CONCURRENT AND CONSTRUCT VALIDITY OF THE AUDIT IN AN URBAN BRAZILIAN SAMPLE

CARLOS TADEU LIMA1,2,*, ANTONIO CARLOS C. FREIRE3, ANDREZZA PAULA B. SILVA3, RODRIGO MAIA TEIXEIRA3, MICHAEL FARRELL2 and MARTIN PRINCE2

1State University of Feira de Santana, Brazil, 2Institute of Psychiatry, Kings College of London, UK and 3School of Medicine and Public Health of Bahia, Brazil

(Received 4 May 2005; first review notified 11 July 2005; in revised form 5 August 2005; accepted 8 August 2005; advance access publication 5 September 2005)

Abstract — Aims: To assess the concurrent and the construct validity of the Alcohol Use Disorders Identification Test (AUDIT) in an urban Brazilian sample. Methods: A random sample of 166 clients of a health management organization, participated in this study. They were visited in their households and completed a self-report questionnaire, which included the AUDIT. Later, they answered the alcohol-related disorders (ARDs) Section of the Composite International Diagnostic Interview. The receiver operating curve (ROC) was used to find the best cut-off point for ICD-10 diagnosis of ARDs. Confirmatory factor analysis was run to assess the construct validity. Results: The ROC analysis showed the same cut-off point (7/8) for ICD-10 diagnosis of ARDs found in previous studies carried out in primary care settings, including in Brazil, with a sensitivity of 100% and a specificity of 76%. The confirmatory factor analysis suggested a two-factor structure. The first factor measured consumption and the second factor alcohol-related problems. Conclusions: The results supported the use of the self-reported version of the AUDIT in epidemiologic studies, and showed a similar cut-off point for detection of ARDs and hazardous drinking.

INTRODUCTION

Alcohol-related disorders (ARD) are the second most prevalent group of mental disorders in Brazil, after anxiety disorders (Almeida Filho et al., 1997). The prevalence of alcohol dependence in the adult population ranges from 8.0 to 9.2% in the biggest cities in the country, but a national survey including 107 cities found a rate of 11.2% (Carlini et al., 2002). The first study was a two-phase survey where a screening instrument, the QMPA (Questionnaire of Psychiatric Morbidity in Adults), was used in the first phase and the Inventory of symptoms of DSM-III was used in the second phase (Almeida Filho et al., 1997). In the second study, the instrument used was the SAMHSA (Substance Abuse and Mental Health Services Administration), which provides diagnosis of substance dependence according to DSM-III-R criteria (Carlini et al., 2002).

Another instrument previously used in Brazilian population-based studies of ARD is the CAGE (Lima et al., 1999). The CAGE was previously validated among male inpatients of a Brazilian psychiatric hospital, comparing alcohol-dependent with non-alcoholic patients. In this study, it demonstrated a sensitivity of 88% and a specificity of 83% (Masur and Monteiro, 1983). Despite its use in epidemiological research in Brazil, the CAGE has not been validated for use in the general population in Brazil. For this purpose, it tends to have lower sensitivity when compared to its performance in primary care settings (Chan et al., 1994; Cherpitel, 1998). The CAGE does not provide a diagnosis of ARD, rather it measures drink-related problems (Mayfield et al., 1974). The MAST (Michigan Alcoholism Screening Test) has been used in clinical settings in Brazil. It is intended only to screen for alcohol dependence (Fortes and Cardo, 1991), for which task it is said to be more specific than the CAGE, but only detects the most severe and insightful cases of dependence (Buchsbaum et al., 1992).

None of the above-mentioned instruments is in agreement with the most recent WHO guidelines for early screening and treatment of alcohol-related problems in primary care, which suggest the need to include a broader range of clinical conditions than alcohol dependence and problem-drinking (Babor et al., 2001). All alcohol-related problems and behaviours, from heavy, hazardous, and binge drinking to harmful drinking and alcohol dependence should be considered for secondary prevention (Babor et al., 1986), with brief targeted interventions where problem-drinking is beginning and dependence is not yet established (Saunders et al., 1993). These conditions cannot be screened using the CAGE or the MAST because they do not discriminate the past from current events and they do not measure alcohol consumption (Barry and Fleming, 1993).

The Alcohol Use Disorders Identification Test (AUDIT) was created to address the early detection of cases (Babor et al., 2001). It was developed from a 150-item interview applied in a sample of 1888 patients of primary care services in Australia, Bulgaria, Kenya, Mexico, Norway, and USA. The choice of the final 10 questions was based on the internal consistency of the scale, and on factor analysis. At the end of the process, face validity was considered in order to assure that the major conceptual domains were included (alcohol consumption, dependence, and alcohol-related problems) (Saunders et al., 1993). Given its cross-cultural development, this instrument covers a variety of drinking cultures and their drinking behaviours. The AUDIT is currently the fourth most commonly used instrument for the detection of alcohol-related disorders and is already validated in several countries and different settings (Allen et al., 1997). It has also been used in Brazil in clinical (Fligie et al., 2000) and population studies (Mendoza-Sassi and Beria, 2003), and has been validated recently in a Brazilian primary care facility (Mendez, 1999). As yet, the AUDIT has not been validated in a general...
population sample and this will allow a better understanding of population-based studies including this instrument. Moreover, Reinerd and Allen (2002) remind us that the psychometric properties of the AUDIT in non-English speaking countries are not well known. The main aim of this study is to validate the AUDIT in a Brazilian urban sample.

METHODS

In a sample of predominantly middle class clients of a health maintenance organization (HMO), we assessed the reliability of the AUDIT and its concurrent validity against the ICD-10 diagnoses provided by the CIDI (Composite International Diagnostic Interview), version 2.1. (WHO, 1997) and its construct validity using factorial analysis.

Participants

A random sample of 259 households was established from a register of clients of an HMO in Salvador, a city in the northeast of Brazil. The clients were contacted first by a letter explaining to them that they were included in the sample for research purposes and stating the general aims of the study. This was followed by a telephone contact to book an interview with all the clients aged 18–60 years, who lived at those addresses.

Procedures

All the interviews took place in the participants’ homes. The interviewer first asked each participant to complete a self-reported questionnaire, which included the AUDIT, among other questions about demographic variables, occupational characteristics, health, and mental health. After this, the respondents underwent a structured interview including the section of CIDI 2.1 pertaining to alcohol-related disorders, in a setting of their choice, where they could not be overheard. The six interviewers were trained in the procedures of both instruments by the main author of this study, who had been trained in the procedures of CIDI 2.1 at the Institute of Psychiatry, Kings College of London, the training centre in United Kingdom for CIDI.

Measures

The AUDIT. The AUDIT was developed by WHO as a screening instrument for alcohol-related disorders in general. Its 10 questions refer to the previous 12 months and address alcohol dependence (questions 4–6), harmful consumption (questions 7–10), both as defined by ICD-10, and hazardous alcohol consumption, considered as a threshold of consumption which predicts future harm (questions 1–3). The maximum score of the AUDIT is 40 (Babor, et al., 2001). In the original study, using a cut-off point of 7/8 its sensitivity was 92%, and its specificity 94%, pooling the results across countries. The gold standard was clinically diagnosed patients with hazardous or harmful alcohol use, defined broadly as one or more of the following: hazardous daily consumption; recurrent intoxication; at least one of the criteria for alcohol dependence in the ICD-10 classification; at least one alcohol-related problem in the previous year; an alcohol-related disease or a perceived drinking problem (Saunders et al., 1993). The main goal of the AUDIT is to screen for hazardous and harmful alcohol use. However, there is already some evidence that its total score reflects a meaningful continuum with increasing scores from the abstainers, normal, hazardous, and harmful drinking, with alcohol dependence achieving the highest mean score (Volk et al., 1997).

The AUDIT was previously validated in a Brazilian primary care facility, in a low-income sample, with low educational level and 20% of illiterate people. The authors used a back-translated version administered by an interviewer. They found a sensitivity of 91.8%, a specificity of 62.3%, and a positive predictive value of 34.9%, for a cut-off point of 7/8, against a semi-structured psychiatric interview addressing ICD-10 criteria (Mendez, 1999). However, the factor structure of the instrument was not analysed.

In our study, the AUDIT was used in self-report format. The questionnaire was translated directly from the instrument proposed by WHO taking into consideration existing Portuguese translations available (Mendez, 1999; Fligie et al., 2000). All versions were submitted to another Brazilian bilingual psychiatrist for review. Minor changes were necessary to adjust wording for expressing drinking behaviour according to the culture of the region and educational level of the sample.

The CIDI. The CIDI is a fully structured and standardized interview, recommended to be applied by lay interviewers (Wittchen et al., 1999). For concurrent validation, given the time frame reference of the AUDIT, we used the 12-month version of the alcohol-related disorders section of the CIDI to provide ICD-10 diagnoses of alcohol dependence and harmful use (WHO, 1997). In Brazil, the CIDI has showed an excellent inter-rater reliability for harmful use (κ = 0.88) and dependence (κ = 0.97) in the previous 12 months (Quitana et al., 2000).

Heavy drinking. The average amount of alcohol consumed per week was calculated using the CIDI questions of quantity and frequency of consumption. Similar procedures were performed with the first two questions of the AUDIT, using the midpoint of the range in the response options, e.g. a range of 30–40 g per occasion meant 35 g and 2–3 drinking occasions per week meant 2.5. Heavy drinking was determined utilizing CIDI data using the threshold of 21 g of alcohol/per day for men and 14 g/per day for women (Reid et al., 1999).

Statistical analyses

The inter-item correlations and Cronbach’s α coefficients were calculated as measures of the internal consistency of the AUDIT. The intraclass correlation between the weekly average intake of alcohol calculated by the CIDI and by the AUDIT data was established as a measurement of reliability (Streiner and Norman, 1995).

To assess concurrent validity, the scores of the AUDIT were compared with the ICD-10 diagnoses provided by the alcohol-related disorders module of the CIDI for the 12-month version, considered here as the gold standard. A positive case was defined as the individual who was diagnosed as having either alcohol harmful use or alcohol dependence. The area under the receiver operating curve (ROC), with the correspondent 95% CI, was estimated for the ICD-10 alcohol-related disorder criterion. The optimal cut-off point was identified and sensitivity,
specificity, positive, and negative predictive values were calculated.

A principal component analysis (PCA) of the AUDIT was performed, in an exploratory approach, using varimax rotation to identify a more suitable solution for the observed data. The Kaiser criteria (eigenvalue >1) was used as a guide as to how many factors should be retained. The results of the PCA were submitted to confirmatory factor analysis in order to measure their goodness-of-fit as well as to compare with the other models of internal structure found in the literature. Apart from the WHO three-dimension model, the alternative models found, and included in this comparison, were a single-factor model (consistent with the AUDIT scoring method), and a two dimension model proposed first by Medina-Mora et al. (1998). The Medina-Mora model suggests that the AUDIT first factor is composed by the items related to consumption level (question 1–3) and the second factor is composed by the remaining questions, covering the problems and consequences related to consumption. The Akaike Information Criteria (AIC) and the Pearson correlation between factors were calculated. The best model was the one that produces the minimum value for the AIC, with high but not perfect Pearson correlation between factors (Dunn et al., 1993). The use of the PCA and confirmatory factor analysis are considered as an assessment of construct validity.

RESULTS

From the 259 households contacted, just 92 (35.5%) accepted to take part in this study. The final sample included in the study had 166 individuals. The great majority of the sample was white (59.3%), followed by those with mixed ancestry (37.0%). Most participants were under 30 years old (38.6%) followed by those 51 years old and above (24.0%). There was a predominance of females (59%) in the group studied. About 63.9% of the participants were employed at the time of the data collection; 81.6% reported earning at least ten times more than the legal minimum salary as family income (Table 1).

Drinking behaviour

Total abstinence in the past year was reported by 33.7%, but the proportion was higher among women (38.8%) than among men (26.5%). The total prevalence of heavy drinkers was 9.0%, 17.6% among men and 3.1% among women. The intra-class correlation coefficient between the weekly quantity of alcohol consumed measured by the AUDIT (self-report) and the CIDI (interviewer administered) was 0.80 (CI 95% 0.70–0.87).

Four cases of harmful use and five cases of alcohol dependence were identified by CIDI according to ICD-10 criteria. The prevalence of any alcohol-related disorder according to ICD-10 was 5.4%. One heavy drinker, diagnosed with harmful use, did not answer the AUDIT.

The comparison of the AUDIT total score and the CIDI, using the criterion of any ICD-10 alcohol-related disorder produced an area under the curve of 0.94 (CI 95% 0.87–0.99). For this criterion, the optimal cut-off point for the sensitivity and specificity suggested by the curve was 7/8 (Figure 1), yielding a 28% prevalence of AUDIT-positive cases in this sample (n = 33). For any ICD-10 alcohol-related disorder, the sensitivity was 100%, as all eight ARD cases were detected by the AUDIT. From the 111 subjects not recognized by the CIDI as having any ARD, 86 were also AUDIT-negative, so the specificity was 76%, and the positive and negative predictive values were 24 and 100%, respectively. Kappa, for the agreement between AUDIT caseness and any ICD-10 disorder, was 0.30 (95% CI 0.06–0.54). For the comparison between the heavy drinker caseness and the AUDIT using
the same 7/8 cutpoint, the kappa was similar, 0.31 (95% CI 0.09–0.53).

For factor analysis, the variable referring to the need for a drink in the morning had to be excluded because its variance was zero. The adequacy of the data for factor analysis was assessed by the inspecting the correlation matrix, which showed that 22 coefficients in 35 were >0.3. Moreover, the Kaiser–Mayer–Oklin value was 0.77 and the Bartlett’s test of sphericity achieved high level of statistical significance ($P<0.0001$), supporting the factorability of the correlation matrix.

Principal components analysis revealed the presence of at least two components with eigenvalues exceeding 1, the first explaining 41.8% and the second 14.1% of the variance. However, given a third component with an eigenvalue of 0.99, explaining 11.0% of the variance, the possibility of three components was also analysed. The inspection of the screeplot also suggested this possibility.

In the one-factor solution, the highest load was for feeling guilty (0.80) and almost all other factors, except for ‘injuring others’ (0.21), achieved higher values than 0.4. Using the varimax rotation with a two-factor solution, the quantity and frequency indexes and the items addressing binge drinking, ‘blackouts’, and ‘injuring others’ loaded strongly on the first factor. The variables ‘feeling guilty’, ‘failing to do expected things’, and ‘unable to stop’ loaded on the second factor. The question about ‘others concern about drinking’ loaded on both factors, although the coefficient was a little higher on the first factor. In the three-factor solution, only the question about ‘injuring others’ loaded on the third factor. The two-factor solution was considered more interpretable and coherent. The first factor was interpreted as relating to level of consumption and its predictable consequences. The second factor was considered to be related to acknowledgement of personal responsibility for the adverse consequences of drinking. When this model was compared with other models (Table 2) existing in the literature using confirmatory factor analysis, the model with the best fit was that proposed by Medina-Mora, with the lowest AIC and with a high correlation between both factors (Table 3). In the model proposed by WHO, the two factors ‘dependence’ and ‘harmful drinking’ correlated perfectly between themselves and had equivalent correlations with the third factor of ‘consumption and frequency’, demonstrating that they were measuring the same construct.

The Cronbach $\alpha$ coefficient for the whole scale was 0.81 and removing any of the items did not improve upon the internal consistency. Comparing the alternative factorial structures, in the WHO model both the dependence (Cronbach’s $\alpha$ 0.41) and harmful use (0.65) subscales showed poor internal consistency. The subscales suggested by the factor structures for the Medina-Mora model: consumption level (0.77) and problems/consequences (0.73); and for the PCA model: first factor (0.78); second factor (0.70); were more internally consistent.

### DISCUSSION

The main result of this study is that the optimal cut-off point of the AUDIT in an urban, middle class sample was the same (7/8) as that suggested in Brazilian primary care settings (Mendez, 1999). Our sample had a higher SES than the general population in Salvador, where the mean income of formally working people is 3.9 legal minimum salaries (IBGE, 2001). The majority of cases of alcohol-related disorder in the general population tend to be mild compared with cases in clinical settings (Caetano and Curandi, 2002). In our sample, this fact is exemplified by the absence of individuals drinking in the morning. Cases would therefore tend to have lower screening scores, but this does not seem to affect the cut-off point in this sample. The poor response rate in our study should not have affected the generalizability of the findings unduly, in so far as the achieved sample provided a broad range of drinking behaviours and a reasonable prevalence of alcohol-related disorder. These parameters of validity seem to be intrinsic characteristics of the instrument.

One of the questions raised in a validation study is about what is the real construct behind the instrument to be validated. The AUDIT was designed to screen for hazardous and harmful drinking, but most validity studies used CIDI, which does not measure hazardous drinking. Hazardous drinking places individuals at risk of developing related diseases. Some authors consider hazardous drinking as synonymous of heavy drinking (Reid et al., 1999), but predictive validity is the key element. AUDIT scores predict social and medical problems (Claussen and Aasland, 1993; Conigrave et al., 1995a), and this instrument also predicts future alcohol intake better than previous alcohol consumption or biological markers (Conigrave et al., 1995b). This discrepancy between constructs may be an explanation for the low kappa for the agreement between AUDIT ‘cases’ and the criterion of any ICD-10 diagnosis. However, the kappa for the agreement between AUDIT and heavy drinking as measured by the CIDI was even lower.

The internal structure of the AUDIT as suggested by factorial analysis has been somewhat controversial. Most authors

### Table 2. Composition of the models submitted to confirmatory factor analysis

<table>
<thead>
<tr>
<th>AUDIT questions</th>
<th>Content</th>
<th>PCA</th>
<th>Single-factor</th>
<th>WHO</th>
<th>Medina-Mora</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Frequency of drinking</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
</tr>
<tr>
<td>Q2</td>
<td>Quantity in a typical day</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
</tr>
<tr>
<td>Q3</td>
<td>Binge drinking</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
</tr>
<tr>
<td>Q4</td>
<td>Unable to stop</td>
<td>F2</td>
<td>F2</td>
<td>F2</td>
<td>F2</td>
</tr>
<tr>
<td>Q5</td>
<td>Failing to do expected things</td>
<td>F2</td>
<td>F2</td>
<td>F2</td>
<td>F2</td>
</tr>
<tr>
<td>Q7</td>
<td>Feeling of guilty</td>
<td>F2</td>
<td>F2</td>
<td>F2</td>
<td>F2</td>
</tr>
<tr>
<td>Q8</td>
<td>Blackouts</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
</tr>
<tr>
<td>Q9</td>
<td>Injuring others</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
</tr>
<tr>
<td>Q10</td>
<td>Others concerned about drinking</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
<td>F1</td>
</tr>
</tbody>
</table>

### Table 3. Comparison of the models submitted to confirmatory factor analysis

<table>
<thead>
<tr>
<th>Models of the AUDIT</th>
<th>AIC</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>14.90</td>
<td>0.54</td>
</tr>
<tr>
<td>Single-factor</td>
<td>54.46</td>
<td>1.00</td>
</tr>
<tr>
<td>WHO</td>
<td>18.38</td>
<td>0.61</td>
</tr>
<tr>
<td>Medina-Mora</td>
<td>14.47</td>
<td>0.63</td>
</tr>
</tbody>
</table>
have reported that the AUDIT is measuring just one construct in samples with a high prevalence of alcohol dependence (Skipsy et al., 1997), or just two constructs in low prevalence samples, the first one being the level of consumption and the second one the related problems or consequences, (Medina-Mora et al., 1998; Maisto et al., 2000; Chung et al., 2002). Both proposals depart from the three-dimension model proposed by the WHO, inspired by considerations of face validity.

The results of our confirmatory factor analysis were similar to those from other samples with a low prevalence of dependence, and different from a sample with a higher prevalence of this disorder, where the AUDIT seemed to have one single-factor structure. The Cronbach α coefficient shows a high internal consistency for the whole AUDIT questionnaire. This does not contradict the results of our factor analysis; high internal consistency can occur when different factors of the same scale are highly correlated (Spector, 1992). The low Cronbach α of the subscales of dependence and harmful use of alcohol is consistent with the poor goodness-of-fit of the factorial structure proposed by the WHO model. The perfect correlation between them also suggests they are measuring the same construct. On the other hand, the higher Cronbach α of the subscales proposed by the two-factor solution found by Medina-Mora is in agreement with the better goodness-of-fit of that model. The AIC indicates an almost similar goodness-of-fit of the PCA model, compared with that of Medina-Mora, and at least two of the items included in the first PCA’s factor, the occurrence of blackouts and of injuries due to drinking (questions 8 and 9) can be related to acute effects of the consumption. Nevertheless, considering not only the slightly smaller magnitude of AIC, but also the highest correlation between factors, the comprehensibility and parsimony of the factors’ distribution, the model proposed by Medina-Mora proved to be the better in the confirmatory factor analysis. The two-factor solution proposed by Medina-Mora and endorsed in this paper does not detract from the notion of the AUDIT as a unidimensional scale suitable for the purpose of screening for hazardous and harmful drinking. Indeed our data provides empirical support for this application for the AUDIT.

In conclusion, this validation study supports the use of the AUDIT in population-based epidemiological research in Brazil, as it has been used in clinical settings and in the same format that was proposed in the original study. The instrument has demonstrated the same properties it showed in other samples with a low prevalence of alcohol dependence, including the same factorial structure. The inconsistency of this factor structure with the theoretically based model proposed by the WHO suggests the need for caution in interpretation of AUDIT ‘subscale’ scores.

Acknowledgements — Dr Carlos Tadeu Lima has been awarded with a scholarship from CAPES Foundation, Ministry of Education, Brazil.

REFERENCES


