AUDIT-3 AND AUDIT-4: EFFECTIVENESS OF TWO SHORT FORMS OF THE ALCOHOL USE DISORDERS IDENTIFICATION TEST

ANTONI GUAL1,2*, LIDIA SEGURA1, MONTSERRAT CONTEL1, NICK HEATHER3 and JOAN COLOM1

1Health and Social Security Department, Program on Substance Abuse, Autonomous Government of Catalonia, 2Alcohol Unit, Hospital Clinic, Barcelona, Spain and 3Centre for Alcohol and Drug Studies, University of Northumbria at Newcastle upon Tyne, UK

(Received 19 September 2001; in revised form 1 February 2002; accepted 27 May 2002)

Abstract — Aims: To identify suitable short versions of the Alcohol Use Disorders Identification Test (AUDIT) and to evaluate their effectiveness as screening tests for 'risky drinking' among men and women in primary health care (PHC) settings. Methods: A total of 255 patients attending five PHC centres in Catalonia (Spain) were interviewed by clinicians regarding health status and drinking pattern. Patients also completed the AUDIT. Clinicians' diagnosis of risky drinking was used as a gold standard to evaluate the effectiveness of three forms of AUDIT. Results: AUDIT-3 and AUDIT-4 performed similarly to AUDIT-10 in detecting risky drinking and had equivalent receiver operating characteristics curves and their areas under the curve. Conclusions: Both short forms of AUDIT seem to be as effective as the full AUDIT for detecting risky drinking among men and women in PHC settings.

INTRODUCTION

The social and economic burden due to alcohol consumption in Western countries has become an increasing cause for concern. Most alcohol-related problems appear in non-alcohol-dependent individuals who fall into the categories of hazardous or harmful drinkers according to World Health Organization terminology.

Concern about those drinking over recommended levels has led to the concept of 'risky drinking' (Higgins-Biddle and Babor, 1996), usually referring to men drinking >280 g (168 g for women) of alcohol per week. Although not defined precisely, the concept of risky drinking is often used to include both hazardous and harmful drinking.

Several studies have shown that drinking above World Health Organization recommendations (Dawson and Archer, 1993; Anderson, 1996; Prada et al., 1996; Portella et al., 1998) raises the risk of alcohol-related physical and psychosocial problems (100% for liver cirrhosis, 20–30% for cancer of the oral cavity, pharynx and larynx, 10% for cancer of the oesophagus, 14% for cancer of the liver, 10–20% for cancer of the female breast and possibly 20% for stroke). Overall, the World Health Organization estimates that, in developed countries, alcohol accounts for 10–11% of all illnesses and deaths each year (Murray and Lopez, 1996). Risky drinking and alcohol misuse or dependence are common in primary health care (PHC) patients. In European countries, the prevalence of risky drinking in PHC settings ranges from 2.1 to 41% among men and from 0.8 to 21% among women (World Health Organization, 2001). This high prevalence of risky drinking has led to the development of screening tools and brief intervention packages (Heather et al., 1987; Gomel et al., 1994; University of Sydney, 1994; National Institute on Alcohol Abuse and Alcoholism, 1995; Òrgan Tècnic de Drogodependencies, 1996), which aim to help general practitioners (GPs) and other health care professionals to identify, assess and advise risky drinkers. The fact that brief interventions have proven efficacy (Wallace et al., 1988; Babor and Grant, 1992; Nuffield Institute for Health, 1993; Altisent et al., 1997; Fernández et al., 1997) and cost-effectiveness (Fleming et al., 2000) clearly increases the need to develop reliable, valid and user-friendly screening tools.

Fiellin et al. (2000) concluded in their systematic review that the literature supports screening for lifetime and current misuse or dependence disorders by means of the CAGE questions and, for less severe alcohol problems, such as at-risk, harmful, and hazardous drinking, by means of the Alcohol Use Disorders Identification Test (AUDIT).

The AUDIT (Saunders and Aasland, 1987; Saunders et al., 1993) was developed as part of the World Health Organization Collaborative Project on the Detection and Management of Alcohol-related Problems in Primary Health Care, to identify hazardous and harmful alcohol use. AUDIT is a 10-item questionnaire that enquires about alcohol consumption and frequency of drinking, the presence of alcohol-related problems and alcohol-dependence symptoms. Several studies have shown its validity and reliability in the detection of risky drinking, alcohol misuse and alcohol dependence (Bohn et al., 1995; Martínez, 1996; Piccinelli et al., 1997; Volk et al., 1997; Rubio et al., 1998; Contel et al., 1999). Depending on the cutoff and the criterion standards used, studies have reported sensitivities between 51 and 97% and specificities between 78 and 96% (Fiellin et al., 2000).

Seppä et al. (1998) developed the Five-Shot Questionnaire for detecting risky drinking, by combining two items from AUDIT asking about drinking amounts and three items from CAGE that correspond to the three different types of question in the AUDIT ([304,86]hazardous alcohol consumption, dependence symptoms and harmful alcohol consumption). This instrument was tested in a middle-aged male population and, although it performed better than the CAGE, its usefulness among other age groups, among women and in PHC settings has not been demonstrated.

The AUDIT-C (Bush et al., 1998; Aertgeerts et al., 2001; Gordon et al., 2001) includes only the three AUDIT alcohol consumption questions and its performance as a screening test...
has been evaluated in three different studies. Bush et al. (1998) evaluated the AUDIT-C for active alcohol misuse or dependence and/or risky drinking in a male population. Although the AUDIT-C performed better than the full AUDIT and the CAGE in identifying risky drinkers, this study was restricted to men, performed at three Veterans Affairs (VA) general medical clinics and the interviews were conducted by telephone. Recent data (Kraus and Augustin, 2001) suggest that telephone interviews can produce a significant bias in results. Gordon et al. (2001) used the AUDIT-C to identify hazardous drinkers in a large PHC sample. The AUDIT-C proved to be as effective as the AUDIT, even though criteria for hazardous drinking were not established on the basis of clinical judgement, but using quantity–frequency measures obtained from a self-administered questionnaire. In general, the AUDIT-C has shown a sensitivity of 54 to 98% and a specificity of 57 to 93% for various definitions of heavy drinking (Fiellin et al., 2000).

In Europe, a large study of alcohol screening questionnaires in PHC carried out in Belgium (Aertgeerts et al., 2001) compared the full AUDIT with two shorter forms (Bush et al., 1998; Gordon et al., 2001) and the 5-shot questionnaire (Seppä et al., 1998). The AUDIT-C performed significantly less well than the full AUDIT among female patients, but compared well with other questionnaires. It should also be noted that this study focused on alcohol misuse and dependence, not on ‘risky drinking’. Nevertheless, the authors suggested that the simplicity of the AUDIT-C supports its routine use by GPs.

Throughout Phase III of the World Health Organization Collaborative Study (Implementing and Supporting Early Intervention Strategies in Primary Health Care) the opinions of all GPs who participated were collected by means of systematic focus groups. Most GPs complained about the fact that it was too long to be used as a systematic screening tool. They also stated that the alcohol consumption questions (1–3) were well accepted by patients, but questions dealing with alcohol dependence and alcohol-related problems (4–10) tended to arouse defensiveness. These observations, together with the encouraging results previously obtained with short forms of the AUDIT (Piccinelli et al., 1997; Bush et al., 1998; Aertgeerts et al., 2001; Gordon et al., 2001) led us to design the present study, in order to identify suitable short versions of AUDIT and, if appropriate, to test their effectiveness as screening tests for risky drinking among men and women in PHC settings. Also, this study provided an opportunity to examine possible differences between a population with a Mediterranean drinking pattern and other populations previously studied.

**SUBJECTS AND METHODS**

**Subjects**

This study was conducted in five PHC centres in Catalonia, Spain. Data presented here are from a validation study carried out as part of a larger study concerned with screening and brief intervention among risky drinkers in PHC settings. According to the Catalan Health Plan Objectives, patients must be asked about their drinking habits every 2 years as part of the clinical routine; no informed consent was then required. A total of 269 patients of both genders were interviewed, of whom 14 (5.2%) were excluded because of lack of relevant information. No *a priori* eligibility criteria were used to select the population, and patients were interviewed when attending their PHC setting.

**Alcohol screening measures**

Before the interview, patients were asked to complete the 10-item AUDIT. GPs interviewed patients about their drinking habits using a Systematic Interview of Alcohol Consumption that included three questions exploring frequency and amount of consumption: ‘If you ever drink alcoholic beverages (wine, beer, etc.), how many beverages a day?’ (measured in standard drinks); ‘How often?’ (number of days in a week); and ‘On weekends (or workdays) do your drinking habits change?’. Detailed information about normal and exceptional drinking patterns was obtained with this recently validated instrument (Gual et al., 2001). The clinician’s diagnosis of risky drinking was used as a gold standard. GPs identified as risky drinkers all patients whose weekly alcohol consumption was above the World Health Organization recommendations (280 g for men and 168 g for women) and/or who fulfilled criteria for hazardous or harmful drinking (World Health Organization, 2000).

**Statistical analysis**

The Statistical Package for the Social Sciences (SPSS 9.0) was used for data analysis. Logistic regression analysis was carried out to identify those items that minimized the probability of misclassification between subjects with and without risky drinking. Following the Piccinelli et al. (1997) study, a stepwise selection of items was adopted by using the likelihood ratio statistic as a test for removal and a *P*-value of 0.10 to remove an item. Cross-tabulations and *t*-tests were used for group comparisons. Relationships between the three forms of AUDIT and reported mean weekly alcohol consumption were examined by regression analysis. Sensitivity, specificity, positive predictive and overall accuracy (OA) values were calculated for the full, 3-item and 4-item AUDIT forms in relation to the gold standard (diagnosis of risky drinking). Receiver operating characteristics (ROC) curves and their areas under the curve (AUROCs) were inspected to choose the optimal screening test and the best cut-off scores according to sensitivity and specificity levels. To compare AUROCs, we used the *Z* statistic defined following the method described by Hanley and McNeil (1983) which takes into account the correlation between the areas that is induced by the paired nature of the data.

**RESULTS**

Of the 255 patients included, 127 (49.8%) were men with a mean age (± SD) of 43.6 ± 13.1 years (range: 17–82) and 128 (50.2%) were women with a mean age of 44.4 ± 14.4 years (range: 18–81). Other variables registered were employment status (73% employed, 12% housewives, 6% retired, 2% students and 2% unemployed), Hollingshead Index of Social Position (Hollingshead, 1957) (24% level I, 20% level II, 18% level III and 14% level IV) and educational level (54% had <9 years of school, 23% were high school graduates and 16% were college graduates).

Reported weekly alcohol consumption was translated into Spanish standard drinks (Gual et al., 1999) (one drink = 10 g of pure ethanol). Alcohol consumption in men ranged from 0 to 105 standard drinks and for women from 0 to 32 (Table 1).
AUDIT-10 scores ranged in men from 0 to 34 (mean ± SD: 6.76 ± 5.64) and in women from 0 to 10 (2.75 ± 1.91).

Comparison of AUDIT-10, AUDIT-3 and AUDIT-4

The logistic regression analysis retained five items in the model ($\chi^2 = 180.18; \text{df } = 5; P < 0.0001$) with 80% sensitivity, 96% specificity and 92% overall accuracy (Table 2). The items analysed were the following: item 1 (How often do you have a drink containing alcohol?); item 2 (How many drinks containing alcohol do you have on a typical day when you are drinking?); item 3 (How often do you have 6 or more drinks on one occasion?); item 6 (How often during the past year have you needed a first drink in the morning to get yourself going after a heavy drinking session?) and item 10 (Has a relative or friend or a doctor or other health worker been concerned about your drinking or suggested that you should cut down?). The same analysis was done separately for both genders. Among males, the logistic regression analysis retained the same five items (sensitivity: 89%; specificity: 90%), whereas among females, four items (item 6 dropped out) were retained in the model (sensitivity: 64%; specificity: 100%) (Table 1). These results led us to evaluate two different short forms of the AUDIT by combining those items significantly associated with the clinical diagnosis of risky drinking and which were retained in the model for both genders. Scores for these short forms were extracted from the full AUDIT for each participant. The AUDIT-3 includes the first three questions which assess alcohol consumption (reported as AUDIT-C in the literature), and AUDIT-4 is constructed by adding the 10th item to these three questions.

AUDIT-3 scores ranged from 0 to 12 (mean ± SD 3.58 ± 2.44) and AUDIT-4 from 0 to 16 (4.24 ± 3.31). Scores on AUDIT-10 ($r = 0.83, n = 255, P < 0.001$), AUDIT-3 ($r = 0.84, n = 255, P < 0.001$) and AUDIT-4 ($r = 0.830, n = 255, P < 0.001$) were strongly correlated with alcohol consumption per week.

According to clinician diagnosis, 41.7% of men and 8.6% of women were classified as risky drinkers. According to the AUDIT cut-off scores established in its Spanish validation (≥9 for men and ≥6 for women), 31.5% of men and 6.3% of women were classified as risky drinkers.

ROC curves comparing the three AUDIT forms with the gold standard are presented with their AUROCs and 95% confidence intervals for men and women in Figs 1 and 2. Sensitivities and specificities at different cut-off scores are detailed for both genders in Table 3. Among men, AUROCs (Fig. 1) were 0.913 for AUDIT-3, 0.924 for AUDIT-4 and 0.920 for AUDIT-10 and no significant differences were found in comparisons among them (AUDIT-3 vs AUDIT-10, $P = 0.646$; AUDIT-4 vs AUDIT-10, $P = 0.751$). When comparing all AUROCs (0.957, 0.945 and 0.871 respectively) among women (Fig. 2), no significant differences were found (AUDIT-3 vs AUDIT-10, $P = 0.889$; AUDIT-4 vs AUDIT-10, $P = 0.90$).

Among men, the best cut-off scores were 7 for the full AUDIT (sensitivity 86.8%; specificity 81.1%) and AUDIT-4 (sensitivity 83.0%; specificity 89.1%), and 5 for AUDIT-3 (sensitivity 92.4%; specificity 74.3%). Sensitivity levels at each cut-off were higher for the full AUDIT, but specificity levels were higher for AUDIT-4. Among women, the best cut-off scores were 5 for the full AUDIT and AUDIT-4 with equal sensitivity and specificity levels (sensitivity 72.7%; specificity 95.73%), and 4 for AUDIT-3 (90.9 and 68.4% respectively). Sensitivity and specificity levels at each cut-off were higher for the full AUDIT and AUDIT-4 than for AUDIT-3.

Risky drinker prevalences according to the best cut-off scores for AUDIT-3, -4 and -10 are also presented in Table 4.

### DISCUSSION

In the Spanish context, the AUDIT questionnaire has been widely used by physicians in conjunction with the World Health Organization Phase III Collaborative Study on alcohol and PHC. As noted above, these doctors complained about its length and an increase in patients’ defensiveness when...
Fig. 1. Receiver operating characteristics curves of the three forms of AUDIT among males. No differences in areas under the curves were found between: AUDIT-3 vs AUDIT-4 ($P = 0.412$); AUDIT-3 vs AUDIT-10 ($P = 0.646$); and AUDIT-4 vs AUDIT-10 ($P = 0.751$).

Fig. 2. Receiver operating characteristics curves of the three forms of AUDIT among females. No differences among areas under the curves were found between: AUDIT-3 vs AUDIT-4 ($P = 0.887$); AUDIT-3 vs AUDIT-10 ($P = 0.889$); and AUDIT-4 vs AUDIT-10 ($P = 0.903$).

Table 3. Performance of three AUDIT forms compared with clinical diagnosis of risky drinking

<table>
<thead>
<tr>
<th>Gender</th>
<th>Version</th>
<th>Cut-off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive predictive value</th>
<th>Overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>AUDIT-3</td>
<td>≥4</td>
<td>1.00</td>
<td>0.53</td>
<td>0.60</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥5</td>
<td>0.92</td>
<td>0.74</td>
<td>0.72</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥6</td>
<td>0.60</td>
<td>0.93</td>
<td>0.86</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>AUDIT-4</td>
<td>≥6</td>
<td>0.87</td>
<td>0.78</td>
<td>0.74</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥7</td>
<td>0.83</td>
<td>0.89</td>
<td>0.85</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥8</td>
<td>0.68</td>
<td>0.94</td>
<td>0.90</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>AUDIT-10</td>
<td>≥6</td>
<td>0.91</td>
<td>0.73</td>
<td>0.71</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥7</td>
<td>0.87</td>
<td>0.81</td>
<td>0.77</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥8</td>
<td>0.73</td>
<td>0.92</td>
<td>0.87</td>
<td>0.84</td>
</tr>
<tr>
<td>Females</td>
<td>AUDIT-3</td>
<td>≥3</td>
<td>0.91</td>
<td>0.52</td>
<td>0.15</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥4</td>
<td>0.91</td>
<td>0.68</td>
<td>0.21</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥5</td>
<td>0.54</td>
<td>0.96</td>
<td>0.54</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>AUDIT-4</td>
<td>≥4</td>
<td>1.00</td>
<td>0.68</td>
<td>0.23</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥5</td>
<td>0.73</td>
<td>0.96</td>
<td>0.61</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥6</td>
<td>0.54</td>
<td>1.00</td>
<td>1.00</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>AUDIT-10</td>
<td>≥4</td>
<td>1.00</td>
<td>0.68</td>
<td>0.23</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥5</td>
<td>0.73</td>
<td>0.96</td>
<td>0.61</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥6</td>
<td>0.54</td>
<td>0.99</td>
<td>0.86</td>
<td>0.95</td>
</tr>
</tbody>
</table>
EFFECTIVENESS OF SHORT AUDIT FORMS

Table 4. Prevalences of risky and non-risky drinkers according to the best cut-off scores in all AUDIT forms and among gender

<table>
<thead>
<tr>
<th>AUDIT form</th>
<th>Males (n = 127)</th>
<th>Females (n = 128)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Risky drinker n (%)</td>
<td>Non-risky drinker n (%)</td>
</tr>
<tr>
<td>AUDIT-3</td>
<td>53.5</td>
<td>46.5</td>
</tr>
<tr>
<td>AUDIT-4</td>
<td>40.9</td>
<td>59.1</td>
</tr>
<tr>
<td>AUDIT-10*</td>
<td>47.2</td>
<td>25.8</td>
</tr>
</tbody>
</table>

* t = 7.62; df = 253; P < 0.001.

For the shorter forms of AUDIT, the cut-off scores were similarly reduced.

The results for male patients found with the AUDIT-3 (or AUDIT-C), including the cut-off point (≥5), fit well with previous studies. For female patients, AUDIT-4 performed exactly as the full AUDIT at the same cut-off points. The results reported here have led us to recommend that clinicians use the AUDIT-3 questionnaire in general health screening interventions. If a score of ≥5 among men and ≥4 among women is observed, a more in-depth assessment of drinking pattern and alcohol-related problems should be carried out. If including question 10 increases the total score to 7 among men and 5 among women, this should definitely lead to a diagnosis of risky drinking. More detailed questioning, including the full AUDIT if necessary, could be used to make decisions about the need for referral to specialist help for alcohol dependence. Screening devices, such as the AUDIT in all of its self-administered forms, facilitate the GP’s daily work, and, as our results confirm, are as reliable as a systematic GPs interview on alcohol consumption.

Acknowledgements — We thank the following primary care physicians for collaborating: N. Bastida, I. Pie, J. M. Segura, E. Esquerra, P. Edo, M. Pallarés and R. Catalá.

REFERENCES


National Institute on Alcohol Abuse and Alcoholism (1995) Screening and Brief Intervention Package. NIAAA, Rockville, MD.


National Institute on Alcohol Abuse and Alcoholism (1995) Screening and Brief Intervention Package. NIAAA, Rockville, MD.