Drug and Alcohol Dependence 152 (2015) 246–256

Contents lists available at ScienceDirect

Drug and Alcohol Dependence

journal homepage: www.elsevier.com/locate/drugalcdep

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[a,b,c,*], Dvora Shmulewitz[a,c], Malka Stohl[c], Eliana Greenstein[c], Christina Aivadyan[c], Kara Morita[c], Tulshi Saha[d], Efrat Aharonovich[a,c], Jeesun Jung[d], Haitao Zhang[d], Edward V. Nunes[a,c], Bridget F. Grant[d]

[a] Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA
[b] Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032, USA
[c] New York State Psychiatric Institute, New York, NY 10032, USA
[d] Laboratory of Epidemiology and Biometry, National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD 20892, USA

ARTICLE INFO

Article history:
Received 23 January 2015
Received in revised form 16 March 2015
Accepted 27 March 2015
Available online 6 April 2015

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

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 with 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 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 ($\kappa = 0.24–0.59$), 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 epidemiological studies. The stronger concordance on dimensional measures supports the current movement toward 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.

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

☆ 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.

* Corresponding author at: Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA. Tel.: +1 646 774 7909; fax: +1 646 774 7920.
E-mail address: deborah.hasin@gmail.com (D.S. Hasin).

http://dx.doi.org/10.1016/j.drugalcdep.2015.03.027
0376-8716/

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 ($\kappa = 0.40–0.65$; 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 with 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 with 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 the 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 geographical 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 10 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 (three for each participant in the mood disorder version; five 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