

Radu Vrsti^a
Bridget F. Grant^b
Somnath Chatterji^{c,d}
Bedirhan T. Üstün^d
Doug Mager^e
Ioan Olteanu^a
Marina Badoi^a

^a Psychiatric Hospital Jebel, Research Department, Jebel, Romania,

^b Division of Biometry and Epidemiology, National Institute on Alcohol Abuse and Alcoholism, Bethesda, Md., USA,

^c National Institute on Mental Health and Neurosciences, Bangalore, India,

^d Epidemiological and Managerial Support Division, Mental Health and Prevention of Substance Abuse, World Health Organization, Geneva, Switzerland,

^e Washington University School of Medicine, Department of Psychiatry, St. Louis, Mo., USA

Reliability of the Romanian Version of the Alcohol Module of the WHO Alcohol Use Disorder and Associated Disabilities: Interview Schedule – Alcohol/Drug-Revised

Abstract

Alcohol Use Disorder and Associated Disabilities Interview Schedule – Alcohol/Drug-Revised (AUDADIS-A/D-R) is a fully structured, standardized and precoded instrument designed to evaluate alcohol and drug use disorders according to DSM-III-R, DSM-IV, and ICD-10 criteria. The AUDADIS-A/D-R has shown good to excellent reliability in both large clinical and general population samples, but prior to the conduct of the present study no data on the reliability of the Romanian version of the AUDADIS-A/D-R existed. The purpose of the present study was to examine the test-retest reliability of the alcohol module of the AUDADIS-A/D-R in a general population and clinical sample in Romania. The overall reliability of ICD-10 and DSM-IV abuse, harmful and dependence diagnoses, was found to be good to excellent, but was somewhat lower for abuse and harmful use diagnoses. The results are discussed in terms of the cultural applicability of the symptom items and within the context of the analysis of discrepant responses between the test and retest interviews.

Key Words

Reliability
Alcohol assessment
Psychiatric research

During the past decade, the reliability and validity of diagnostic interviews have been the focus of a great deal of studies. The criterion-based nosologies, like the Diagnostic and Statistical Manual of Mental Disorders of the American Psychiatric Association, 3rd and 4th editions (DSM-III-R and DSM-IV, respectively) [1, 2], and the 10th revision of the International Classification of Diseases (ICD-10) [3], require a special structure of those instruments that are devoted to eliciting proper answers to questions that load on a diagnostic criterion.

The World Health Organization/US National Institutes of Health Joint Project on Diagnostic and Classifica-

tion of Mental Disorders, Alcohol- and Drug-Related Problems was one of the largest international endeavors designed to develop reliable, valid and cross-culturally applicable instruments for assessment of alcohol and drug use disorders [4]. The first phase of this project examined the cross-cultural applicability of the alcohol and drug sections of three assessment instruments: the Composite International Diagnostic Interview [5, 6]; the Schedules for Clinical Assessment in Neuropsychiatry [7, 8], and Alcohol Use Disorder and Associated Disabilities Interview Schedule – Alcohol/Drug-Revised (AUDADIS-A/D-R) [9]. The second phase of this project focused on the

reliability and validity of the three instruments and the overall multisite results of this phase have recently been published [10–15].

AUDADIS-A/D-R is an instrument designed to evaluate alcohol and drug use disorders according to DSM-III-R, DSM-IV, and ICD-10 criteria. Two recent large test-retest studies in a general population sample [16] and clinical sample [17] have shown good to excellent reliability for measures of alcohol and drug use related diagnosis of abuse, harmful use, and dependence.

The present study focuses on the Romanian version of the alcohol module of the AUDADIS-A/D-R. The major objective of this study was to examine the test-retest reliability of the AUDADIS-A/D-R at the criterion and diagnostic level of alcohol use disorders defined by ICD-10 and DSM-IV.

Method

Study Design

The present study used a test-retest design. Each subject was interviewed twice with a time interval of 3 days to 3 weeks and a mean duration of 14 days. The mean duration of interview administration was 87 min. Different interviewers administered test and retest interviews so that no subject was interviewed twice by the same interviewer. In addition, each interviewer remained blind to the results of the other interviewers as required by the test-retest design.

Other procedures used in this study included the use of the Discrepancy Interview Protocol (DIP) [18]. The DIP was developed to assess possible reasons why responses to preselected questions appearing on the AUDADIS-A/D-R might differ at test and retest. After the second interview, the interviewer would identify discrepant responses and reasons why certain questions were answered differently during each interview. There were three different sets of 9–12 questions selected for the DIP interview which were each administered to approximately one-third of the sample.

The Debriefing Questionnaire (DQ) was used to assess the interview process as a whole, highlighting those characteristics of respondents and investigators that could jeopardize the reliability of the verbal reports of a subject. The DQ contained two parts: one applied to respondents and encompassing questions about respondents' opinions about the interview (structure, length, understandability, offensiveness, proposing additional questions, and general feeling about it), and another for interviews that rated the respondents' behavior and opinion just after the interview session. The DQ was completed by each interviewer after every AUDADIS-A/D-R session, that is after each test and retest.

Sample

Approximately 25% (n = 44) of the 149 subjects were systematically drawn from specialized alcohol and drug treatment units of the Psychiatric Hospital Jebel, a 500-bed general hospital with an active and dynamic research department involved in several international multisite studies. The remaining 75% (n = 105) of the sample were patients from general or primary care facilities without a clinical his-

Table 1. Subject characteristics

Subjects	149
Male, %	68
Mean age, years	39.91
Marital status, %	
Never married	12
Married/living with	73
Divorced/separated	11
Widowed	4
Mean years of education	12.46
Recruitment setting, %	
General practitioner's office	50
Community mental health center	2
Psychiatric inpatient ward	30
Psychiatric outpatient clinic	12
Rehabilitation institution	5
Other	1

tory of alcoholism. Recruiting samples from both alcohol settings and nonalcohol settings were implemented to obtain an overall sample with a broad range of abuse and dependence symptomatology necessary for the reliability of the statistical analyses. Characteristics of responders appear in table 1. Those subjects who had cognitive impairment or severe behavioral disturbances were eliminated. Because lifetime abstainers could not provide useful information about alcohol and drug symptoms, they were excluded from the general population through a screening procedure leaving only lifetime alcohol users.

Interview Characteristics and Training

Four female interviewers who had no experience with psychiatric patients administered the AUDADIS-A/D-R in this study. The mean age of these interviewers was 26 years with an average of 13 years of education. At least one principal investigator from each site participated in the same extensive 1-week training session on the administration of the AUDADIS-A/D-R, including mock interview sessions, reviewing interview specifications and videotapes and supervised interviews with drug and alcohol patients. Training on the AUDADIS-A/D-R at each site also entailed several practice interviews, quality control and standardization assessment. This training was then replicated at each site with the principal investigators serving as trainers. Prior to the training sessions the AUDADIS-A/D-R was translated and back-translated by an expert bilingual group at the study site. Moreover, the Romanian version of AUDADIS-A/D-R was back-translated by an independent expert at the US National Institute of Alcohol Abuse and Alcoholism (NIAAA), Rockville, Md. Although translations agreed well, minor discrepancies were resolved jointly between each site and the NIAAA.

Diagnostic Assessment

The AUDADIS-A/D-R is a fully structured instrument designed to be administered by nonclinicians or clinicians. The precoded response categories facilitate fully computerized diagnostic analysis.

The AUDADIS-A/D-R assesses harmful use, abuse and dependence diagnoses for alcohol and drugs.

Generally, diagnostic questions associated with alcohol use disorders are asked of all respondents drinking at least 12 drinks during any 1-year period of their lives. In order to achieve an ICD-10 or DSM-IV diagnosis of alcohol dependence, three dependence criteria relevant to each classification must have been positive during the year preceding the interview. To correspond to a diagnosis of dependence, the three dependence criteria must have been met with the additional qualification that they occurred: (1) most days for at least 1 month; (2) repeatedly for a few months or longer, or (3) around the same time. Harmful use and abuse diagnoses were similarly constructed, but in concert with the diagnostic definitions, one or more of the harmful use or abuse criteria must have been met within each time frame in order to achieve a positive diagnosis. Subjects were classified with a lifetime harmful use, abuse or dependence diagnosis if they met criteria for the relevant diagnosis in the past year and/or prior to the past year.

Statistical Analysis

In this study, kappa coefficient of reliability was used to measure interrater agreement corrected for chance. Kappa has a range from -1 to +1, and a kappa of 0.0 indicates agreement at the chance level. According to Fleiss [19], kappas of 0.75 and above indicate excellent

agreement, from 0.65 to 0.74 good agreement, from 0.40 to 0.64 fair agreement and less than 0.40 poor agreement. Although straightforward, kappa is influenced by prevalence. For this reason, base rates for each criterion or diagnosis were calculated in presented.

Results

Table 2 shows the kappas and base rates of ICD-10 and DSM-IV diagnostic criteria for alcohol use disorders for each time frame: past year and lifetime. With the exception of the withdrawal criterion, the reliability of all diagnostic criteria for harmful use and dependence syndromes (ICD-10) and for abuse and dependence (DSM-IV) was good to excellent, kappas ranging from 0.66 to 0.85 for both high-prevalence and low-prevalence diagnostic criteria. Reliability of withdrawal criterion was fair for ICD-10 as well as DSM-IV alcohol dependence in both time frames, kappas ranged from 0.62 for lifetime frame to 0.64 for past year frame.

Table 2. Reliability of ICD-10 and DSM-IV diagnostic criteria for alcohol use disorders among users by time frame

International classificatory system criterion	Past year		Lifetime			
	kappa	base rate		kappa	base rate	
		1st	2nd		1st	2nd
<i>ICD-10</i>						
Harmful use	0.79 (0.06)	0.51	0.50	0.73 (0.06)	0.65	0.63
Dependence						
Compulsion	0.76 (0.06)	0.55	0.53	0.74 (0.06)	0.63	0.61
Impaired control	0.62 (0.08)	0.79	0.73	0.68 (0.09)	0.85	0.82
Withdrawal	0.64 (0.08)	0.80	0.73	0.62 (0.09)	0.85	0.83
Tolerance	0.74 (0.06)	0.41	0.36	0.71 (0.06)	0.51	0.48
Neglect interest	0.78 (0.06)	0.36	0.40	0.68 (0.06)	0.45	0.47
Continued use	0.79 (0.06)	0.51	0.50	0.73 (0.06)	0.65	0.63
<i>DSM-IV</i>						
Abuse						
Role obligation	0.77 (0.06)	0.37	0.33	0.75 (0.06)	0.53	0.48
Hazardous use	0.80 (0.06)	0.25	0.24	0.75 (0.06)	0.38	0.35
Legal problems	0.85 (0.15)	0.03	0.02	0.70 (0.14)	0.05	0.05
Continued use	0.66 (0.07)	0.46	0.45	0.66 (0.07)	0.58	0.55
Dependence						
Tolerance	0.74 (0.06)	0.41	0.36	0.71 (0.06)	0.51	0.48
Withdrawal	0.64 (0.08)	0.80	0.73	0.62 (0.09)	0.85	0.83
Impaired control	0.69 (0.07)	0.55	0.53	0.71 (0.07)	0.76	0.69
Compulsion	0.68 (0.07)	0.71	0.65	0.71 (0.07)	0.74	0.70
Time involved	0.73 (0.08)	0.21	0.20	0.66 (0.07)	0.31	0.27
Important activities	0.85 (0.05)	0.29	0.31	0.79 (0.06)	0.32	0.37
Continued use	0.70 (0.06)	0.46	0.42	0.72 (0.06)	0.56	0.50
SE in parentheses.						

Table 3 shows the reliability of diagnoses of harmful use, abuse and dependence. Kappas for harmful use in the past year and for abuse in all time frames were fair, ranging between 0.38 and 0.58, and poor for prior to past year and lifetime harmful use. In contrast to the results for abuse and harmful use, kappas for dependence ranged from good to excellent (0.66–0.79) for both classification systems.

The reasons for discrepancies between responses provided during test and retest are shown in table 4. There were 770 questions checked which produced 123 discrepancies (15.9%), almost 0.85 discrepancies per subject, on average. Among them, 59% were attributed to the respondent, 34% to the assessment (interview), and 6% to the interviewer. Surprisingly, patients (clinical sample) reported more discrepancies (20.6%) than those from the general population (14.0%), but with respect to reasons no differences were found for the comparison of clinical (patients) and nonclinical (users from general population) subjects.

Questions with the highest rate of discrepancies included the number of drinks consumed in a day in the last 12 months, the usual number of drinks having much less effect than it once did, spending a lot of time drinking or hungover, and having a strong desire or need to drink.

During the entire study, 144 subjects completed the DQ after the AUDADIS-A/D-R session test interview and 140 subjects completed the same questionnaire after the retest. The AUDADIS-A/D-R was judged by the majority of respondents as being appropriate concerning length of interview (90.7%), with very few finding the questions offensive or hard to understand (table 5). The interviewers considered that the AUDADIS-A/D-R was

easy to administer (64.4% very easy, 32.0% rather easy, 3.6% rather difficult) and felt the instrument elicited reliable answers (63.7% very reliable, 30.2% rather reliable, 4.2% rather unreliable, 1.9% very unreliable). Respondents seemed to be generally interested in the interview (96.8%) and appeared not to be offended by the questions (98.6%).

Table 3. Reliability of ICD-10 and DSM-IV diagnoses of harmful use, abuse, and dependence (syndrome) among users by time frame

Classification time frame	Kappa	Base rate	
		1st	2nd
<i>ICD-10</i>			
Harmful use			
Past year	0.48 (0.22)	0.02	0.04
Prior to past year	0.17 (0.14)	0.05	0.02
Lifetime	0.38 (0.20)	0.04	0.04
Dependence			
Past year	0.71 (0.07)	0.64	0.60
Prior to past year	0.75 (0.06)	0.39	0.39
Lifetime	0.73 (0.06)	0.66	0.64
<i>DSM-IV</i>			
Abuse			
Past year	0.45 (0.13)	0.09	0.11
Prior to past year	0.58 (0.08)	0.33	0.27
Lifetime	0.49 (0.10)	0.15	0.15
Dependence			
Past year	0.66 (0.07)	0.63	0.58
Prior to past year	0.79 (0.05)	0.36	0.36
Lifetime	0.68 (0.06)	0.64	0.58
SE in parentheses.			

Table 4. Frequency distribution of reasons for discrepancies – AUDADIS test-retest

Sample	Checked	Discrepancies		Reasons for discrepancies					
				interviewer		respondent		assessment	
				n	%	n	%	n	%
Alcohol users (clinical sample – patients)	228	47	20	3	6	28	59	16	34
Alcohol users (nonclinical – general population)	542	76	14	5	6	45	59	26	34
Total sample	770	123	16	8	6	73	59	42	34

Interviewer: Interviewer was wrong, answer miscoded/misunderstood.

Respondent: I couldn't remember the answer; Since first interview, symptoms/substance use changed; Not paying attention; Said 'no' to shorten the interview; Too embarrassed to give correct answer; Thought the interviewer would disapprove; I didn't know the right answer.

Assessment: I didn't understand the question; Questions were different.

Table 5. Respondents' and interviewers' opinions about interview process

<i>Respondent (284 DQ completed)</i>	
Enjoy being interviewed	
Yes, %	98.7
Length of interview	
Too long, %	2.8
Just right, %	90.7
Questions hard to understand	
Yes, %	2.1
Offensive questions	
Yes, %	3.6
Need additional questions on alcohol	
Yes, %	7.1
<hr/>	
<i>Interviewer (284 DQ completed)</i>	
Respondent is interested in the interview	
No, %	3.2
Unclear or confusing questions	
Yes, %	3.6
Respondent appears offended by questions	
Yes, %	1.4
Questions that subject refused to answer	
Yes, %	0.7
How difficult to administer	
Very easy, %	64.4
Rather easy, %	32.0
Rather difficult, %	3.6
How reliable are subject's answers	
Very reliable, %	63.7
Rather reliable, %	30.2
Rather unreliable, %	4.2
Very unreliable, %	1.9

Discussion

The overall reliability of the Romanian version of the AUDADIS-A/D-R was good to excellent for most diagnostic criteria and diagnoses of dependence, regardless of time frame examined. In contrast, the reliability of diagnoses of harmful use (ICD-10) and abuse (DSM-IV) was lower, which may reflect the hierarchical nature of the diagnosis rather than the unreliability of its criteria. These results are consistent with those of recent studies concerning the reliability of the AUDADIS in other countries [10–15]. Although the reliability of ICD-10 and DSM-IV alcohol use disorder diagnostic criteria was good to excellent, reliability for the withdrawal criterion was only fair. One explanation for this finding is that this criterion was found to be more culturally inappropriate than other diagnostic criteria in the cross-cultural applicability phase of the joint project [10, 16, 17].

The DIP data showed that, on average, each subject reported less than one question answered differently from one interview to the next. This is a very low rate of discrepancies. The greatest proportion of discrepancies (59%) were related to the respondent (e.g., did not remember the questions, not pay attention, etc.). This suggests that the instrument was generally appropriate and capable of eliciting the proper answers at test and retest, or in other words, it showed reproducibility of measurement. Analyzing those questions associated with the greatest number of discrepancies, our results show that the questions which related to tolerance, withdrawal, and compulsion to use criteria yielded more discrepancies than others (data not shown). For these criteria the high discrepancy rate resulted from not understanding the question, not remembering the answers, and not paying attention. Again, those criteria found to be most culturally inapplicable during the first phase of the Joint Project [4] were associated with the greatest number of discrepancies.

The reliability results of this study in conjunction with the analysis of the assessment process suggest that the Romanian version of AUDADIS-A/D-R, alcohol module, may be used with confidence in clinical and general populations for the assessment of alcohol use disorders in the Romanian culture. Future analyses of the Joint Project reliability data in Romania will include more detailed analyses of the relationship between reliability and discrepancies in addition to a more disaggregated analysis of reliability results by important sociodemographic characteristics.

References

- 1 American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, ed 3, Rev (DSM-III-R). Washington, APA, 1981.
- 2 American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, ed 4, Rev (DSM-IV). Washington, APA, 1994.
- 3 World Health Organization: The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research. Geneva, World Health Organization, 1993.
- 4 Room R, Janca A, Bennet LA, Schmidt L, Sartorius N: WHO cross-cultural applicability research on diagnosis and assessment of substance use disorders: An overview of methods and selected results. *Addiction* 1996;91:199-220.
- 5 World Health Organization: Composite International Diagnostic Interview (CIDI), core version 1.0. Geneva, World Health Organization, 1990.
- 6 Robins LN, Wing JK, Wittchen HU, Helzer JE, Babor TF, Burke J, Farmer A, Jablensky A, Pickens R, Regier DA, Sartorius N, Towle LH: The Composite International Diagnostic Interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and terms and in different culture. *Arch Gen Psychiatry* 1988;45:1069-1077.
- 7 World Health Organization: Schedules for Clinical Assessment in Neuropsychiatry. Geneva, World Health Organization, 1992.
- 8 Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, Jablenski A, Regier D, Sartorius N: SCAN: Schedules for clinical assessment in neuropsychiatry. *Arch Gen Psychiatry* 1990; 47:589-593.
- 9 Grant BF, Hasin DS: The Alcohol Use Disorders and Associated Disabilities Interview Schedule. Rockville, National Institute on Alcohol and Alcoholism, 1991.
- 10 Chatterji S, Saunders J, Vrsti R, Grant BF, Hasin D, Mager D: Reliability of the alcohol and drug modules of the Alcohol Use Disorder and Associated Disabilities Interview Schedule-Alcohol/Drug-Revised (AUDADIS-A/D-R): An international comparison. *Drug Alcohol Depend* 1997;47:171-185.
- 11 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 criteria and diagnoses as measured by the AUDADIS-A/D-R, CIDI, and SCAN. *Drug Alcohol Depend* 1997;47:195-205.
- 12 Easton C, Maza E, Mager B, Ulug B, Kilic C, Gogus A, Babor T: Test-retest reliability of the alcohol and drug use disorder sections of the Schedules for Clinical Assessment in Neuropsychiatry (SCAN). *Drug Alcohol Depend* 1997;47:187-194.
- 13 Hasin D, Grant BF, Cottler L, Blaine J, Towle L, Ustun B, Sartorius N: Nosological comparisons of alcohol and drug diagnoses: A multi-site, multi-instrument study. *Drug Alcohol Depend* 1997;47:217-226.
- 14 Pull CB, Saunders JB, Mavreas V, Cottler LB, Grant BF, Hasin DS, Blaine J, Mager D, Ustun B: Concordance of between ICD-10 alcohol and drug use disorder criteria and diagnoses as measured by the AUDADIS-A/D-R, CIDI and SCAN: Results of a cross-national study. *Drug Alcohol Depend* 1997;47:207-216.
- 15 Ustun B, Compton W, Mager T, Babor T, Baiyewu O, Chatterji S, Cottler L, Gogus A, Maureas V, Peters L, Pull C, Saunderson JB, Smoets R, Rubio-Stipec M, Vrsti R, Hasin D, Room R, Van der Brink W, Regier D, Blaine J, Grant BF, Sartorius N: WHO study on the reliability and validity of the alcohol and drug disorder instruments: Overview of methods and results. *Drug Alcohol Depend* 1997;47:161-169.
- 16 Grant BF, Harford TC, Dawson DA, Chou PS, Pickering RP: The alcohol use disorder and associated disabilities interview schedule (AUDADIS): Reliability of alcohol and drug modules in a general population sample. *Drug Alcohol Depend* 1995;39:37-44.
- 17 Hasin DS, Carpenter KM, McCloud S, Smith M, Grant BF: The Alcohol Use Disorder and Associated Disabilities Interview Schedule: Reliability of alcohol and drug modules in a clinical sample. *Drug Alcohol Depend* 1996; 11:199-220.
- 18 Cottler LB, Compton WM, Brown L, Shell A, Keating S, Shillington A, Hummel R: The discrepancy interview protocol: A method for evaluating and interpreting discordant survey responses. *Int J Methods Psychiatr Res* 1994;4: 173-182.
- 19 Fleiss JL: *Statistical Methods for Rates and Proportions*. New York, Wiley, 1981.