CAGE Questionnaire

Description of Measure

Purpose
To quickly screen for problem drinking.

Conceptual Organization
The questions were developed to be brief, easy to administer, and easy to remember with the idea that physicians could use them as a brief screening tool in a clinical setting. The CAGE is comprised of four non-threatening questions about drinking, which loosely correspond to the acronym, “CAGE.”

Item Origin/Selection Process
Item origin and selection process is described by Ewing (1984). The CAGE questions were developed in 1968 in a clinical study. Study participants were 130 randomly selected medical and surgical patients selected from a general hospital population. Each patient agreed to a lengthy interview including questions developed for the detection of alcoholism. As reported by Ewing, sixteen of the 130 patients were clearly identified as suffering from alcoholism by their own physicians (confirmed by chart review). Analyses contrasted the 16 identified alcoholics with the remainder of the sample, and four questions emerged as most useful in differentiating the two groups. The four questions were retested using a clinical population and were found to reliably distinguish alcoholics from non-alcoholics.

Materials
**Time Required**

2 minutes

**Administration Method**

Interviewer-administered

**Training**

Minimal

**Scoring**

**Score Types**

Yes responses are coded 1; No responses are coded 0. Responses to the four items are summed to derive a total score with values ranging from 0-4.

**Score Interpretation**

A higher score indicates greater risk for alcoholism. Mayfield and others used two or more positive responses as indicative of alcoholism or problem drinking (King, 1986; Mayfield, McLeod, & Hall, 1974).

A critical review of alcohol screening questionnaires in female patients (Bradley et al., 1998) recommended that the cutoff point for the CAGE be lowered from two to one or more positive responses. By lowering the cutoff point from two to one, sensitivity of the measure (the number of true positives) increased by as much as 20%, with only a 10% (at most) reduction in specificity (the number of true negatives).

**Psychometric Support**

**Reliability**
In a general population survey of 703 drinkers over the age of 18, a factor analysis of the CAGE indicated that the four items measure a single dimension ($\chi^2 = 1.22, 2 \text{ df, } p = 0.54$), presumably problem drinking, and exhibit good internal reliability, factor loadings ranging from .55 to .92 (Smart et al., 1991).

**Validity**

The CAGE has been validated in both the US (Beresford et al., 1982) and the UK (Barrison et al., 1982; Bernadt et al., 1982; King, 1986), and research using clinical samples has indicated that the CAGE identifies most alcoholics. Early studies found that the instrument functioned most effectively at a cut-off point of two or more affirmative replies with a sensitivity of 84%, a specificity of 95%, and a positive predictive value of 45% (King, 1986; Mayfield, McLeod, & Hall, 1974). A more recent critical review of the literature has suggested that lower thresholds be used for a positive screening result for women (Bradley et al., 1998, see Score Interpretation above).

Although the test is reported to reliably distinguish groups diagnosed as alcoholic from other clinical samples, it may produce a high number of false positives, especially individuals who formerly had drinking problems (Strang, Bradley, & Stockwell, 1989). In a comparison between the CAGE and the Clyde Bank Questionnaire (CBQ) (a disguised alcohol questionnaire), the CAGE correlated well with the CBQ for subjects who were chemically dependent on alcohol ($r = .77$), but poorly for problem drinkers with no symptoms of physical dependence ($r = .25$). Very few studies have examined the use of the CAGE in a general population survey and compared results with other aggregate estimates of alcohol problems (Smart et al., 1991). Saunders and Kershaw (1978, 1980) used the CAGE in a community sample and reported poor validity, with hospital records as the source of validation. However, this study was limited in that not all alcoholics or problem drinkers have clinical records confirming their behavior. It has been noted that there is probably no absolute method of validation of an alcoholism screening test, short of direct observation (King, 1986).
LONGSCAN Use

Data Points

Pre-Age 4: NW & MW sites only
Age 4: all sites

Respondent

Primary maternal caregiver

Mnemonic and Version

Administered as part of the Caregiver Physical Health Assessment (MHLA), as items 5-8.

Rationale

LONGSCAN investigators are interested in exploring the relationship between a history of alcoholism or problem drinking and child abuse and neglect. The CAGE was chosen because it is a widely used measure that is brief and easy to administer.

Administration and Scoring Notes

A screening item asks if the respondent has ever consumed alcoholic beverages. Every respondent who answers “yes,” to this question is then administered the CAGE questionnaire.

Results

Descriptive Statistics

Table 1 displays a breakdown of CAGE scores at the Age 4 interview by race and study site. One-third of the caregivers reported never having consumed alcoholic beverages at all. Of the participants who answered the screening question affirmatively, 40% scored 0, indicating no
problems at all with the targeted drinking behaviors. Across all sites, 17 to 27% of respondents, depending on whether a cutoff point of 1 or 2 is used, had scores indicative of problem drinking. The SO site had the highest rate of abstinence (51%) among the five sites, followed by the EA site (41%). The NW site had the lowest rate of abstinence among the sites with almost one-quarter of the NW caregivers scoring 2 or higher. Black and Hispanic caregivers reported higher rates of abstinence from drinking than other racial groups.

Table 1 about here

Reliability

Internal consistency reliability was examined by computing Cronbach’s alpha coefficients. (See Table 1.) Alphas were acceptable across sites and racial groups, especially for a 4-item scale (entire sample $\alpha = .78$). One would not expect extremely high internal consistency on such a test because higher endorsement of symptoms likely reflects greater severity of alcohol abuse, that is, the presence of one symptom should not necessarily predict the presence of another.

References and Bibliography


Table 1. Frequency of CAGE Scores and Cronbach’s Alpha Coefficients by Race and Study Site

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>% No drinking ever</th>
<th>CAGE Score</th>
<th></th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>&gt;2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1147</td>
<td>33.3</td>
<td>39.9</td>
<td>10.2</td>
<td>16.5</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>397</td>
<td>25.2</td>
<td>46.6</td>
<td>9.8</td>
<td>18.4</td>
</tr>
<tr>
<td>Black</td>
<td>584</td>
<td>37.7</td>
<td>36.1</td>
<td>10.4</td>
<td>15.7</td>
</tr>
<tr>
<td>Hispanic</td>
<td>82</td>
<td>41.5</td>
<td>41.4</td>
<td>4.9</td>
<td>12.2</td>
</tr>
<tr>
<td>Multiracial</td>
<td>37</td>
<td>24.3</td>
<td>43.3</td>
<td>16.2</td>
<td>16.2</td>
</tr>
<tr>
<td>Other</td>
<td>44</td>
<td>18.2</td>
<td>47.7</td>
<td>15.9</td>
<td>18.1</td>
</tr>
<tr>
<td><strong>Study Site</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA</td>
<td>237</td>
<td>41.3</td>
<td>37.6</td>
<td>9.7</td>
<td>11.3</td>
</tr>
<tr>
<td>MW</td>
<td>123</td>
<td>32.5</td>
<td>40.7</td>
<td>9.8</td>
<td>17.2</td>
</tr>
<tr>
<td>SO</td>
<td>221</td>
<td>50.7</td>
<td>27.6</td>
<td>7.2</td>
<td>14.4</td>
</tr>
<tr>
<td>SW</td>
<td>316</td>
<td>30.1</td>
<td>42.7</td>
<td>11.7</td>
<td>15.5</td>
</tr>
<tr>
<td>NW</td>
<td>250</td>
<td>12.4</td>
<td>52.0</td>
<td>11.6</td>
<td>24.0</td>
</tr>
</tbody>
</table>

*Source.* Based on data received at the LONGSCAN Coordinating Center by 7/8/97.