Supplementary Online Content

Hasin DS, Sarvet AL, Meyers JL, et al. Epidemiology of adult *DSM-5* major depressive disorder and its specifiers in the United States. *JAMA Psychiatry*. Published online February 14, 2018. doi:10.1001/jamapsychiatry.2017.4602

eTable. Income Categories Above $70,000

eAppendix. Supplementary Material

This supplementary material has been provided by the authors to give readers additional information about their work.

eTable. Income Categories Above $70,000		
	12-Month MDD	Lifetime MDD
Income Category	Prevalence % (SE)	Prevalence % (SE)
$75,000 to $99,999	7.74 (0.54)	20.01 (0.87)
$100,000 to $149,000	7.01 (0.58)	19.09 (1.03)
$150,000 or more	7.30 (0.98)	18.93 (1.03)

eAppendix. Supplementary Material
The NESARC-III Clinical (Procedural) Validity Study:
712 blinded clinician re-appraisals of AUDADIS diagnoses.

The NESARC-III Clinical (Procedural) Validity Study has been described in detail in two previous publications[1,2]. Readers are encouraged to consult these two publications for detailed information about the study. In these two publications, "clinical validity" was termed "procedural validity", consistent with earlier publications that have used these two terms interchangeably (e.g., [3-6]). Here, we summarize the clinical (procedural) validity study briefly below for the convenience of the readers of the present article.

Background. The ability of epidemiological studies to produce clinically meaningful diagnoses is critical to their credibility and utility for clinicians, researchers, and policymakers. Since clinicians cannot be used to make diagnoses in large-scale general population studies, diagnostic procedures administered by lay interviewers must be used. Thus, clinical (procedural) validation of lay-administered psychiatric diagnostic instruments is an important element of general population surveys. One way of establishing such validity is by comparing diagnoses generated by clinician re-appraisals of a series of participants to the diagnoses generated by the lay-administered interviews. Few clinical (procedural) validation studies of psychiatric disorders have been conducted in general population samples.

In a study conducted in the early 1980s, Diagnostic Interview Schedule (DIS) diagnoses were compared to structured psychiatrist re-interviews of Epidemiologic Catchment Area participants (n=370), with psychiatrists completely blinded to DIS interviews[7]. Complete blinding of the second assessor increases the independence of the second assessment from the first. Such independence is a desirable design feature if one is interested in clinician diagnoses that were not influenced by information from the initial interview. In this study, chance-corrected lay/psychiatrist concordance ranged for major depressive disorder (κ=0.28-0.50)[7].

In studies comparing Composite International Diagnostic Interview (CIDI) to clinician-administered Structured Clinical Interview for DSM (SCID) re-interviews[8], SCID interviewers were informed of responses to CIDI gateway questions for each module, and actually reminded participants of these responses. (Informing re-test interviewers about results of the first interview and reminding participants of their responses reduces the independence of the second assessment). CIDI - SCID chance-corrected agreement on major depression disorder in this study was κ=0.56[8].

PRISM-5: the clinical validation diagnostic procedure for the NESARC-III clinical validation study. In the NESARC-III AUDADIS-5 clinical (procedural) validity study, the clinician-administered re-interview was the Psychiatric Research Interview for Substance and Mental Disorders (PRISM), DSM-5 version[1,2].

The PRISM is a semi-structured clinician-administered diagnostic interview that includes full separate modules to assess the diagnostic criteria in detail for unipolar and bipolar mood disorders, anxiety disorders (generalized anxiety, panic disorder, specific phobia, social anxiety), post-traumatic stress disorder, substance use disorders (DSM-IV abuse, dependence; DSM-5 substance use disorder) and personality disorders (e.g., antisocial, borderline). The PRISM was initially designed to overcome well-known challenges in assessing mood and anxiety disorders among drug users and heavy drinkers[9-11]. The PRISM was successful in accomplishing this purpose. Test-retest reliability of major depressive disorder was established in two studies conducted in patients who were all drug users or heavy drinkers. In these test-retest reliability studies, two interviewers conducted two separate interviews on a series of participants, with the second interviewer always completely blinded to the results of the first interview. These studies established that the PRISM had excellent test-retest reliability for current and lifetime major depressive disorder: κ=0.81 and κ=0.74, respectively[12], and κ=0.75 and κ=0.70, respectively[13]. Further, a group independent of the PRISM developers conducted a validity study using the Longitudinal Expert All Data (LEAD) procedure (see Spitzer, 1983[14] for the design of LEAD studies). In this study, the PRISM demonstrated better validity than the SCID in diagnosing major depressive disorders in substance-using patients[15].

The version of the PRISM used in the clinical (procedural) validity study was the PRISM-5. This version is computer-assisted to improve efficiency, and updated for DSM-5 criteria. To reduce participant burden for the clinical (procedural) validity study, two shortened versions of the PRISM-5 were prepared: one that included mood disorders but not anxiety disorders, and the other with anxiety disorders but not mood

disorders. These two versions were randomly assigned to participants.

Sample selection. After completing face-to-face AUDADIS-5 interviews for the NESARC-III, 777 participants were selected for the clinical validation study using an algorithm based on the AUDADIS-5 psychiatric and substance module screening questions designed to increase the prevalence of psychopathology. Participants were also required to reside within the Eastern Time Zone to facilitate telephone re-interviews from New York City, the location of the PRISM-5 interviewers. NIH, Westat and New York State Psychiatric Institute IRBs approved all procedures; all respondents gave informed consent to participate. The response rate was excellent, 92.5% (712/777).

Data collection. PRISM-5 clinical re-appraisals were conducted by telephone from a call center established in Dr. Hasin's research laboratory at New York State Psychiatric Institute, permitting centralized, closely-supervised interviewing of participants dispersed over a wide geographic area. These interviews were conducted after receiving notification from Westat that the participant's NESARC-III interview was complete. The mean interval between AUDADIS-5 and PRISM-5 interviews was 10.5 days (s.e. = 0.27); median was 9 days, and the range was 2−69 days. With consent, PRISM-5 interviews were recorded for quality assurance purposes. Of the 712 participants, 700 (98.3%) consented and their interviews were recorded.

Blinding procedures. All PRISM-5 team members (not just the interviewers) were completely blinded to AUDADIS-5 interview results, diagnoses and values in the selection algorithm. To ensure that participants knew this, and the reason for it, the PRISM-5 interviewer told each respondent at the beginning of the interview that the re-interview was to help understand the quality of the previous interview, and that he/she (the interviewer) did not have information from the previous 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.

PRISM-5 interviewers. The ten PRISM-5 interviewers for the clinical (procedural) validity all had master's degrees in clinical fields. In addition, they all had clinical experience with psychiatric and/or substance abuse patients (mean length of clinical experience, 4.15 years, range, 2-14 years).

PRISM-5 trainer/supervisors. Training on PRISM-5, study procedures and confidentiality was conducted by two PRISM trainer/supervisors (Christina Aivadyan, M.S., and Eliana Greenstein, M.A., M.P.H.). These two trainer/supervisors were co-authors of the NESARC-III clinical (procedural) validity study[1,2]. They both have a master's degree in a clinical field, clinical experience with psychiatric and substance abuse patients, and many years of experience managing research studies and training and supervising clinician- and lay-administered diagnostic interviewers for research studies in clinical and general population settings in the U.S. and elsewhere (e.g., Spain, Norway, Israel).

PRISM-5-training/certification. After reviewing a training manual and answering self-study questions, the PRISM-5 trainer/supervisors conducted 4 days of PRISM-5 training that included didactic presentation of material (covering confidentiality, how to engage participants, diagnostic criteria, managing the computer), and extensive group role-plays to ensure that interviewers were familiar with the interview and also with techniques to conduct the interview efficiently while making participants feel at ease during the interview. Trainees were allowed to begin interviewing for the clinical validity study only after they were certified by the study trainer/supervisors. To certify trainees, the trainer/supervisors listened to five PRISM-5 interviews, and rated their quality using structured interview quality rating forms. Interview quality was required to be satisfactory or better. Thus, training was structured and rigorous.

PRISM-5 quality assurance. During the course of the study, the trainer/supervisors used the same structured interview quality rating forms to rate recordings from 214 randomly selected PRISM-5 interviews for quality assurance, using the information from these reviews to provide feedback to interviewers in regular group supervision meetings. Thus, quality assurance procedures were extensive. In addition, two psychiatrists 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 59 cases, 234 relevant mood and anxiety disorder diagnoses were possible. Of these mood and anxiety diagnoses, psychiatrists and supervisors agreed on 95.3% of them; 1.3% were made by a psychiatrist but not a supervisor, while 3.4% were made by a supervisor but not a psychiatrist. Thus, the excellent agreement between the two study trainer/supervisors and the psychiatrists further attests to the excellent level of supervision/quality control of the clinical (procedural) validity study.

References

1. Hasin DS, Greenstein E, Aivadyan C, et al. 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.
2. Hasin DS, Shmulewitz D, Stohl M, et al. Procedural validity of the AUDADIS-5 depression, anxiety and post-traumatic stress disorder modules: Substance abusers and others in the general population. *Drug Alcohol Depend.* 2015;152:246-256.
3. 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(2):95-110.
4. 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(6):1169-1177.
5. Wittchen HU. Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): a critical review. *J Psychiatr Res.* 1994;28(1):57-84.
6. Peters L, Andrews G. Procedural validity of the computerized version of the Composite International Diagnostic Interview (CIDI-Auto) in the anxiety disorders. *Psychol Med.* 1995;25(6):1269-1280.
7. Helzer JE, Robins LN, McEvoy LT, et al. A comparison of clinical and diagnostic interview schedule diagnoses. Physician reexamination of lay-interviewed cases in the general population. *Arch Gen Psychiatry.* 1985;42(7):657-666.
8. Haro JM, Arbabzadeh-Bouchez S, Brugha TS, et al. 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(4):167-180.
9. Kadden RM, Kranzler HR, Rounsaville BJ. Validity of the distinction between "substance-induced" and "independent" depression and anxiety disorders. *Am J Addict.* 1995;.4(2):pp.
10. Bryant KJ, Rounsaville B, Spitzer RL, Williams JB. Reliability of dual diagnosis. Substance dependence and psychiatric disorders. *J Nerv Ment Dis.* 1992;180(4):251-257.
11. Weiss RD, Mirin SM, Griffin ML. Methodological considerations in the diagnosis of coexisting psychiatric disorders in substance abusers. *Br J Addict.* 1992;87(2):179-187.
12. 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(9):1195-1201.
13. 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(4):689-696.
14. Spitzer RL. Psychiatric diagnosis: are clinicians still necessary? *Compr Psychiatry.* 1983;24(5):399-411.
15. 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(7):1231-1237.