Published in final edited form as:

Drug Alcohol Depend. 2015 July 1; 152: 246–256. doi:10.1016/j.drugalcdep.2015.03.027.

Procedural validity of the AUDADIS-5 depression, anxiety and post-traumatic stress disorder modules: substance abusers and others in the general population*

Deborah S. Hasin^{**},^{1,2,3}, Dvora Shmulewitz^{1,3}, Malka Stohl³, Eliana Greenstein³, Christina Aivadyan³, Kara Morita³, Tulshi Saha⁴, Efrat Aharonovich^{1,3}, Jeusun Jung⁴, Haitao Zhang⁴, Edward V. Nunes^{1,3}, and Bridget F. Grant⁴

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

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

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

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

Abstract

Background—Little is known about the procedural validity of lay-administered, fully-structured assessments of depressive, anxiety and post-traumatic stress (PTSD) disorders in the general population as determined by comparison to clinical re-appraisal, and whether this differs between current regular substance abusers and others. We evaluated the procedural validity of the Alcohol Use Disorder and Associated Disabilities Interview Schedule, DSM-5 Version (AUDADIS-5) assessment of these disorders through clinician re-interviews.

Methods—Test-retest design among respondents from the National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III): (264 current regular substance abusers, 447

*Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

**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.

Disclaimer: The views and opinions expressed in this report are those of the authors and should not be construed to represent the views of sponsoring organizations, agencies, or the U.S. government.

Contributors

Drs. Saha, 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, Ms. Aivadyan, and Ms. Morita collected, cleaned and analyzed the data and critically reviewed drafts of the manuscript. Dr. Shmulewitz and 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.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

others). Clinicians blinded to AUDADIS-5 results administered the semi-structured Psychiatric Research Interview for Substance and Mental Disorders, DSM-5 version (PRISM-5). AUDADIS-5/PRISM-5 concordance was indicated by kappa (κ) for diagnoses and intraclass correlation coefficients (ICC) for dimensional measures (DSM-5 symptom or criterion counts). Results were compared between current regular substance abusers and others.

Results—AUDADIS-5 and PRISM-5 concordance for DSM-5 depressive disorders, anxiety disorders and PTSD was generally fair to moderate ($\kappa = 0.24\text{--}0.59$), with concordance on dimensional scales much better (ICC = 0.53–0.81). Concordance differed little between regular substance abusers and others.

Conclusions—AUDADIS-5/PRISM-5 concordance indicated procedural validity for the AUDADIS-5 among substance abusers and others, suggesting that AUDADIS-5 diagnoses of DSM-5 depressive, anxiety and PTSD diagnoses are informative measures in both groups in epidemiologic studies. The stronger concordance on dimensional measures supports the current movement towards dimensional psychopathology measures, suggesting that such measures provide important information for research in the NESARC-III and other datasets, 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

To diagnose depressive, anxiety and post-traumatic stress disorders in national surveys, trained lay interviewers administer structured diagnostic interviews. Little is known about the validity of these diagnoses in general population samples. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS; Grant et al., 2001) is one such interview. AUDADIS-IV (DSM-IV criteria; American Psychiatric Association, 1994) was used in the U.S. National Longitudinal Alcohol Epidemiologic Survey (NLAES; 1991–1992) and National Epidemiologic Survey on Alcohol and Related Conditions (NESARC; Compton et al., 2004; Grant et al., 2004a, 2009, 2004b). In 2012–2013, NIAAA fielded NESARC-III (Grant, 2014), a survey of 36,309 new respondents, using AUDADIS-5 (Grant et al., 2011) to assess DSM-5 diagnoses (American Psychiatric Association, 2013). All these studies required valid measurement in regular substance abusers and others.

In the U.S. general population, AUDADIS test-retest reliability studies with blinded re-interviews conducted by a second, different interviewer maximized independence of the test and retest interviews, and therefore the rigor of the comparison. Using this methodology, AUDADIS-IV depressive and anxiety diagnoses had moderate to substantial reliability ($\kappa = 0.40\text{--}0.65$; Grant et al., 2003). Dimensional measures of these disorders (criteria or symptom counts) had higher reliability than binary diagnoses (Grant et al., 2003). AUDADIS-5 mood and anxiety disorders were recently shown to have comparable test-retest reliability (Grant et al., 2015).

Scientific utility requires replicable, reliable results across independent interviewers. However, reliability does not guarantee validity. An important strategy to determine validity of a lay-administered diagnostic procedure is comparison to a clinician-administered procedure, often termed procedural validity.

Few procedural validity studies of depressive, anxiety or trauma-related disorders have been conducted in general population samples, and none compared regular substance abusers to others, important because substance abuse can complicate mood and anxiety diagnoses (Hasin et al., 2006, 1996; Torrens et al., 2004). Diagnostic Interview Schedule (DIS) diagnoses were compared to structured psychiatrist re-interviews of Epidemiologic Catchment Area participants (n=370), with psychiatrists blinded to initial DIS interviews. Lay/psychiatrist concordance ranged considerably ($\kappa=0.10-0.50$; Helzer et al., 1985). In studies comparing Composite International Diagnostic Interview (CIDI) to clinician-administered Structured Clinical Interview for DSM (SCID) re-interviews, SCID interviewers were informed of responses to CIDI gateway questions for each module, and reminded participants of these responses, procedures that could increase CIDI/SCID concordance. For example, in National Comorbidity Study participants, CIDI/SCID concordance on Generalized Anxiety Disorder (GAD) was $\kappa=0.33-0.47$ (n=30; Wittchen et al., 1995), and $\kappa=0.45-0.63$ (n=40) for phobic disorders (Wittchen et al., 1996). In other national survey participants, CIDI/SCID concordance for depressive and anxiety disorders was $\kappa=0.42-0.56$ (n=143; Europe), and $\kappa=0.33-0.61$ (n=325; U.S.; Haro et al., 2006). We know of no CIDI procedural validity studies utilizing fully blinded re-evaluations.

Determining the procedural validity of AUDADIS-5 mood and anxiety disorder diagnoses in general population substance abusers and others is important to aid in interpreting NESARC-III (Grant, 2014.) findings. A subset of NESARC-III participants underwent independent clinical re-appraisals using the DSM-5 Psychiatric Research Interview for Substance and Mental Disorders (PRISM-5; Hasin et al., 2011), a semi-structured interview designed to address assessment issues in substance abusers (Hasin et al., 2006). AUDADIS-5 substance disorders showed moderate to substantial procedural validity ($\kappa=0.40-0.72$; Hasin et al., 2015). We now examine the procedural validity of AUDADIS-5 depressive, anxiety, and post-traumatic stress disorders and dimensional disorder measures in the full sample, and in regular substance abusers and others. Procedural validity was also explored by days between AUDADIS-5 and PRISM-5 interviews, since longer intervals could decrease agreement.

2. MATERIAL AND METHODS

2.1 Sample; procedures

NESARC-III, conducted by Westat (Westat) included non-institutionalized civilians 18 years selected via multistage probability sampling (Grant, 2014), with Hispanics, Blacks, and Asians oversampled. The NESARC-III response rate was 60.1%, comparable to many U.S. health surveys (Centers for Disease Control and Prevention; Division of Health Interview Statistics). Participants completed face-to-face AUDADIS-5 interviews (N=36,309); 25,769 consented to re-interviews. From these, 777 potential validity participants were selected with an algorithm using AUDADIS-5 psychiatric and substance

module screening questions (Hasin et al., 2015) to increase prevalence of psychopathology, and residence within the Eastern Time Zone (to facilitate telephone interviews from New York City). NIH, Westat and New York State Psychiatric Institute IRBs approved all procedures; respondents gave informed consent to participate (Hasin et al., 2015). The response rate was 92.5% (712/777) (Hasin et al., 2015). One respondent stopped after the substance modules, leaving n=711 for present analyses. Respondents were classified as current regular substance abusers if they reported at least weekly illicit/non-medical drug use or binge drinking (≥ 5 drinks for men; ≥ 4 for women) in the past year (N=264).

2.2 Re-interview

PRISM-5 clinical re-appraisals were conducted by telephone, permitting centralized, closely-supervised interviewing over a wide geographic area (Kessler et al., 2009). The test-retest interval was 2–69 days (median, 9 days). With consent (n=700; 98.3%), PRISM-5 interviews were recorded.

All PRISM-5 team members were blinded to AUDADIS-5 results. Prior to starting, PRISM-5 interviewers told respondents that the re-interview was to help understand the quality of the previous interview, and that he/she did not have information from that 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.” This procedure was designed to maximize the independence of AUDADIS-5 and PRISM-5 assessments.

2.3 Diagnostic assessment

2.3.1 AUDADIS-5—AUDADIS-5 substance use measures were used to classify current regular users. The AUDADIS-5 mood and anxiety disorders included major depressive episode, persistent depression, panic, generalized anxiety disorder (GAD), social anxiety, specific phobia, agoraphobia, and post-traumatic stress disorder (PTSD). While DSM-5 was finalized in 2012, the diagnostic criteria were anticipated in 2011, and incorporated into AUDADIS-5.

2.3.2 PRISM-5—The validation procedure was the PRISM-5, a semi-structured interview initially designed for DSM-IV (Hasin et al., 2006, 1996). In PRISM, probes for gateway questions, symptoms and criteria are asked as written. However, unlike fully structured interviews, PRISM interviewers add unstructured follow-up probing, informed by their clinical expertise, to obtain more information and clarify responses. The PRISM has fair-substantial test-retest reliability (Hasin et al., 2006) and validity (Hasin et al., 2006; Torrens et al., 2004). DSH and BFG (both involved with DSM-5 development) supervised PRISM adaptation into the computer-assisted PRISM-5 (Hasin et al., 2011) to assess DSM-5 criteria. To reduce participant burden, two shortened versions were created: one with mood disorders; the other with anxiety disorders and PTSD. Versions were randomly assigned to participants.

2.3.3 PRISM-5 interviewers and quality assurance—The ten interviewers had master’s degrees in clinical fields and experience with psychiatric and/or substance abuse

patients (mean years, 4.15, range, 2–14 years; Hasin et al., 2015). Training on PRISM-5, study procedures and confidentiality included a manual, lectures, and group role-plays (Hasin et al., 2015). Trainees were certified after recordings of five interviews were rated satisfactory by trainer/supervisors (EG, CA; Hasin et al., 2015). During the study, supervisors rated recordings from 214 randomly selected PRISM-5 interviews for quality assurance, providing feedback to interviewers in regular meetings. Also, two psychiatrists (Hasin et al., 2015) who received PRISM-5 training, each with >10 years of clinical experience, independently reviewed 107 randomly selected PRISM-5 recordings. Of these, 59 were also reviewed by PRISM-5 supervisors. Among these, 234 relevant diagnoses were possible (3 for each participant in the mood disorder version; 5 for each participant in the anxiety disorder version. Of these mood and anxiety diagnoses, psychiatrists and supervisors agreed on 95.3%; 1.3% were made by the psychiatrist but not a supervisor; 3.4% were made by a supervisor but not the psychiatrist.

2.4 Statistical analyses

Timeframes included past year, prior to past year and lifetime. Kappa (κ) indicated AUDADIS-5/PRISM-5 concordance (Fleiss, 1981b). McNemar's test of paired binary variables determined if AUDADIS-5 and PRISM-5 prevalences differed. Intraclass correlation coefficients (ICC) indicated AUDADIS-5/PRISM-5 concordance on dimensional measures. Signed rank tests, which do not require normality, tested AUDADIS-5/PRISM-5 differences in dimensional measures.

The validity design assumed that interviewers were randomly drawn from a larger population of interviewers, so we used a one-way random effects ANOVA model to derive ICCs (Shrout and Fleiss, 1979). κ and ICC values share the same interpretation (Davis and Fleiss, 1982). Values 0.61 indicate substantial agreement; 0.41–0.60, moderate agreement; and 0.21–0.40, fair agreement (Landis and Koch, 1977).

This study sample was large relative to prior procedural validation studies of depressive and anxiety disorders (Haro et al., 2006; Helzer et al., 1985; Wittchen et al., 1995, 1996), and the selection algorithm enriched the sample. Nevertheless, prevalence of some disorders, e.g., bipolar, was too low to yield stable estimates (<3.0%) and are not reported.

To investigate the impact of current regular substance abuse on procedural validity, we divided the sample into regular substance abusers and others, testing differences in κ and ICC between these sub-groups using Fleiss's (Fleiss, 1981b); and Shoukri's procedures (Shoukri, December 2010. p.50), respectively. We explored whether interval (days) between interviews affected procedural validity by splitting the sample at the median (9 days), and testing differences in κ and ICC using the same procedures.

Among respondents screening into specific AUDADIS-5 modules, we compared lifetime AUDADIS-5 dimensional measures for that module between those screened and not screened into the corresponding PRISM-5 modules, using Kruskal-Wallis tests for bivariate relationships and logistic regressions to control for age, sex, race, substance abuse, and interval between interviews. We also used these tests to compare AUDADIS-5 severity

scores between those with specific lifetime AUDADIS-5 diagnoses who did and did not have corresponding PRISM-5 diagnoses.

3. RESULTS

3.1 Sample characteristics

About two-thirds were non-Hispanic white, about half female, most 30–64 years old, and about half had education past high school (Table 1). The validity sample did not differ from the remaining NESARC-III sample ($n=35,597$) on gender ($\chi^2=0.98$, $p=0.32$) or education ($\chi^2=1.89$, $p=0.49$), but did have somewhat larger proportions of white ($\chi^2=118.8$, $p<0.001$) and younger participants ($t=-3.31$, $df=36,307$, $p<0.001$).

3.2 Concordance on depressive diagnoses

AUDADIS-5/PRISM-5 concordance for depressive disorders was fair to moderate across timeframes ($\kappa=0.35$ – 0.46) (Table 2). AUDADIS-5 diagnosed depressive disorders significantly more often than PRISM-5. Kappas did not differ significantly between current regular substance abusers and others (Table 3) or by interval between interviews (Table 4). Removing the impairment criterion from the diagnostic algorithms produced little difference in kappas in the full sample (Supplementary Table 1¹), by substance abuse (Supplementary Table 2²) or by interval between interviews (Supplementary Table 3³).

3.3 Concordance on depressive dimensional measures

While the signed rank tests indicated that AUDADIS-5 tended to assess more depressive symptoms than PRISM-5, ICCs for major depressive and persistent depressive symptoms indicated substantial agreement (ICCs 0.60–0.64; Table 5). ICCs differed little between substance abusers and others; interval between interviews also had little effect on ICCs (Table 5).

3.4 Screening into the depression modules

AUDADIS-5/PRISM-5 concordance on screening into the major depression and persistent depression modules was fair ($\kappa=0.39$, 0.46 ; Table 6), with little difference by whether participants were regular substance abusers or not, or interval between interviews (Supplementary Tables 4, 5⁴). Of the 81 respondents screened into AUDADIS-5 but not PRISM-5 major depression modules, 72 audible recordings were available. Of these, 45 (62.5%) simply said no to all PRISM-5 screening questions. The others initially said yes, but responses to interviewer probing suggested uncertainty about low mood/anhedonia persistence for 2 weeks, leading to a negative code in PRISM-5's conservative procedures. Among those screening into the AUDADIS-5 major depression module, AUDADIS-5 dimensional severity was significantly higher among those screening positive in PRISM-5 than those screening negative (Table 7). Among those who screened positive for major depressive episode in both interviews, AUDADIS-5 dimensional severity was higher in those

¹Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

²Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

³Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

⁴Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

with lifetime diagnoses in both AUDADIS-5 and PRISM-5 than in those with AUDADIS-5 diagnoses only (Table 8).

3.5 Concordance on DSM-5 anxiety disorders

Except for GAD, AUDADIS-5/PRISM-5 concordance on diagnoses of anxiety disorders was fair to moderate across timeframes ($\kappa=0.24\text{--}0.59$; Table 2), with concordance lowest for GAD ($\kappa=0.18\text{--}0.24$) and highest for Panic Disorder ($\kappa=0.55\text{--}0.59$). AUDADIS-5 diagnosed some but not all anxiety disorders significantly more often than PRISM-5. Kappas did not differ significantly between regular substance abusers and others (Table 3). Interval between interviews also had little effect on κ (Table 4). Removing the impairment criterion from the diagnostic algorithms produced little difference in kappas (Supplementary Tables 1–3⁵).

3.6 Concordance on DSM-5 anxiety disorder dimensional measures

While the signed rank tests indicated that AUDADIS-5 tended to assess more anxiety symptoms than PRISM-5 for some disorders, ICCs for anxiety dimensional measures (except GAD) indicated substantial agreement (ICCs 0.75–0.81; Table 5). This differed little between substance abuser and others (Table 5). Days between interviews had little effect on ICCs (Table 5).

3.7 Screening into anxiety disorder modules

3.7.1 Generalized anxiety disorder (GAD)—AUDADIS-5/PRISM-5 concordance on screening into the GAD modules was low ($\kappa=0.10$; Table 6), little affected by regular substance abuse or interval between interviews (Supplementary Tables 4, 5⁶). Of the 150 respondents screened into the AUDADIS-5 but not PRISM-5 module, 40 were randomly selected for additional review, with audible recordings available for 37, of whom 18 (48.6%) said no to all PRISM-5 screening questions. The others initially said yes, but after probing, were coded no because respondents did not consider worries excessive ($n=11$), persistence of anxiety was uncertain ($n=3$), or worried about only one thing ($n=5$). In those screening into the AUDADIS-5 module, AUDADIS-5 dimensional severity was higher in those screening positive in PRISM-5 (Table 7). Among those screening positive in both interviews, AUDADIS-5 dimensional severity did not differ between those with diagnoses in both AUDADIS-5 and PRISM-5 and those with AUDADIS-5 diagnoses only (Table 8).

3.7.2 Panic disorder—AUDADIS-5/PRISM-5 concordance on screening into the panic disorder modules was moderate, 0.58 (Table 6), with little difference by regular substance abuse or interval between interviews (Supplementary Tables 4, 5⁷). Of the 41 respondents screened into the panic disorder module in AUDADIS-5 but not PRISM-5, 20 were randomly selected for additional review, with audible recordings available for 18, of whom 6 (33.3%) said no to all PRISM-5 screening questions. The others initially said yes, but after probing, responses were coded no because panic attacks were not “out of the blue” ($n=11$) or the respondent misunderstood the initial probe ($n=1$). In those screening into the

⁵Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

⁶Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

⁷Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

AUDADIS-5 module, AUDADIS-5 dimensional severity was greater among those screening positive in PRISM-5 (Table 7). Among those screening positive in both interviews, AUDADIS-5 dimensional severity did not differ between those with diagnoses in both AUDADIS-5 and PRISM-5 and those with AUDADIS-5 diagnoses only (Table 8).

3.7.3 Social Anxiety—AUDADIS-5/PRISM-5 concordance on screening into the social anxiety module was moderate ($\kappa=0.46$; Table 6), with little difference by regular substance abuse or interval between interviews (Supplementary Tables 4, 5⁸). Of the 58 respondents screened into the social anxiety module in AUDADIS-5 but not PRISM-5, 30 were randomly selected for additional review, with audible recordings available for 27, of whom 16 (59.3%) said no to all PRISM-5 screening questions. The others initially said yes, but after probing, responses were coded no because they calmed down quickly in the situation ($n=9$) or didn't consider the fear "strong" ($n=2$). In those screening into the AUDADIS-5 module, AUDADIS-5 dimensional severity was significantly higher in those screening positive in PRISM-5 than in those screening negative (Table 7). Among those screening positive in both interviews, AUDADIS-5 dimensional severity did not differ between those with diagnoses in both AUDADIS-5 and PRISM-5 and those with AUDADIS-5 diagnoses only (Table 8).

3.7.4 Specific phobia—AUDADIS-5/PRISM-5 concordance on screening into the specific phobia modules was fair ($\kappa=0.38$; Table 6), with little difference by regular substance abuse or interval between interviews (Supplementary Tables 4, 5⁹). Of the 25 respondents screened into the specific phobia module in AUDADIS-5 but not PRISM-5, audible recordings were available for 23, of whom 19 (82.6%) responded no to all PRISM-5 screening questions. The others initially said yes, but after probing, responses were coded no because fear arose inconsistently ($n=3$) or ceased during the first few minutes in the situation ($n=1$). In those screening into the AUDADIS-5 module, AUDADIS-5 dimensional severity was significantly higher in those screening positive in PRISM-5 than in those screening negative (Table 7). Among those who screened positive in both interviews, AUDADIS-5 dimensional severity did not differ between those with diagnoses in both AUDADIS-5 and PRISM-5 and those with AUDADIS-5 diagnoses only (Table 8).

3.7.5 Agoraphobia—AUDADIS-5/PRISM-5 concordance on screening into the agoraphobia module was moderate ($\kappa=0.49$; Table 6), with little difference by regular substance abuse or interval between interviews (Supplementary Tables 4, 5¹⁰). Of the 50 respondents screened into the agoraphobia module in AUDADIS-5 but not PRISM-5, 11 were randomly selected for additional review, with audible recordings available for 10, of whom 4 (40.0%) said no to all PRISM-5 screening questions. In other cases, probing indicated that fear ceased quickly in the situation ($n=4$) or only one agoraphobic situation was feared ($n=2$) so responses were coded no. Among those screening into the AUDADIS-5 module, AUDADIS-5 dimensional severity was significantly higher in those screening positive in PRISM-5 than in those screening negative (Table 7). Among those who screened

⁸Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

⁹Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

¹⁰Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

positive in both interviews, AUDADIS-5 dimensional severity did not differ significantly between those with diagnoses in AUDADIS-5 and PRISM-5 and those with AUDADIS-5 diagnoses only (Table 8).

3.8 Concordance on Post-traumatic Stress Disorder (PTSD)

AUDADIS-5/PRISM-5 concordance on PTSD was fair to moderate across timeframes ($\kappa=0.34-0.46$) (Table 2). McNemar tests showed that AUDADIS-5 diagnosed PTSD more often than PRISM-5. Dividing the sample into current substance abusers and non-abusers produced higher kappas in substance abusers than in others for past-year but not other timeframes (Table 3). Concordance did not differ by interval between interviews (Table 4). Removing the impairment criterion from the diagnostic algorithm produced little difference in kappas (Supplementary Tables 1–3¹¹).

3.9 Concordance on PTSD dimensional measures

Concordance of AUDADIS-5 and PRISM-5 on dimensional measures of PTSD was substantial (Table 5), for all symptoms combined (ICC=0.69), and for different components (ICC=0.53–0.66). Sub-setting by current regular substance abuse or by interval between interviews had little effect on the ICCs (Table 5).

3.10 Screening into the PTSD module

AUDADIS-5/PRISM-5 concordance on screening into the PTSD module was fair ($\kappa=0.33$ Table 6), with little difference by substance abuse or interval between interviews (Supplementary Tables 4, 5¹²). Of the 14 respondents screened into the PTSD module in AUDADIS-5 but not PRISM-5, audible recordings were available for 12; 10 (83.3%) responded no to all PRISM-5 screening questions. Of participants who initially said yes, one refused to select a most traumatic experience; the other initially said yes to seeing a dead body/parts, but was re-coded to no after probing indicated that this did not occur outside of a hospital or wake. In those that screened into the AUDADIS-5 module, AUDADIS-5 dimensional severity was significantly higher in those screening positive in PRISM-5 than in those screening negative (Table 7). Among those screening positive in both interviews, AUDADIS-5 dimensional severity was higher in those with diagnoses in both AUDADIS-5 and PRISM-5 than in those with AUDADIS-5 diagnoses only (Table 8).

4. DISCUSSION

A large procedural validity study of AUDADIS-5 diagnoses of DSM-5 depression and anxiety disorders and PTSD was conducted among selected NESARC-III participants. The validation procedure, PRISM-5, was administered by clinicians blinded to AUDADIS-5 results, providing an independent, rigorous test of agreement between the two procedures. AUDADIS-5 and PRISM-5 depressive disorders demonstrated fair to moderate concordance, with substantial concordance on dimensional depression measures. Except for GAD, anxiety disorders and PTSD showed fair to moderate AUDADIS-5/PRISM-5 concordance for

¹¹Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

¹²Supplementary material can be found by accessing the online version of this paper at <http://dx.doi.org> and by entering doi:...

diagnoses, and substantial concordance for dimensional measures. When AUDADIS-5/PRISM-5 discrepancies arose, AUDADIS-5 made diagnoses more frequently than PRISM-5, and indicated greater numbers of symptoms and criteria. A question of key interest was whether procedural validity differed between current regular substance abusers and others. This did not occur, and in addition, neither days between interviews nor exclusion of the impairment criterion influenced results.

Concordance in the present study was comparable to or higher than the only previous general population procedural validity study with blinded clinician re-appraisals (Helzer et al., 1985). Some concordances reported here are also comparable to procedural validity results when clinician re-interviewers and participants were informed of initial interview results (Haro et al., 2006; Wittchen et al., 1995, 1996), although the latter design did produce higher concordance for many disorders, as would be expected. In our study, many discordant cases arose because respondents simply said no to PRISM-5 screening questions worded similarly to AUDADIS-5 screening questions, suggesting an order effect more pronounced for mood and anxiety disorders that came later in the interview than for substance disorders (always administered first; Hasin et al., 2015). Other respondents initially said yes to PRISM-5 screening, but were scored no after PRISM-5's more conservative probing procedures. Of respondents screened into AUDADIS-5 mood and anxiety modules, AUDADIS-5 dimensional scores were significantly higher in respondents who screened positive in PRISM-5 than in those who did not, suggesting that PRISM-5 positive-screens were more severe cases. Overall, the considerable concordance of AUDADIS-5 and PRISM-5 on symptom counts indicates validity of NESARC-III dimensional measures. AUDADIS-5 binary diagnoses appear to have a lower threshold, casting a wider net than the conservative clinical PRISM-5. This AUDADIS-5 feature may be advantageous in indicating overall public health need, by identifying conditions clinicians might consider sub-threshold that might still benefit from appropriate intervention.

Dimensional measures generally showed greater AUDADIS-5/PRISM-5 concordance than binary diagnoses, underscoring value of dimensional approaches and consistent with increasing research interest in dimensional assessment (Faraone, 2013; Fazzino et al., 2014; Hasin et al., 2013; Regier et al., 2012). Dimensional approaches may also be useful clinically, since mild and severe disorders require different clinical management. The dimensional results are also relevant to phenotypes analyses using DNA collected from 24,381 NESARC-III participants.

Although "validation" is used to characterize this study, no psychiatric instrument provides a perfect gold standard. PRISM has moderate to substantial reliability for depressive, anxiety and PTSD diagnoses (Hasin et al., 2006), and better validity than SCID for major depression (Torrens et al., 2004). However, imperfect reliability decreases ability to show associations with other variables, including diagnoses from fully-structured interviews. Thus, while clinician re-appraisal is important, antecedent, concurrent, and prospective validation also merits investigation.

Study limitations are noted. First, individuals without phone access were excluded. However, few in the U.S. lack phone access (McInnes et al., 2013; United States Census

Bureau, 2011), so this does not greatly limit generalizeability. Second, compared to the full NESARC-III sample, the validity subsample had more white, young respondents, consistent with enrichment of the subsample for the common psychiatric disorders prevalent in those groups. Future studies should focus on older participants, and race/ethnic minorities. Third, the study did not target treated individuals. While a general population sample was of prime importance to NESARC-III, clinician re-appraisal studies in patient samples also contribute important information and should be conducted (Canino et al., 1999; Cottler et al., 1997). Fourth, the present sample had too few cases of some disorders (e.g., bipolar) to analyze. These merit examination in specialized samples. Fifth, we did not examine individual items or criteria, which warrant future attention. Sixth, only one dichotomization of the interval between interviews is presented. However, sensitivity analyses exploring six alternative dichotomizations showed no additional significant findings for kappas, and a few additional significant findings for ICCs that did not form a consistent pattern. Seventh, we did not present procedural validity by participants' functioning measured on an external measure, e.g., the SF-12, although exploratory analyses indicated no consistent pattern of SF-12 influences on procedural validity. This merits more detailed attention in future studies. Eighth, AUDADIS-5 was always the first interview and PRISM-5 the second. Future studies should balance interview order to differentiate procedural discrepancies from order effects. Ninth, AUDADIS-5 was administered in person, and PRISM-5 by telephone. Centralized PRISM-5 telephone administration facilitated efficient, cost-effective, closely-supervised interviews, but left unanswered whether concordance would be higher if administration mode were constant in the two interviews. Nevertheless, telephone vs. in-person modes appear not to influence findings (Aziz and Kenford, 2004; Kessler et al., 2009; Sobin et al., 1993; Wells et al., 1988), so discordance was likely due to other factors. Finally, other participant characteristics may impact validity, warranting future investigation.

Study strengths are also noted. The large sample provided stable concordance estimates even for relatively low-prevalence conditions. Among those selected for re-interview, the response rate was high. Participants came from the general population, providing important information on the main type of participant for whom AUDADIS-5 was designed. The PRISM-5 team was blind to AUDADIS-5 results, eliminating potential bias, including influences of AUDADIS-5 results on PRISM-5 interviewer questions and probes. The short re-test interval reduced the influence of new illness onset on concordance. The semi-structured clinician reappraisal procedure, PRISM-5, has many strengths (Hasin et al., 2006), has been used extensively in drinking and drug-abusing samples (Drake et al., 2011; Hasin et al., 2002; Nunes et al., 2006; Torrens et al., 2011), was a source of information to the DSM-5 SUD workgroup that influenced decisions on DSM-5 (Hasin et al., 2012), and has been the gold standard in measurement studies of other instruments (Cuenca-Royo et al., 2012; Mestre-Pinto et al., 2014). PRISM-5 interviewers all had advanced clinical degrees and considerable patient experience. PRISM-5 training and supervision was standardized and rigorous. Finally, by comparing procedural validity between regular substance abusers and others, we addressed an important validity question not previously examined. These strengths contributed to a rigorous clinical re-appraisal study for indicating the procedural validity of AUDADIS-5 measures of depressive and anxiety disorders and PTSD.

The previous NLAES and NESARC surveys produced extensive epidemiologic information about psychiatric and substance use disorders (Hasin, In Press). NESARC-III will extend this knowledge with new data on prevalence, environmental, clinical and genetic risk factors. Given the widespread importance of national findings for research and policy, understanding AUDADIS-5 procedural validity in regular substance abusers and others is crucial. To further understand AUDADIS-5 measures, future studies should include multivariate analyses of factors influencing validity. Meanwhile, the concordance with clinician re-appraisals of AUDADIS-5 depressive, anxiety and traumatic stress measures suggests that AUDADIS-5, especially its dimensional measures, is a useful diagnostic instrument for research purposes with substance abusers and others in the NESARC-III and other general population samples.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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. No authors have any relevant financial interests.

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.

References

- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorder. 4. American Psychiatric Association Publishing; Washington, DC: 1994.
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorder. 5. American Psychiatric Association Publishing; Washington, DC: 2013.
- 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, Bravo M, Ramirez R, Febo VE, Rubio-Stipec M, Fernandez RL, Hasin D. 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, June 2013.
- 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, Blaine J, 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. National Center for Health Statistics; Hyattsville, Maryland: 2012.
- Drake RE, Caton CL, Xie H, Hsu E, Gorroochurn 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]
- 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. John Wiley and Sons, Inc; New York: 1981a. The Measurement Of Interrater Agreement, Formula 13.22; p. 222
- Fleiss, JL. Statistical Methods For Rates And Proportions. John Wiley and Sons, Inc; New York: 1981b.
- Grant, BF. Source And Accuracy Statement: National Epidemiologic Survey On Alcohol And Related Conditions-III (NESARC-III). National Institute on Alcohol Abuse and Alcoholism; Rockville, MD: 2014. <http://www.niaaa.nih.gov/research/nescarc-iii>
- 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). National Institute on Alcohol Abuse and Alcoholism; Bethesda, MD: 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 alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules 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.; Goldstein, R.; Aivadyan, C.; Hasin, D. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition Version (AUDADIS-5). National Institute on Alcohol Abuse and Alcoholism; Rockville, MD: 2011.
- Grant BF, Goldstein R, Smith SM, Jung J, Zhang H, Chou SP, Pickering RP, Ruan WJ, Huang B, Saha T, Aivadyan C, Greenstein E, Hasin DS. 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. *Drug Alcohol Depend*. 2015; 148:27–33. [PubMed: 25595052]
- Grant BF, Goldstein RB, Chou SP, Huang B, Stinson FS, Dawson DA, Saha TD, 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, 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]
- Haro JM, Arbabzadeh-Bouchez S, Brugha TS, de Girolamo G, Guyer ME, Jin R, Lepine JP, Mazzi F, Reneses B, Vilagut G, Sampson NA, Kessler RC. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys. *Int J Methods Psychiatr Res*. 2006; 15:167–180. [PubMed: 17266013]
- 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.; Aivadyan, C.; Greenstein, E.; Grant, BF. Psychiatric Research Interview for Substance Use and Mental Disorders, Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (PRISM-5) version. Columbia University, Department of Psychiatry; New York, NY: 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, Greenstein E, Aivadyan C, Stohl M, Aharonovich E, Saha T, Goldstein R, Nunes EV, Jung J, Zhang H, Grant BF. 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. *Drug Alcohol Depend*. 2015; 148:40–46. [PubMed: 25604321]
- 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*. 2013; 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):e11–24.
- Mestre-Pinto JI, Domingo-Salvany A, Martin-Santos R, Torrens M, PsyCoBarcelona G. 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]
- 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]
- Shoukri, MM. Measures of Inter-observer Agreement and Reliability. 2. Chapman & Hall/CRC Press; Boca Raton, Florida: Dec. 2010 p. 50
- 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. Psychiatric comorbidity in illicit drug users: substance-induced versus independent disorders. *Drug Alcohol Depend*. 2011; 113:147–156. [PubMed: 20801586]

- Torrens M, Serrano D, Astals M, Perez-Dominguez 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 on 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>
- 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]
- Westat. [accessed on January 20, 2015] <https://www.westat.com/>
- Wittchen HU, Kessler RC, Zhao S, Abelson J. Reliability and clinical validity of UM-CIDI DSM-III-R generalized anxiety disorder. *J Psychiatr Res*. 1995; 29:95–110. [PubMed: 7666382]
- Wittchen HU, Zhao S, Abelson JM, Abelson JL, Kessler RC. Reliability and procedural validity of UM-CIDI DSM-III-R phobic disorders. *Psychol Med*. 1996; 26:1169–1177. [PubMed: 8931163]

- Procedural validity of DSM-5 AUDADIS-5 mood and anxiety diagnoses was examined
- Sample: current regular substance abusers and others in the general population
- Validation procedure: blinded clinician-administered PRISM-5 interviews
- AUDADIS-5 diagnostic validity on: fair-moderate in substance abusers and others
- AUDADIS-5 dimensional validity: moderate-substantial in substance abusers and others
- AUDADIS-5 measures are useful in regular substance abusers and others

Table 1

Characteristics of procedural validity study respondents (n=712)

Sociodemographic characteristics	N	% (SE)
Sex		
Male	324	45.5 (1.87)
Female	388	54.5 (1.87)
Age		
18–24	81	11.4 (1.19)
25–44	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)

Table 2

AUDADIS-5 and PRISM-5 concordance on DSM-5 diagnoses of depressive, anxiety disorders and PTSD: full sample

Disorder	Prevalence (%)		κ (95% CI)	Difference in prevalence, McNemar's Test, p-values
	AUDADIS-5	PRISM-5		
Past year				
Major Depression	42.0	23.1	.35 (.26, .45)	<0.0001
Persistent depression	14.1	10.4	.36 (.22, .50)	0.063
Any depressive	45.6	26.5	.40 (.31, .49)	<0.0001
Specific phobia	20.2	16.6	.34 (.22, .46)	0.12
Social anxiety	12.4	17.4	.25 (.12, .38)	0.029
Agoraphobia	07.6	07.6	.24 (.07, .41)	1.0
Panic	19.4	13.2	.59 (.48, .70)	0.0005
GAD	29.2	19.4	.22 (.12, .33)	0.0006
Any anxiety	51.4	39.9	.32 (.23, .42)	0.0002
PTSD	28.9	14.6	.34 (.23, .44)	<0.0001
Prior to the past year				
Major Depression	67.3	44.8	.35 (.26, .43)	<0.0001
Persistent depression	25.2	17.5	.46 (.35, .57)	0.0008
Any depressive	71.3	47.9	.36 (.28, .45)	<0.0001
Specific phobia	22.8	19.1	.36 (.25, .48)	0.13
Social anxiety	18.3	21.1	.36 (.24, .48)	0.24
Agoraphobia	09.8	09.3	.35 (.19, .51)	0.75
Panic	27.5	24.7	.55 (.45, .65)	0.20
GAD	39.6	27.5	.20 (.10, .30)	0.0002
Any anxiety	63.5	50.8	.27 (.18, .37)	<0.0001
PTSD	35.1	24.2	.46 (.37, .56)	<0.0001
Lifetime				
Major Depression	71.6	47.9	.35 (.27, .43)	<0.0001
Persistent depression	25.1	17.5	.46 (.35, .57)	0.0008
Any depressive	75.2	51.0	.36 (.28, .45)	<0.0001
Specific phobia	22.8	19.1	.36 (.25, .48)	0.13
Social anxiety	18.3	21.4	.36 (.24, .47)	0.20
Agoraphobia	09.8	09.3	.35 (.19, .51)	0.75
Panic	28.7	26.1	.56 (.46, .65)	0.26
GAD	41.9	29.5	.18 (.08, .28)	0.0002
Any anxiety	64.6	52.3	.27 (.17, .37)	0.0001
PTSD	36.2	24.2	.44 (.35, .54)	<0.0001

Note: CI = Confidence Interval; any depressive includes respondents with major or persistent depression; any anxiety includes respondents with specific phobia, social anxiety, agoraphobia, Panic disorder, or GAD.

n=355 for major or persistent depression, n=356 for anxiety disorders and PTSD

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

AUDADIS-5 and PRISM-5 concordance on DSM-5 diagnoses of depressive, anxiety disorders and PTSD: By current regular substance abuse status ^a

Table 3

	Participants who were current regular substance abusers ^a				Participants who were not current regular substance abusers				χ^2 test of difference in κ between substance abusers & others p-values ^b
Disorder	Prevalence (%)		Difference in prevalence, McNemar's Test, p-values	κ (95% CI)	Prevalence (%)		Difference in prevalence, McNemar's Test, p-values	κ (95% CI)	
	AUDADIS-5	PRISM-5			AUDADIS-5	PRISM-5			
	<i>Past year</i>								
Major Depression	48.4	24.2	<.0001	.38 (.24, .52)	38.3	22.5	<.0001	.33 (.21, .45)	0.54
Persistent depression	17.2	09.4	.018	.40 (.18, .62)	12.3	11.0	.59	.34 (.16, .52)	0.62
Any depressive	51.6	28.1	<.0001	.41 (.28, .55)	42.3	25.6	<.0001	.39 (.27, .51)	0.79
Specific phobia	16.9	18.4	.72	.24 (.04, .44)	22.3	15.5	.019	.40 (.25, .55)	0.09
Social anxiety	14.7	17.7	.45	.24 (.04, .45)	10.9	17.3	.027	.26 (.09, .42)	0.85
Agoraphobia	08.1	08.1	1.00	.31 (.04, .58)	07.3	07.3	1.00	.19 (−.02, .40)	0.37
Panic	25.7	17.7	.016	.55 (.38, .72)	15.5	10.5	.012	.62 (.46, .77)	0.52
GAD	33.8	23.5	.029	.18 (.01, .35)	26.4	16.8	.005	.24(.10, .39)	0.53
Any anxiety	54.4	41.9	.011	.35 (.20, .50)	49.6	38.6	.006	.31 (.18, .43)	0.62
PTSD	38.2	21.3	<.0001	.44 (.29, .59)	23.2	10.5	<.0001	.21 (.07, .36)	0.03
<i>Prior to the past year</i>									
Major Depression	63.3	38.3	<.0001	.32 (.18, .46)	69.6	48.5	<.0001	.36 (.25, .47)	0.57
Persistent depression	25.8	16.4	.014	.44 (.26, .63)	24.7	18.1	.015	.47 (.33, .60)	0.73
Any depressive	68.0	43.0	<.0001	.38 (.24, .51)	73.1	50.7	<.0001	.35 (.24, .46)	0.68
Specific phobia	20.6	20.6	1.00	.28 (.09, .47)	24.1	18.2	.047	.42 (.27, .56)	0.13
Social anxiety	17.7	18.4	.85	.33 (.13, .53)	18.6	22.7	.18	.38 (.23, .53)	0.59
Agoraphobia	11.0	09.6	.59	.44 (.20, .69)	09.1	09.1	1.00	.29 (.08, .49)	0.26
Panic	30.9	27.2	.34	.52 (.36, .68)	25.5	23.2	.40	.57 (.44, .70)	0.55
GAD	40.4	29.4	.039	.15 (−.01, .32)	39.1	26.4	.001	.23(.10, .36)	0.34
Any anxiety	61.0	50.7	.043	.29 (.13, .45)	65.0	50.9	.0006	.26 (.14, .38)	0.71
PTSD	41.2	30.2	.005	.54 (.40, .69)	31.4	20.5	.0009	.39 (.26, .53)	0.10
<i>Lifetime</i>									
Major Depression	69.5	40.6	<.0001	.32 (.18, .45)	72.7	52.0	<.0001	.36 (.25, .47)	0.59

Disorder	Participants who were current regular substance abusers ^a				Participants who were not current regular substance abusers				χ^2 test of difference in κ between substance abusers & others p-values ^b
	Prevalence (%)		Difference in prevalence, McNemar's Test, p-values	κ (95% CI)	Prevalence (%)		Difference in prevalence, McNemar's Test, p-values	κ (95% CI)	
	AUDADIS-5	PRISM-5			AUDADIS-5	PRISM-5			
Persistent depression	25.8	16.4	.014	.44 (.26, .63)	24.7	18.1	.019	.47 (.33, .60)	0.73
Any depressive	73.4	45.3	<.0001	.37 (.24, .50)	76.2	54.2	<.0001	.36 (.24, .47)	0.89
Specific phobia	20.6	20.6	1.00	.28 (.09, .47)	24.1	18.2	.047	.42 (.27, .56)	0.13
Social anxiety	18.4	17.7	.85	.33 (.13, .53)	18.6	23.2	.14	.37 (.22, .52)	0.67
Agoraphobia	11.0	09.6	.59	.44 (.20, .69)	09.1	09.1	1.00	.29 (.08, .49)	0.26
Panic	32.4	29.4	.45	.52 (.36, .67)	26.4	24.1	.40	.58 (.45, .70)	0.47
GAD	42.7	31.6	.043	.14 (-.02, .31)	41.4	28.2	.0013	.20 (.08, .33)	0.47
Any anxiety	62.5	52.2	.043	.29 (.13, .44)	65.9	52.3	.0008	.26 (.14, .39)	0.71
PTSD	43.4	30.2	.0015	.50 (.36, .65)	31.8	20.5	.0006	.39 (.25, .52)	0.23

Notes: CI = Confidence Interval; any depressive includes respondents with major or persistent depression; any anxiety includes respondents with specific phobia, social anxiety, agoraphobia, Panic disorder, or GAD.

^a current at least weekly illicit drug use or risky drinking (5 drinks per occasion for men, 4 drinks for women)

For major or persistent depression, n=128 regular substance abusers, n=227 others

For anxiety disorders and PTSD, n=136 regular substance abusers, n=220 others

^b χ^2 statistic: formula from 13.22 in Fleiss, 1981 (Fleiss, 1981a); values <0.05 are shown in italics

Table 4

AUDADIS-5 and PRISM-5 concordance on lifetime diagnoses of depressive, anxiety disorders and PTSD: By interval between interviews ^a

Disorder	Interval: at or below median (9 days)				Interval: greater than 9 days			
	Prevalence (%)		k (95% CI)	Difference in prevalence, McNemar's test p-value	Prevalence (%)		k (95% CI)	Difference in prevalence, McNemar's test p-values
	AUDADIS-5	PRISM-5			AUDADIS-5	PRISM-5		
Major Depression	70.6	46.2	.39 (.28, .50)	<.0001	72.8	50.0	.29 (.16, .42)	<.0001
Persistent depression	25.4	19.8	.44 (.29, .58)	.08	24.7	14.6	.49 (.32, .65)	.0017
Any depressive	73.6	49.8	.38 (.27, .50)	<.0001	77.2	52.5	.34 (.21, .46)	<.0001
Specific phobia	23.2	19.7	.33 (.17, .48)	.30	22.2	18.4	.41 (.24, .59)	.27
Social anxiety	18.2	22.2	.34 (.19, .50)	.22	18.4	20.3	.37 (.19, .55)	.59
Agoraphobia	09.1	09.6	.20 (-.01, .40)	.85	10.8	08.9	.54 (.31, .76)	.41
Panic	28.8	26.3	.58 (.46, .71)	.38	28.5	26.0	.52 (.37, .67)	.47
GAD	39.4	31.3	.21 (.07, .35)	.06	44.9	27.2	.15 (.01, .30)	.0005
Any anxiety	61.1	50.5	.22 (.09, .35)	.02	69.0	54.4	.33 (.19, .47)	.0013
PTSD	36.4	23.7	.45 (.32, .58)	.0003	36.1	24.7	.44 (.29, .59)	.0035
								.92

Notes: CI = Confidence Interval; any depressive includes respondents with major or persistent depression; any anxiety includes respondents with specific phobia, social anxiety, agoraphobia, Panic disorder, or GAD.

^aFor major or persistent depression, n=197 at or below the median, n=158 above; for anxiety disorders and PTSD, n=198 at or below the median, n=158 above

^bTested using the chi-square statistic from formula 13.22 in Fleiss 1981(Fleiss, 1981a); values <0.05 are shown in italics

AUDADIS-5 and PRISM-5 concordance (ICC^a) on lifetime dimensional (severity) measures^b of DSM-5 depressive, anxiety disorders and PTSD

Table 5

Total sample				By whether participants were regular substance abusers ^c		By interval between interviews ^d				
Mean symptoms				ICC (95% CI)		ICC (95% CI)				
Disorder	AUDADIS-5	PRISM-5	ICC (95% CI)	Difference in severity, signed rank test ^e , p-values	Difference in ICCs ^f , substance abusers and others, p-value		Difference in ICCs ^f , those at or below the median vs. those above, p-value			
					Regular Substance abusers	Other participants		Interval: at or below median (9 days)	Interval: above median (9 days)	
Major Depression	5.8	4.0	0.60 (.51–.68)	<0.01	0.59 (.42–.71)	0.61 (.49–.70)	0.78	0.64 (.52–.73)	0.56 (.39–.68)	0.24
Persistent depression	1.3	0.9	0.64 (.56–.71)	<0.01	0.67 (.53–.77)	0.62 (.51–.71)	0.44	0.60 (.47–.69)	0.70 (.58–.78)	0.10
Specific phobia	4.2	4.2	0.81 (.77–.85)	0.87	0.78 (.69–.84)	0.84 (.79–.88)	0.11	0.80 (.73–.85)	0.83 (.77–.88)	0.40
Social anxiety	4.5	3.9	0.75 (.70–.80)	<0.01	0.75 (.65–.82)	0.76 (.68–.81)	0.83	0.75 (.67–.81)	0.76 (.67–.82)	0.83
Agoraphobia	1.3	1.2	0.77 (.72–.82)	0.08	0.78 (.70–.85)	0.76 (.69–.82)	0.65	0.80 (.73–.85)	0.74 (.64–.81)	0.17
Panic	3.3	2.8	0.76 (.70–.80)	0.03	0.74 (.64–.82)	0.76 (.69–.82)	0.68	0.75 (.68–.80)	0.76 (.67–.82)	0.83
GAD	3.9	2.1	0.19 (0.01–.35)	<0.01	0.35 (.09–.50)	0.08 (–.20–.29)	0.01	0.32 (.10–.48)	0.03 (–.33–.29)	<0.01
PTSD (20 symptoms)	9.9	8.7	0.69 (.62–.75)	<0.01	0.71 (.60–.80)	0.66 (.55–.74)	0.39	0.64 (.53–.73)	0.74 (.65–.81)	0.07
B. Intrusion symptoms (5)	2.9	2.7	0.66 (.58–.72)	0.22	0.62 (.47–.73)	0.68 (.58–.75)	0.34	0.66 (.55–.74)	0.66 (.54–.75)	1.00
C. Avoidance or efforts to avoid (2)	1.1	1.1	0.53 (.42–.62)	0.87	0.62 (.47–.73)	0.45 (.29–.58)	0.03	0.56 (.42–.67)	0.49 (.30–.63)	0.37
D. Change in cognition/mood (7)	3.4	3.0	0.63 (.54–.70)	<0.01	0.65 (.51–.75)	0.60 (.47–.69)	0.45	0.58 (.45–.68)	0.68 (.56–.77)	0.12
E. Change in arousal/activity (6)	2.5	1.9	0.62 (.53–.69)	<0.01	0.61 (.45–.72)	0.60 (.48–.69)	0.89	0.50 (.34–.62)	0.74(.64–.81)	<0.01

^aICC = Intraclass correlation coefficient; CI = Confidence interval^bDefinition of dimensional measures:

Major depression= count of major depression symptoms endorsed, from DSM-5 criterion A (maximum of 9)

Persistent depression=count of persistent depression symptoms endorsed, from DSM-5 criterion B (maximum of 6)

Specific phobia=count of specific phobia situations endorsed, from DSM-5 criterion A (maximum of 17)

Social anxiety=count of social anxiety situations endorsed, from DSM-5 criterion A (maximum of 16)

Agoraphobia=count of agoraphobia situations endorsed, from DSM-5 criterion A (maximum of 6)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Panic= count of panic disorder symptoms endorsed, from DSM-5 criterion A (maximum of 13)
GAD= count of GAD symptoms endorsed, from DSM-5 criterion C (maximum of 6)
PTSD= count of PTSD symptoms endorsed, from DSM-5 criteria B, C, D, and E (maximum of 20)

^c current weekly or greater binge drinking (5 drinks per occasion for men, 4 drinks for women) or illicit drug use; depressive disorders: regular substance abusers, n=128; others, n=227; anxiety disorders and PTSD: regular substance abusers, n=136; others, n=220

^d depressive disorders: at or below median, n=197; above median, n=158; anxiety disorders and PTSD: at or below median, n=198; above median, n=158

^e Signed rank test p-values indicate if distributions of AUDADIS-5 and PRISM-5 dimensional measures differ significantly (non-parametric equivalent of a paired t-test).

^f Test of ICCs between independent samples is from Shoukri MM: Measures of Inter-observer Agreement and Reliability, Second Edition, p. 50(Shoukri, December 2010. p.50); values <0.05 are shown in italics

Table 6

AUDADIS-5 and PRISM-5 concordance on screening into diagnostic modules

Disorder	Prevalence (%)		k (95% CI)	Difference in prevalence, McNemar's Test, p-values
	AUDADIS-5	PRISM-5		
Major Depression	82.5	62.5	.39 (.29, .48)	<0.0001
Persistent depression	25.9	17.5	.46 (.35, .57)	0.0002
Specific phobia	88.2	87.4	.38 (.24, .53)	0.66
Social anxiety	67.4	59.8	.46 (.37, .56)	0.004
Agoraphobia	36.5	31.2	.49 (.40, .59)	0.035
Panic	37.4	33.4	.58 (.50, .67)	0.09
GAD	83.4	46.4	.10 (.03, .17)	<0.0001
PTSD	94.1	93.8	.33 (.14, .52)	0.85

Note: CI = Confidence Interval; n=355 for major or persistent depression, n=356 for anxiety disorders and PTSD

AUDADIS-5 dimensional measures (severity)^a among those who screened into the AUDADIS-5 module, by screening into the PRISM-5 module

Table 7

Screen into PRISM-5 module	N	Mean (SD)	Median	range	K-W ^b χ^2	OR ^c (95%CI)
<i>Major depression (N=293)</i>						
Yes	212	7.3 (1.6)	8.0	1–9	22.9 ^{***}	1.48 (1.26, 1.72) ^{***}
No	81	6.1 (2.0)	6.0	2–9		
<i>Persistent depression (N=92)</i>						
Yes	44	5.4 (1.0)	6.0	1–6	3.2	1.53 (1.02, 2.30) [*]
No	48	4.8 (1.5)	5.0	0–6		
<i>Specific phobia (N=314)</i>						
Yes	289	5.0 (3.5)	4.0	1–16	9.7 ^{**}	1.30 (1.07, 1.57) ^{**}
No	25	2.9 (2.3)	2.0	1–9		
<i>Social anxiety (N=240)</i>						
Yes	182	7.3 (4.4)	7.0	1–16	11.6 ^{***}	1.13 (1.05, 1.23) ^{**}
No	58	5.0 (3.6)	4.0	1–14		
<i>Agoraphobia (N=130)</i>						
Yes	80	3.4 (1.5)	3.0	1–6	17.2 ^{***}	1.99 (1.35, 2.91) ^{***}
No	50	2.4 (0.8)	2.0	1–5		
<i>Panic (N=133)</i>						
Yes	92	9.1 (2.5)	9.0	2–13	1.4	1.17 (1.01, 1.35) [*]
No	41	8.1 (3.5)	9.0	0–13		
<i>GAD (N=297)</i>						
Yes	147	4.9 (1.4)	5.0	0–6	6.0 [*]	1.18 (1.00, 1.39) [*]
No	150	4.5 (1.5)	5.0	0–6		
<i>PTSD (N=335)</i>						
Yes	321	10.7 (6.1)	12.0	0–20	4.4 [*]	1.13 (1.02, 1.25) [*]

Screen into PRISM-5 module	N	Mean (SD)	Median	range	K-W ^b χ^2	OR ^c (95%CI)
No	14	7.2 (5.7)	5.5	0–20		

^aDefinition of AUDADIS-5 severity measures:
Major depression= count of major depression symptoms endorsed, from DSM-5 criterion A (maximum of 9)
Persistent depression=count of persistent depression symptoms endorsed, from DSM-5 criterion B (maximum of 6)
Specific phobia=count of specific phobia situations endorsed, from DSM-5 criterion A (maximum of 17)
Social anxiety=count of social anxiety situations endorsed, from DSM-5 criterion A (maximum of 16)
Agoraphobia=count of agoraphobia situations endorsed, from DSM-5 criterion A (maximum of 6)
Panic = count of panic disorder symptoms endorsed, from DSM-5 criterion A (maximum of 13)
GAD= count of GAD symptoms endorsed, from DSM-5 criterion C (maximum of 6)
PTSD= count of PTSD symptoms endorsed, from DSM-5 criteria B, C, D, and E (maximum of 20)

^bK-W χ^2 = Kruskal-Wallis χ^2 , testing for differences in the distributions of the severity indicators

^cOR= odds ratio; CI = Confidence interval; from logistic regression, AUDADIS-5 severity predicting PRISM-5 screening status, correcting for age, sex, race, current regular substance abuse, interval between interviews.

p-values:
* <0.05;
** <0.01;
*** <0.001

AUDADIS-5 severity^a among those who screened into both AUDADIS-5 and PRISM-5 modules and received an AUDADIS-5 lifetime diagnosis, by PRISM-5 lifetime disorder diagnosis

Table 8

PRISM-5 lifetime diagnosis	N	Mean (SD)	Median	Range	K-W ^b	χ^2	OR ^c (95%CI)
<i>Major depression (N=194)</i>							
Yes	153	7.8 (1.2)	8.0	5–9	13.2 ***		1.93 (1.39, 2.67) ***
No	41	7.0 (1.3)	7.0	5–9			
<i>Specific phobia (N=76)</i>							
Yes	37	5.7 (3.5)	6.0	1–14	0.1		0.96 (0.83, 1.10)
No	39	6.2 (4.1)	6.0	1–16			
<i>Social anxiety (N=54)</i>							
Yes	34	9.8 (4.0)	11.0	3–15	2.4		1.14 (0.99, 1.31)
No	20	7.6 (4.8)	7.0	1–15			
<i>Agoraphobia (N=29)</i>							
Yes	14	4.4 (1.2)	4.5	2–6	2.9		2.28 (0.96–5.40)
No	15	3.7 (1.2)	3.0	2–6			
<i>Panic (N=74)</i>							
Yes	66	9.4 (2.4)	9.0	4–13	0.3		1.11 (0.79, 1.56)
No	8	8.9 (2.1)	9.0	5–12			
<i>GAD (N=84)</i>							
Yes	59	5.2 (0.9)	5.0	3–6	0.13		1.22 (0.74, 1.99)
No	25	5.0 (1.3)	6.0	3–6			
<i>PTSD (N=128)</i>							
Yes	65	17.4 (2.5)	18.0	10–20	22.3 ***		1.40 (1.20, 1.63) ***
No	63	15.2 (2.6)	15.0	9–20			

^aDefinition of AUDADIS-5 severity measures:

Major depression= count of major depression symptoms endorsed, from DSM-5 criterion A (maximum of 9)

Specific phobia=count of specific phobia situations endorsed, from DSM-5 criterion A (maximum of 17)

Social anxiety=count of social anxiety situations endorsed, from DSM-5 criterion A (maximum of 16)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Agoraphobia=count of agoraphobia situations endorsed, from DSM-5 criterion A (maximum of 6)
Panic = count of panic disorder symptoms endorsed, from DSM-5 criterion A (maximum of 13)
GAD= count of GAD symptoms endorsed, from DSM-5 criterion C (maximum of 6)
PTSD= count of PTSD symptoms endorsed, from DSM-5 criteria B, C, D, and E (maximum of 20)
(Note that persistent depression was not included since all those who received an AUDADIS-5 diagnosis [N=43] also received a PRISM-5 diagnosis.)
 χ^2 = Kruskal-Wallis χ^2 , testing for differences in the distributions of the severity indicators
 χ^2 = odds ratio; CI = Confidence interval; from logistic regression, AUDADIS-5 severity predicting PRISM-5 diagnosis status, correcting for age, sex, race, current regular substance abuse, interval between interviews.

p-values:
***<0.001