A.) and review of issues arising from review of recordings of PRISM-5 interviews.

Recordings from 214 randomly selected PRISM-5 interviews were reviewed and rated for quality assurance purposes. Of these, 107 were reviewed by the PRISM-5 trainer/supervisors, using the same structured form used to rate recordings of trainees. They used information from these reviews to give individual feedback to interviewers as needed, and to give feedback to all interviewers during group supervision meetings.

Additional quality assurance involved psychiatrist review of 107 PRISM-5 recordings by two addiction psychiatrists, each with over 10 years of clinical experience. After structured training on the PRISM-5, the two psychiatrists independently reviewed recordings of PRISM-5 interviews, observing responses to items as they listened to the interview, and

evaluating interviewer performance. Of these recorded interviews, 59 were also independently rated by one of the clinical supervisors. In 90% of these 59 cases, the psychiatrists and clinical supervisors agreed that all 27 SUD diagnoses (9 substances × 3 timeframes) derived from the PRISM-5 were correct. In 3 cases (5.1%), the psychiatrist viewed a SUD diagnosis as correct while the clinical supervisor did not; in 3 other cases (5.1%), the clinical supervisor viewed a SUD diagnosis as correct while the psychiatrist did not. In no instance did the psychiatrist and clinical supervisor agree that PRISM-5 SUD diagnoses were incorrect. The psychiatrist evaluations were provided to the supervisors and interviewers.

2.4 Statistical analyses

The present study used standard statistical methodology that has been used in prior test-retest and procedural validity studies of the AUDADIS. For dichotomous diagnoses, kappa (κ) was used as the concordance coefficient, defined as a measure of pairwise agreement corrected for chance (Fleiss, 1981). McNemar's test, a test for paired comparisons of binary variables, was used to determine if the prevalence of diagnoses differed between the AUDADIS-5 and the PRISM-5. For continuous measures, intraclass correlation coefficients (ICC) were used to examine the concordance of dimensional diagnostic criteria scales. Paired comparisons of differences in the AUDADIS-5 and PRISM-5 dimensional measures were tested with signed rank tests, which do not require meeting normality assumptions. The timeframes assessed for substance use disorders were past year, prior to past year and lifetime.

The validity design assumed that interviewers were randomly drawn from a larger population of interviewers. We therefore used a one-way random effects ANOVA model to derive intraclass correlation coefficients (Shrout and Fleiss, 1979). κ and ICC values share the same interpretation (Davis and Fleiss, 1982). κ and ICC values range from 1.00 (perfect agreement) to − 1.00 (total disagreement) with values of zero indicating agreement equivalent to chance. Excellent agreement is indicated by κ or ICC ≥0.75; fair to good agreement, κ=0.40 to 0.74; and poor agreement, κ<0.39 (Fleiss, 1981; Landis and Koch, 1977).

The N of this procedural validity study was large relative to prior studies (Canino et al., 1999; Cottler et al., 1997), and the selection algorithm enriched the sample in terms of disorder prevalence. Nevertheless, the prevalence of some past-year SUDs in the AUDADIS-5 or PRISM-5 was too low (<2%) to yield stable estimates. Consequently, κ for these diagnoses was not reported. This constraint did not apply to the ICCs, since dimensional criteria scales are assessed separately from diagnoses.

3. RESULTS

3.1 Sample characteristics

I In terms of respondent characteristics (Table 1), about two-thirds were non-Hispanic whites, slightly over half were female, most were 25–64 years old, about 39% were married or cohabiting, and slightly over half had some education past high school. There were no

differences between the procedural validity sample (n=712) and the remaining NESARC-III sample (n=35,597) on gender (χ^2=0.98, p=0.32) or education (χ^2=1.89, p=0.49). There were differences between these two samples on race-ethnicity (χ^2=118.8, p<0.001), age (t= −3.31, df=36,307, p<0.001) and employment status (χ^2=4.63, p=0.03), with the procedural validity sample containing a somewhat larger proportion of whites, younger, and employed participants.

3.2 Concordance on DSM-5 SUD diagnoses

The concordance between AUDADIS-5 and PRISM-5 diagnoses of DSM-5 SUDs is shown in Table 2. For past-year diagnoses, κ statistics indicated fair to good concordance, including κ≥0.60 for alcohol, cannabis, cocaine and tobacco, and κ=0.40 for opioids. For prior to the past year diagnoses, κ statistics indicated fair to good concordance (κ≥0.40) for all substances except hallucinogens and stimulants. For lifetime diagnoses, κ statistics indicated fair to good concordance (κ≥0.40) for all SUD diagnoses except stimulants.

Results of the McNemar tests indicated that of the five SUDs with sufficient prevalence to analyze, past-year prevalence was higher in AUDADIS-5 than in PRISM-5 for three, and did not differ for the other two. Prevalence of prior to past year diagnoses was significantly higher in AUDADIS-5 than in PRISM-5 for alcohol use disorder, significantly higher in PRISM-5 than AUDADIS-5 for cannabis and heroin use disorders and did not differ in the remaining disorders. Prevalence of lifetime disorders differed only for tobacco and alcohol use disorders, which was higher in AUDADIS-5 than in PRISM-5.

3.3 Concordance on DSM-5 SUD dimensional measures

For past-year dimensional DSM-5 SUD measures, ICCs (Table 3) indicated excellent AUDADIS-5/PRISM-5 concordance (ICC≥0.79) on alcohol, cannabis, cocaine, hallucinogen, heroin and tobacco use disorders, fair to good concordance on stimulant, tobacco and opioid disorders (ICC=0.0.43–0.68), and sedative use disorders the only exception with ICC<0.40. For the prior to past-year dimensional measures, ICCs indicated excellent concordance (ICC≥0.76) for alcohol, cannabis, cocaine, heroin, opioid, sedative and tobacco use disorders, and very good concordance for hallucinogen and stimulant use disorders (ICC=0.72). For the lifetime dimensional measures, ICCs indicated excellent concordance (ICC≥0.76) for alcohol, cannabis, cocaine, heroin, opioid, sedative and tobacco use disorders, and very good concordance for hallucinogen and stimulant use disorders (ICC=0.72). The signed rank tests showed that in most instances, the AUDADIS-5 and PRISM-5 did not differ significantly on the number of criteria rated positive. Exceptions to this were found for alcohol and tobacco use disorders across timeframes (higher values in the AUDADIS-5), past-year opioid disorder (higher in the AUDADIS-5), prior to past-year heroin and sedative use disorders (higher in the PRISM-5) and lifetime cocaine and sedative use disorders (higher in the PRISM-5).

4. DISCUSSION

In a large, rigorous study in the general population, AUDADIS-5 diagnoses of DSM-5 substance use disorders demonstrated fair to very good concordance with PRISM-5 clinician

re-appraisals for current, past and lifetime diagnoses of the DSM-5 substance use disorders examined. Concordance levels on diagnoses in the present study are comparable to those from prior procedural validity studies of the AUDADIS-IV (Canino et al., 1999; Cottler et al., 1997). This suggests that overall, results from the NESARC-III, based on AUDADIS-5 diagnoses, can be considered valid when interpreting findings of analyses on the full NESARC-III sample.

Importantly, dimensional representations of these disorders showed greater AUDADIS-5/PRISM-5 concordance than binary diagnoses, with excellent concordance for alcohol, cannabis, cocaine, heroin and hallucinogens, and fair to good concordance on most remaining substances. The stronger validity of the dimensional variables attests to the value of a dimensional approach in the measurement of SUDs. Although diagnoses of substance disorders are useful for communication among clinicians, researchers and for policy decisions (Hasin et al., 2013b), results for the dimensional measures were likely obtained because they do not impose arbitrary thresholds (Hasin et al., 2013b), and thus are more statistically informative than categorical measures. The higher procedural validity of the dimensional SUD measures supports the growing movement towards dimensional assessment of psychopathology (Faraone, 2013; Fazzino et al., 2014), which facilitates indicating substance disorder severity and change over time. The dimensional approach is also useful clinically, for example, mild alcohol use disorders require different clinical management than moderate or severe disorders. The greater validity of the DSM-5 SUD dimensional measures is also relevant to NESARC-III phenotype development for genetic analyses of DNA collected from 24,381 of the participants.

While the AUDADIS-5 was administered in person, the PRISM-5 was administered by telephone. PRISM-5 telephone administration facilitates an efficient, cost-effective centralized team of interviewers conducting closely-supervised interviews throughout a large geographic region (Kessler et al., 2009). The present study leaves unanswered whether concordance would be higher if administration mode had been constant in the two interviews. Nevertheless, telephone vs. in-person mode of administration has been shown repeatedly not to influence findings, including for alcohol harms (Midanik and Greenfield, 2003), substance use disorders (Kessler et al., 2009; Sobin et al., 1993), and other symptomatology (Aziz and Kenford, 2004; Wells et al., 1988). As such, the literature suggests that discordant cases were likely due to other factors.

Of the twenty-three tests of whether prevalence of binary diagnoses differed between the AUDADIS-5 and the PRISM-5, only eight were significant, with six showing higher prevalence in the AUDADIS-5 and two showing lower prevalence, while in 27 tests of whether the corresponding dimensional means differed, nine were significant, with alcohol and tobacco use disorders higher in the AUDADIS-5, and instances of heroin, cocaine and sedative use disorders higher in the PRISM-5. While future efforts should address whether specific criteria led to the differences that were found, on the whole, results did not suggest that AUDADIS-5 diagnoses or dimensional measures were systematically biased by greater inclusiveness than clinician evaluations.

Concordance between categorical measures is impacted to a greater degree by mild or borderline cases than by more severe cases (Helzer et al., 1985), and the diagnostic threshold/borderline for DSM-5 SUD is lower (hence, milder) than it was for DSM-IV dependence. The impact of varying the diagnostic threshold for DSM-5 substance use disorders on procedural validity across different substances warrants investigation in future studies.

A limitation in interpreting study findings is that the AUDADIS-5 was always the first interview and PRISM-5 the second. This was unavoidable given the study design (clinical re-appraisals of participants in a national survey). Future studies should balance interview order to allow differentiating between differences due to the procedures and differences due to any order effects.

Although the word “validation” is used to characterize the results reported here, no instrument in psychiatry provides a perfect gold standard of true DSM disorders. The PRISM has shown very good to excellent but not perfect reliability for substance dependence diagnoses (Hasin et al., 2006; Torrens et al., 2004), and imperfect reliability may attenuate the ability to show associations with other variables, including diagnoses from fully-structured diagnostic interviews. Thus, while clinician re-appraisal is an important aspect of validation, especially to establish credibility among clinical investigators, other validation strategies are also needed, including those involving antecedent, concurrent, and prospective validators. Such studies are outside the scope of the present report, but should be conducted in the future.

Study limitations are noted. First, individuals without phone access were excluded. However, very few in the U.S. (even those in poverty) have no phone access (McInnes et al., 2013; United States Census Bureau, 2011), so this feature does not greatly limit generalizeability. Second, relative to the full NESARC-III sample, the validity subsample had more white and young respondents. These groups have higher prevalence of many common disorders so these differences were expected. Future studies should focus on older participants, and those from race/ethnic minority groups. Third, the validity study did not target individuals in treatment. Conducting the study in the general population was of prime importance, given the role of the AUDADIS-5 in NESARC-III. However, some validity studies of the AUDADIS-IV were conducted in clinical settings (Canino et al., 1999; Cottler et al., 1997). For complete information on AUDADIS-5 procedural validity, clinician re-appraisal studies should also be conducted in patients. Fourth, the present general population sample had too few cases of some disorders to analyze. To determine the validity of remaining AUDADIS-5 SUD diagnoses, studies are needed in higher-prevalence samples. Fifth, the present study did not examine individual items or criteria. Such examination merits attention in future studies. Sixth, multiple factors may impact validity, e.g., participant characteristics. These warrant future investigation. Finally, we did not use biological measures as validators. A neurobiological test indicating addiction would be a valuable validator for AUDADIS-5 SUD diagnoses, but such a test does not yet exist. Biological indicators of substance use have many shortcomings for the purposes of the present study, including time windows for detection that are too short for most SUDs of interest in the NESARC-III.

Study strengths are also noted. First, the large sample provided stable concordance estimates even for many relatively low-prevalence conditions. Second, a high response rate was obtained among those selected for re-interview. Third, the sample was selected from the general population, providing important information on the main type of study participant for whom the AUDADIS-5 was designed. Fourth, the PRISM-5 team was entirely blind to the results of the AUDADIS-5 interviews and values in the selection algorithm. This eliminated potential sources of bias, including influences of the AUDADIS-5 interview on PRISM-5 interviewer questions and probes. Fifth, the study employed a test/re-test administration interval that reduced the influence of new illness onset on concordance. Sixth, the semi-structured clinician re-appraisal procedure, the PRISM-5, has many strengths. It overcomes measurement problems in heavy drinkers and drug users (Hasin et al., 2006), has been used extensively to provide results in such samples (Drake et al., 2011; Hasin et al., 2002; Nunes et al., 2006; Torrens et al., 2011), provided information to the DSM-5 SUD workgroup that influenced decisions on DSM-5 (Hasin et al., 2012), and has been used as the gold standard in measurement studies of other instruments (Cuenca-Royo et al., 2012; Mestre-Pinto et al., 2014). Finally, PRISM-5 interviewers all had advanced clinical degrees and experience working with substance abusing populations. PRISM-5 training, supervision and quality control procedures were standardized and rigorous. All of these strengths contributed to a clinical re-appraisal study with a rigorous design for indicating the procedural validity of AUDADIS-5 SUD measures.

In view of the widespread importance and influence of national epidemiologic studies for research, policy and clinical work, understanding the procedural validity of the AUDADIS-5 SUD measures is crucial. While the procedural validity of the AUDADIS-IV was known, the changes in DSM-5 necessitated a new study. Future analyses of these data will focus on multivariate analyses of factors affecting validity to further understand the new DSM-5 measures of substance use disorders in the AUDADIS-5. Meanwhile, the concordance of AUDADIS-5 DSM-5 SUD diagnoses and dimensional measures with clinician re-appraisals suggests that the AUDADIS-5 is a useful diagnostic tool for a variety of research purposes. Most importantly, the AUDADIS-5 has demonstrated procedural validity in this general population sample, the target sample for which it was primarily designed.

Acknowledgements

We acknowledge the work of Shabnam Shakibaie Smith, M.D. and Anatoliy Vasilov, M.D., for their assistance in the quality assurance of the PRISM-5 interviews for this study.

Role of Funding Source

Support is acknowledged from the National Institutes of Health, Contract HHSN275200900007C, K05AA014223 (Hasin), K24DA022412 (Nunes) and the New York State Psychiatric Institute (Hasin, Nunes). The National Epidemiologic Survey on Alcohol and Related Conditions-III was sponsored by the National Institute on Alcohol Abuse and Alcoholism and funded, in part, by the Intramural Program, NIAAA, National Institutes of Health.

REFERENCES

Aziz MA, Kenford S. Comparability of telephone and face-to-face interviews in assessing patients with posttraumatic stress disorder. J. Psychiatr. Pract. 2004; 10:307–313. [PubMed: 15361745]

Canino G, et al. The Spanish Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability and concordance with clinical diagnoses in a Hispanic population. J. Stud. Alcohol. 1999; 60:790–799. [PubMed: 10606491]

Centers for Disease Control and Prevention. U.S. Department of Health and Human Services. 2013 Jun.

Chatterji S, Saunders JB, Vrasti R, Grant BF, Hasin D, Mager D. Reliability of the alcohol and drug modules of the Alcohol Use Disorder and Associated Disabilities Interview Schedule--Alcohol/Drug-Revised (AUDADIS-ADR): an international comparison. Drug Alcohol Depend. 1997; 47:171–185. [PubMed: 9306043]

Compton WM, Grant BF, Colliver JD, Glantz MD, Stinson FS. Prevalence of marijuana use disorders in the United States: 1991–1992 and 2001–2002. JAMA. 2004; 291:2114–2121. [PubMed: 15126440]

Cottler LB, Grant BF, Mavreas V>, Pull C, Hasin D, Compton WM, Rubio-Stipec M, Mager D. Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. Drug Alcohol Depend. 1997; 47:195–205. [PubMed: 9306045]

Cuenca-Royo AM, Sanchez-Niubo A>, Forero CG, Torrens M, Suelves JM, Domingo-Salvany A. Psychometric properties of the CAST and SDS scales in young adult cannabis users. Addict. Behav. 2012; 37:709–715. [PubMed: 22386300]

Davis M, Fleiss JL. Measuring agreement for multinomial data. Biometrics. 1982; 38:1047–1051.

Division of Health Interview Statistics. National Health Interview Survey (NHIS) Public Use Data Release. Hyattsville, Maryland: National Center for Health Statistics; 2012.

Drake RE, Caton CL, Xie H, Hsu E, Gorrochurn P, Samet S, Hasin DS. A prospective 2-year study of emergency department patients with early-phase primary psychosis or substance-induced psychosis. Am. J. Psychiatry. 2011; 168:742–748. [PubMed: 21454918]

Easton C, Meza E, Mager D>, Ulug B, Kilic C, Gogus A, Babor TF. Test-retest reliability of the alcohol and drug use disorder sections of the schedules for clinical assessment in neuropsychiatry (SCAN). Drug Alcohol Depend. 1997; 47:187–194. [PubMed: 9306044]

Faraone SV. Real progress in molecular psychiatric genetics. J. Am. Acad. Child. Adolesc. Psychiatry. 2013; 52:1006–1008. [PubMed: 24074467]

Fazzino TL, Rose GL, Burt KB, Helzer JE. Comparison of categorical alcohol dependence versus a dimensional measure for predicting weekly alcohol use in heavy drinkers. Drug Alcohol Depend. 2014; 136:121–126. [PubMed: 24485060]

Fleiss, JL. Statistical Methods For Rates And Proportions. New York: John Wiley and Sons Inc.; 1981.

Grant, BF. Source And Accuracy Statement: National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 2014.

Grant, BF., Dawson, DA., Hasin, DS. The Alcohol Use Disorder and Associated Disabilities Interview Schedule Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition Version (AUDADIS-IV). Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2001.

Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering r. The alcohol use disorder and associated disabilities interview schedule-IV (AUDADIS-IV): reliability of new psychiatric diagnostic modules and risk factors in a general population sample. Drug Alcohol Depend. 2003; 71:7–16. [PubMed: 12821201]

Grant BF, Dawson DA, Stinson FS, Chou SP, Dufour MC>, Pickering RP. The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991–1992 and 2001–2002. Drug. Alcohol. Depend. 2004a; 74:223–234. [PubMed: 15194200]

Grant, BF., et al. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition Version (AUDADIS-5). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 2011.

Grant BF, et al. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-5 (AUDADIS-5): reliability of substance use and psychiatric disorder modules in a general population sample. In Press.

Grant BF, Goldstein RB, Chou SP, Huang B, Stinson FS, Dawson dA, Saha D, Smith SM, Pulay AJ, Pickering RP, Ruan WJ, Compton WM. Sociodemographic and psychopathologic predictors of

first incidence of DSM-IV substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. Mol. Psychiatry. 2009; 14:1051–1066. [PubMed: 18427559]

Grant BF, Harford TC, Dawson DA, Chou PS, Pickering RP. The Alcohol Use Disorder and Associated Disabilities Interview schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. Drug Alcohol Depend. 1995; 39:37–44. [PubMed: 7587973]

Grant BF, Stinson FS, Dawson DA, Chou SP, Dufour MC, Compton W, Pickering RP, Kaplan K. Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch. Gen. Psychiatry. 2004b; 61:807–816. [PubMed: 15289279]

Hasin D, Carpenter KM, McCloud S, Smith M, Grant BF. The alcohol use disorder and associated disabilities interview schedule (AUDADIS): reliability of alcohol and drug modules in a clinical sample. Drug Alcohol Depend. 1997; 44:133–141. [PubMed: 9088785]

Hasin D, Keyes KM, Hatzenbuehler ML, Aharonovich EA, Alderson D. Alcohol consumption and post-traumatic stress after terrorism exposure: effects of physical proximity, loss, and previously-assessed psychiatric history. Am. J. Public. Health. 2007; 97:2268–2275. [PubMed: 17971553]

Hasin D, Liu X, Nunes E, McCloud S, Samet S, Endicott J. Effects of major depression on remission and relapse of substance dependence. Arch. Gen. Psychiatry. 2002; 59:375–380. [PubMed: 11926938]

Hasin D, Samet S, Nunes E, Meydan J, Matseoane K, Waxman R. Diagnosis of comorbid psychiatric disorders in substance users assessed with the Psychiatric Research Interview for Substance and Mental Disorders for DSM-IV. Am. J. Psychiatry. 2006; 163:689–696. [PubMed: 16585445]

Hasin DS. The NESARC, Waves 1 and 2: a decade of research findings. Addiction. In Press.

Hasin DS, Aharonovich E, O'Leary A, Greenstein E, Pavlicova M, Arunajadai S, Waxman R, Wainberg M, Helzer J, Johnston B. Reducing heavy drinking in HIV primary care: a randomized trial of brief intervention, with and without technological enhancement. Addiction. 2013a; 108:1230–1240. [PubMed: 23432593]

Hasin, DS., et al. Psychiatric Research Interview for Substance Use and Mental Disorders, Diagnostic and Statistical Manual of Mental Disorders, Fifth Edtion (PRISM-5) version. New York, NY: Columbia University, Department of Psychiatry; 2011.

Hasin DS, Fenton MC, Beseler C, Park JY, Wall MM. Analyses related to the development of DSM-5 criteria for substance use related disorders: 2. Proposed DSM-5 criteria for alcohol, cannabis, cocaine and heroin disorders in 663 substance abuse patients. Drug. Alcohol. Depend. 2012; 122:28–37. [PubMed: 21963333]

Hasin DS, O'Brien CP, Auriacombe M, Borges G, Bucholz K, Budney A>, Compton WM, Crowley T, Ling W, Petry NM, Schuckit M, Grant BF. DSM-5 criteria for substance use disorders: recommendations and rationale. Am. J. Psychiatry. 2013b; 170:834–851. [PubMed: 23903334]

Hasin DS, Trautman KD, Miele GM, Samet S, Smith M, Endicott J. Psychiatric Research Interview for Substance and Mental Disorders (PRISM): reliability for substance abusers. Am. J. Psychiatry. 1996; 153:1195–1201. [PubMed: 8780425]

Helzer JE, Robins LN, McEvoy LT, Spitznagel EL, Stoltzman RK, Farmer A, Brockington IF. A comparison of clinical and diagnostic interview schedule diagnoses. Physician reexamination of lay-interviewed cases in the general population. Arch. Gen. Psychiatry. 1985; 42:657–666. [PubMed: 4015307]

Kessler RC, Avenevoli S, Green J, Gruber MJ, Guyer M, He Y, Jin R, Kaufman J, Sampson NA, Zaslavsky AM. National comorbidity survey replication adolescent supplement (NCS-A): III. Concordance of DSM-IV/CIDI diagnoses with clinical reassessments. J. Am. Acad. Child. Adolesc. Psychiatry. 2009; 48:386–399. [PubMed: 19252450]

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

McInnes DK, Li AE, Hogan TP. Opportunities for engaging low-income, vulnerable populations in health care: a systematic review of homeless persons' access to and use of information technologies. Am. J. Public Health. 2013; 103(Suppl. 2):e11–e24. [PubMed: 24148036]

Mestre-Pinto JI, Domingo-Salvany A, Martin-Santos R, Torrens M. PsyCoBarcelona Group. Dual diagnosis screening interview to identify psychiatric comorbidity in substance users: development and validation of a brief instrument. Eur. Addict. Res. 2014; 20:41–48. [PubMed: 23921461]

Meyers JL, Shmulewitz D, Wall MM, Keyes KM, Aharonovich E, Spivak B, Weizman A, Frisch A, Edenberg HJ, Gelernter J, Grant BF, Hasin D. Childhood adversity moderates the effect of ADH1B on risk for alcohol-related phenotypes in Jewish Israeli drinkers. Addict. Biol. 2013 Epub ahead of print.

Midanik LT, Greenfield TK. Telephone versus in-person interviews for alcohol use: results of the 2000 National Alcohol Survey. Drug Alcohol Depend. 2003; 72:209–214. [PubMed: 14643937]

Nunes EV, Liu X, Samet S, Matseoane K, Hasin D. Independent versus substance-induced major depressive disorder in substance-dependent patients: observational study of course during follow-up. J. Clin. Psychiatry. 2006; 67:1561–1567. [PubMed: 17107247]

Regier DA, Kuhl EA, Narrow WE, Kupfer DJ. Research planning for the future of psychiatric diagnosis. Eur. Psychiatry. 2012; 27:553–556. [PubMed: 21676595]

Ruan WJ, Goldstein RB, Chou SP, Smith SM, Saha TD, Pickering RP, Dawson DA, Huang B, Stinson FS, Grant BF. The alcohol use disorder and associated disabilities interview schedule-IV (AUDADIS-IV): reliability of new psychiatric diagnostic modules and risk factors in a general population sample. Drug Alcohol Depend. 2008; 92:27–36. [PubMed: 17706375]

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 1979; 86:420–428. [PubMed: 18839484]

Sobin C, Weissman MM, Goldstein RB, Adams P, Wickramaratne P, Warner V, Lish JD. Diagnostic interviewing for family studies: comparing telephone vs. face-to-face methods for the diagnosis of lifetime psychiatric disorders. Psychiatr. Genet. 1993; 3:227–233.

Torrens M, Gilchrist G, Domingo-Salvany A. psyCoBarcelona Group. Psychiatric comorbidity in illicit drug users: substance-induced versus independent disorders. Drug Alcohol Depend. 2011; 113:147–156. [PubMed: 20801586]

Torrens M, Serrano D, Astais M, Perez-Dopminguez G, Martin-Santos R. Diagnosing comorbid psychiatric disorders in substance abusers: validity of the Spanish versions of the Psychiatric Research Interview for Substance and Mental Disorders and the Structured Clinical Interview for DSM-IV. Am. J. Psychiatry. 2004; 161:1231–1237. [PubMed: 15229056]

United States Census Bureau. [Accessed August 5, 2014] Extended Measures of Well-being: Living Conditions in the United States, 2011. 2011. http://www.census.gov/hhes/well-being/publications/extended-11.html. accessed on

Vrasti R, Grant BF, Chatterji S, Ustun BT, Mager D>, Olteanu I, Badoi M. Reliability of the Romanian version of the alcohol module of the WHO Alcohol Use Disorder and Associated Disabilities: Interview Schedule --Alcohol/Drug-Revised. Eur. Addict. Res. 1998; 4:144–149. [PubMed: 9852366]

Wells KB, Burnam MA, Leake B, Robins LN. Agreement between face-to-face and telephone-administered versions of the depression section of the NIMH Diagnostic Interview Schedule. J. Psychiatr. Res. 1988; 22:207–220. [PubMed: 3225790]

Highlights

- AUDADIS-5 DSM-5 Substance Diagnoses were compared to PRISM-5 clinician re-evaluations
- AUDADIS-5/PRISM-5 concordance on DSM-5 SUD diagnoses ranged from fair to good
- AUDADIS-5/PRISM-5 concordance on SUD dimensional scales was generally excellent
- AUDADIS-5 SUD diagnoses and dimensional measures are useful measurement tools

Table 1

Sociodemographic characteristics of reliability study respondents (n=712)

Sociodemographic characteristics	N	% (SE)
Sex		
Male	324	45.5 (1.87)
Female	388	54.5 (1.87)
Age		
18–29	81	11.4 (1.19)
30–44	289	40.6 (1.84)
45–64	284	39.9 (1.84)
65 +	58	8.1 (1.03)
Race-ethnicity		
White	483	67.8 (1.75)
Black	160	22.5 (1.57)
American Indian/Alaska native	15	2.1 (0.54)
Asian/Pacific Islander	5	0.7 (0.31)
Hispanic	49	6.9 (0.95)
Education		
Less than high school	103	14.5 (1.32)
High school	206	28.9 (1.70)
Some college or higher	403	56.6 (1.86)

Note: SE = Standard error of percentage

Table 2

Concordance of AUDADIS-5 and PRISM-5 DSM-5 substance use disorder diagnoses N=712**

Disorder	A (95% CI)	Prevalence (%) AUDADIS-5	Prevalence (%) PRISM-5	McNemar's Test, p-values
Past year				
Alcohol	**0.62** (0.56–0.68)	33.0	25.3	<.01
Cannabis	**0.60** (0.51–0.70)	10.5	10.0	0.58
Cocaine	**0.66** (0.48–0.83)	03.2	02.3	0.05
Hallucinogen	_ [a]	_ [a]	_ [a]	_ [a]
Heroin	_ [a]	_ [a]	_ [a]	_ [a]
Opioids	**0.40** (0.26–0.56)	06.3	03.9	>0.01
Sedative	_ [a]	_ [a]	_ [a]	_ [a]
Stimulant	_ [a]	_ [a]	_ [a]	_ [a]
Tobacco	**0.68** (0.61–0.75)	62.5	47.6	<.01
Prior to the past year				
Alcohol	**0.42** (0.35–0.49)	57.0	47.5	<.01
Cannabis	**0.45** (0.37–0.53)	18.8	22.8	0.01
Cocaine	**0.54** (0.44–0.64)	10.5	11.8	0.26
Hallucinogen	**0.39** (0.20–0.59)	02.3	03.2	0.14
Heroin	**0.50** (0.31–0.69)	02.1	03.4	0.04
Opioids	**0.50** (0.40–0.62)	08.7	09.1	0.69
Sedative	**0.49** (0.36–0.63)	05.9	05.9	1.00
Stimulant	**0.35** (0.20–0.51)	03.5	03.3	0.40
Tobacco	**0.54** (0.45–0.63)	53.0	53.8	0.74
Lifetime				
Alcohol	**0.49** (0.43–0.55)	65.5	55.8	<.01
Cannabis	**0.51** (0.43–0.58)	22.9	25.1	0.16
Cocaine	**0.58** (0.48–0.67)	11.9	12.2	0.80
Hallucinogen	**0.44** (0.25–0.62)	02.8	03.4	0.41
Heroin	**0.56** (0.38–0.74)	02.4	03.5	0.06
Opioids	**0.49** (0.36–0.61)	11.4	10.8	0.64
Sedative	**0.48** (0.35–0.61)	06.9	06.6	0.77
Stimulant	**0.36** (0.22–0.52)	04.2	05.6	0.09
Tobacco	**0.72** (0.65–0.80)	69.6	62.8	<.01

Note: CI = Confidence Interval

[*] N for tobacco =355

[a] Prevalence at test and/or retest of <0.02

Table 3

Concordance of AUDADIS-5 and PRISM-5 DSM-5 substance use disorder dimensional measures N=712[a]

Disorder	ICC (95% CI)	Mean, AUDADIS-5	Mean, PRISM-5	Signed rank test, p-values
Past year				
Alcohol	**0.85** (0.82–0.87)	1.70	1.33	<0.01
Cannabis	**0.79** (0.75–0.81)	0.53	0.52	0.49
Cocaine	**0.90** (0.88–0.91)	0.18	0.17	0.55
Hallucinogen	**0.91** (0.89–0.92)	0.03	0.03	0.92
Heroin	**0.87** (0.85–0.89)	0.09	0.08	0.41
Opioids	**0.68** (0.63–0.72)	0.29	0.22	0.02
Sedative	**0.38** (0.28–0.46)	0.10	0.07	0.26
Stimulant	**0.44** (0.35–0.51)	0.06	0.06	0.85
Tobacco	**0.80** (0.76–0.84)	3.52	2.19	<0.01
Prior to the past year				
Alcohol	**0.82** (0.79–0.84)	5.40	3.86	<0.01
Cannabis	**0.76** (0.73–0.80)	1.61	1.75	0.13
Cocaine	**0.85** (0.83–0.87)	1.14	1.29	0.10
Hallucinogen	**0.72** (0.68–0.76)	0.27	0.28	0.38
Heroin	**0.92** (0.90–0.93)	0.31	0.38	0.08
Opioids	**0.79** (0.76–0.82)	0.89	0.87	0.70
Sedative	**0.76** (0.72–0.79)	0.46	0.58	0.04
Stimulant	**0.72** (0.67–0.76)	0.41	0.47	0.42
Tobacco	**0.84** (0.80–0.87)	4.60	3.53	<0.01
Lifetime				
Alcohol	**0.81** (0.78–0.84)	5.64	4.03	<0.01
Cannabis	**0.78** (0.74–0.81)	1.71	1.82	0.30
Cocaine	**0.86** (0.84–0.88)	1.16	1.31	0.08
Hallucinogen	**0.74** (0.70–0.78)	0.28	0.29	0.46
Heroin	**0.92** (0.91–0.93)	0.31	0.39	0.10
Opioids	**0.79** (0.76–0.82)	0.93	0.97	0.77
Sedative	**0.76** (0.72–0.79)	0.49	0.61	0.04
Stimulant	**0.71** (0.67–0.75)	0.41	0.50	0.27
Tobacco	**0.85** (0.82–0.88)	4.73	3.69	<0.01

ICC = Intraclass correlation coefficient; CI = Confidence interval

[a]N=355 for tobacco