Using the Neurobehavioral Cognitive Status Examination as a Screening Measure for Older Adults

Daniel L. Drane
Medical College of Georgia, Augusta

Sheryl S. Osato
West Los Angeles VA Medical Center and UCLA School of Medicine

We evaluated the ability of the Neurobehavioral Cognitive Status Examination (NCSE) to accurately distinguish between healthy older adults and geriatric patients suffering from dementia. Although the NCSE correctly identified all dementia patients, it produced an unacceptably high rate of false positives among the healthy elderly (70%). Despite the NCSE's lack of specificity when using the recommended classification criteria, significant group differences were found on several individual subscales and on the total number of subscales passed. These findings suggest the need to further evaluate the appropriateness of the geriatric norms for the NCSE and highlight some of the unique considerations involved in the assessment of older adults. © 1997 National Academy of Neuropsychology

Clinical neuropsychologists are poised to play a pivotal role in the care and treatment of the ever-increasing geriatric segment of the population through the process of ongoing cognitive assessment and functional evaluation. More than ever before, geriatric care facilities are regularly faced with the need to obtain neuropsychological evaluation of their residents for purposes that include differential diagnosis, placement, and an assessment of needs regarding nursing requirements and level of general care. The demand for sensitive cognitive screening measures that can be employed in an expedient, valid fashion is reinforced by the realization that the elderly are particularly susceptible to potentially reversible causes of cognitive impairment related to such etiologies as depression and a sensitivity to the effects of medication.

Several widely used screening instruments for use with older adults suffer from a lack of sensitivity and an inability to provide more than a dichotomous classification of patients as either cognitively impaired or non-impaired (Osato, Yang, & La Rue, 1993). The Neurobe-
Behavioral Cognitive Status Examination (NCSE) (Northern California Neurobehavioral Group, 1988) has attempted to address many difficulties that have plagued these traditional bedside screening measures and is reportedly sensitive to cognitive impairment (Schwamm, Van Dyke, Kiernan, Merrin, & Mueller, 1987). The NCSE increases the chance of identifying focal deficits by providing results for several individual domains rather than a single global score. The NCSE is also set apart from previous screening measures by its greater sampling of construction and language functions. It should, therefore, be more sensitive to right hemisphere dysfunction and language impairment.

Despite the apparent usefulness of the NCSE, its utility within a geriatric population requires further validation. Various age cohorts often respond differently to the stimulus materials of any given psychometric instrument due to differences in education and life experiences (Kaszniak, 1990). Likewise, due to the general physical impairment and slowing which accompanies normal aging, older adults often perform significantly worse than other segments of the population on tasks requiring psychomotor speed, motor coordination, and visuo-perceptual skills (Zarit, Eiler, & Hassinger, 1985). The present study represents a preliminary examination of the validity of the NCSE among older adults by examining its ability to accurately distinguish between healthy elderly residents and patients with dementia in a VA nursing home.

**METHOD**

**Subjects**

Nursing home patients received a primary diagnosis of dementia as defined by *DSM-III-R* criteria (American Psychiatric Association, 1987) after completing a standardized diagnostic evaluation. This included a complete medical and psychiatric history, physical and neurological examinations, and laboratory tests including SMA-12, complete blood count, electrolytes, VDRL, T4, chest x-ray, and electrocardiogram. Patients were excluded if there was evidence of a co-existing psychiatric disturbance. They ranged in age from 60 to 100 (mean = 76.9, SD = 9.8) and had obtained a mean of 12.3 years of education (SD = 3.9).

The healthy control group consisted of residents of a residential retirement center. Subjects with histories of either psychiatric illness or serious medical condition were excluded on the basis of a careful review of available medical records. In addition, each subject received a structured interview and those with evidence of a psychiatric disturbance or failing to score at least 24 points on the Mini-Mental State Examination were excluded. This group ranged in age from 60 to 96 (mean = 79.3, SD = 11.6) and had obtained a mean of 12.6 years of education (SD = 3.9).

**Procedure**

Individual subtest performance was classified as either impaired or non-impaired on the basis of age-corrected cut-off scores provided in the NCSE test manual (Northern California Neurobehavioral Group, 1988). Categorization of subjects as impaired or non-impaired was compared with the original diagnostic group classifications (i.e., dementia patients vs. normal elderly).

**RESULTS**

The equivalence of the two diagnostic groups was examined along a variety of descriptive parameters (e.g., age, education, and gender). With the exception of gender ($\chi^2 = 12.29, p <$...
Cognitive Screening Among Older Adults

**TABLE 1**

Comparison of Healthy Older Adults and Dementia Patients on NCSE Subscales

<table>
<thead>
<tr>
<th>NCSE Subtest</th>
<th>Healthy Below Cut-Off</th>
<th>Dementia Below Cut-Off</th>
<th>M</th>
<th>SD</th>
<th>Healthy M</th>
<th>Dementia M</th>
<th>SD</th>
<th>Significance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>10%</td>
<td>27%</td>
<td>6.8</td>
<td>1.5</td>
<td>5.8</td>
<td>1.8</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Orientation</td>
<td>0%</td>
<td>68%</td>
<td>11.6</td>
<td>1.0</td>
<td>6.5</td>
<td>3.5</td>
<td>p&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td>0%</td>
<td>50%</td>
<td>5.9</td>
<td>0.4</td>
<td>4.2</td>
<td>1.0</td>
<td>p&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Repetition</td>
<td>0%</td>
<td>77%</td>
<td>11.6</td>
<td>0.6</td>
<td>9.3</td>
<td>1.7</td>
<td>p&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Naming</td>
<td>5%</td>
<td>41%</td>
<td>7.7</td>
<td>0.9</td>
<td>6.1</td>
<td>1.9</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>32%</td>
<td>82%</td>
<td>3.8</td>
<td>1.8</td>
<td>1.0</td>
<td>1.5</td>
<td>p&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Memory</td>
<td>59%</td>
<td>82%</td>
<td>7.4</td>
<td>2.4</td>
<td>3.0</td>
<td>2.5</td>
<td>p&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Calculations</td>
<td>14%</td>
<td>64%</td>
<td>3.6</td>
<td>1.0</td>
<td>2.0</td>
<td>1.3</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Similarities</td>
<td>14%</td>
<td>50%</td>
<td>5.5</td>
<td>1.3</td>
<td>3.1</td>
<td>2.6</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>Judgment</td>
<td>5%</td>
<td>55%</td>
<td>4.8</td>
<td>0.6</td>
<td>3.5</td>
<td>1.9</td>
<td>ns</td>
<td></td>
</tr>
</tbody>
</table>

*Reflects a modified Bonferroni correction.

.001), no relevant differences were found. More men were included in the impaired group, while the normal sample included a higher proportion of women.

Accuracy ratings for the NCSE were compiled by comparing the NCSE classifications (i.e., impaired versus non-impaired) with the original diagnostic categorizations of each patient. As noted, impairment on the NCSE is indicated by below average performance in any one given domain of functioning. The NCSE correctly identified impairment in 100% (21 of 21) of the dementia group, but rated 70% (14 of 20) of the healthy group as impaired as well. Of those subjects who were classified as impaired by the NCSE, the average number of failed subscales was 6.5 ($SD = 2.4$) for the dementia patients, but only 1.5 ($SD = 1.5$) for their healthy counterparts. The percentage of subjects from each subject group falling below NCSE normative cut-off criteria on specific subscales is shown in Table 1.

Multiple t test comparisons were made in order to determine whether the healthy subjects and dementia patients differed significantly on individual NCSE subscales. A modified Bonferroni criterion was calculated in order to offset the increased chance of a Type I error engendered by multiple comparisons. Differences on 5 of the 10 NCSE subscales remained after the correction factor was added [Orientation ($t_{39} = 6.38, p < .05$), Comprehension ($t_{39} = 6.80, p < .05$), Repetition ($t_{39} = 5.78, p < .05$), Construction ($t_{39} = 5.58, p < .05$), and Memory ($t_{39} = 5.75, p < .05$)]. Group differences on these dependent measures remained strong, even after descriptive variables (including gender, education, and age) were partialled-out using regression analysis. Mean scores on individual NCSE subscales for each diagnostic group, along with the results of the aforementioned statistical comparisons, are also provided in Table 1.

**DISCUSSION**

The NCSE successfully identified cognitive dysfunction in all patients with a diagnosis of dementia (100%), yet also generated an unacceptable level of false positives among the healthy older adults (70%). While conclusions regarding the NCSE remain tenuous due to the small sample size of the current study, this lack of specificity raises concerns regarding the usefulness of the NCSE as a cognitive screening device for older adults. Problems with specificity do not appear attributable to dissimilarity in group composition between the current study and the original NCSE standardization sample since the present healthy sample appears comparable to that described in the NCSE manual. Thus, possible explanations for
the NCSE’s observed lack of specificity include the appropriateness of the geriatric norms included in its manual and some procedural and methodological assessment issues impacting the accuracy of individual subscale scores.

Despite the failure of the recommended classification criterion to accurately distinguish between healthy and impaired subjects in our study, sizable differences existed on all subscales with the exception of Attention. Even after Type I error rate correction, group differences remained on 5/7 NCSE subscales. The largest group differences were present on Memory and Construction subscales, which were primarily responsible for the high rate of false positives observed in the healthy control group. Expanding the normal range for elderly adults on these two subsections, therefore, would improve the specificity of the NCSE without compromising its current level of sensitivity.

The substantially higher number of subscales failed by the dementia patients (6.5 scales vs. only 1.5 scales, respectively) suggests that the calculation of a percentage of subscales passed may be helpful in group differentiation. In our sample, a performance ratio of 8/10 subscales passed would have reduced the false positive ratio among healthy controls from 70% to 20% while continuing to correctly identify all members of the dementia group. The addition of a global ratio as well as improved geriatric subscale norms would increase the utility of the NCSE as a screening measure for both focal and diffuse deficits among the elderly. These findings appear consistent with other work in this area (Osato et al., 1993; Van Gorp, Sultzer, Marcotte, Hinkin, Mahler, & Cummings, 1994).

One procedural factor involving the NCSE’s screen and metric approach may be hindering performance on the Memory subscale. This involves administering a series of additional tasks within a given conceptual domain whenever an initial screening item is failed. Although allowing for a more rapid evaluation, this procedure leads to variations in the duration of the overall stimulus retention interval. That is, the length of time which subjects have to retain stimulus words depends on whether they successfully pass the screening items on the intervening tasks. Subjects who successfully completed the screen had a shorter retention time and avoided the possibility of fatigue and distraction related to the performance of additional intervening tasks. Variations in both the length of stimulus retention and the extent of distractor tasks may confound individual comparisons on the Memory subscale.

Despite the questions that this study raises regarding the use of the NCSE as a screening measure for older adults, additional investigation is warranted due to sample size limitations. It appears that many of these concerns could be overcome without extensive effort. Improving geriatric norms for the NCSE is a major concern given the results of the current study and the limited size of the original older adult normative sample. Furthermore, the addition of a composite performance score and a standardized delay period for the memory subscale might also enhance diagnostic utility. Because the NCSE was sensitive to impairment, and showed promise for improved specificity, such efforts appear justified.

REFERENCES


