Do Physicians Referred for Competency Evaluations Have Underlying Cognitive Problems?
Lauri L. Korinek, PhD, Laetitia L. Thompson, PhD, Cynthia McRae, PhD, and Elizabeth Korinek, MPH

Abstract

Purpose
Research suggests that there are concerns about the neuropsychological functioning of physicians who undergo physician competency evaluation. Academic health center faculty often participate in the evaluation and remediation of these physicians. The purpose of this study was to compare the cognitive abilities between a group of physicians referred for competency evaluations and a control group.

Method
Using the MicroCog, a computerized neuropsychological screen originally designed for physicians, the authors compared the cognitive performance of 267 physicians referred for competency evaluations with a control group of 68 recruited physicians. Physicians referred for competency evaluations took the MicroCog as a part of their evaluation at CPEP, the national Center for Personalized Education for Physicians, from January 1997 to January 2004. The control group comprised practicing physicians whose competency was not in question.

Results
Compared with the control group, the competency evaluation group had a greater proportion of physicians with scores suggesting possible cognitive impairment and performed significantly lower on scores of processing speed, processing accuracy, and cognitive proficiency. The control group of physicians performed significantly better than the age- and education-corrected normative sample.

Conclusions
Because there were significant neuropsychological differences between physicians referred for competency evaluations and physicians whose competency was not in question, it is important that neuropsychological screening be included as part of physicians’ competency evaluations.


Studies suggest that about 7% to 10% of physicians practicing medicine in the United States are impaired.1,2 According to the American Medical Association Counsel, an impaired physician is “one who is unable to practice medicine with reasonable skill and safety to patients because of physical or mental illness including deterioration through the aging process, or loss of motor skill, or excessive use of drugs including alcohol.”3 In the evaluation of clinical competence of physicians, such as when medical licensure is under review, cognitive deficits that could impair physician functioning have often been overlooked or given only minimal attention. This is a concern because physicians are asked to make difficult, often quick, and sometimes life-and-death decisions that demand high and complex levels of cognitive functioning. Minor changes in cognitive functioning that might otherwise go unnoticed in most individuals may significantly affect a physician’s ability to provide competent care.4

Medical boards and hospitals frequently refer physicians whose competency is in question to specialized programs for competency assessment and remediation. These programs evaluate a physician’s clinical knowledge, reasoning, judgment, documentation, and patient care. Academic health center faculty are often asked to assist in these evaluations. Until recently, most physician competency evaluation programs did not include neuropsychological testing. However, recent research suggests that neuropsychological issues are a factor among physicians evaluated for competency.5,6 In one study,7 five physicians whose competency was seriously questioned completed an extensive three-year program to improve their clinical performance. At the conclusion of the program, only one physician’s performance improved, one physician’s performance remained the same, and three physicians’ performance declined. The authors suggested that “it is possible that their [the physicians’] incompetence arose from early age-related cognitive decline, early organic dementia, severe mood disturbance, or other conditions associated with neuropsychological impairment.”

In another study, Turnbull and colleagues5 reported that 7 of 27 physicians assessed in a physician competency evaluation program demonstrated moderate to severe cognitive problems. Therefore, a considerable proportion of physicians evaluated (26%) demonstrated a problematic level of neuropsychological functioning.

Williams and colleagues6 compared the neuropsychological performance of
Reducing malpractice liability. The MicroCog, a neuropsychological screening measure, was designed to help identify physicians with potential cognitive impairment. The Risk Management Foundation of the Harvard Medical Institutions funded the development of the MicroCog, and the instrument was purchased by clinicians to provide a screening tool for physicians with a valid, reliable, and well-normed assessment tool.

We used the MicroCog: Assessment of Cognitive Functioning to provide more information about the cognitive differences between physicians referred for competency evaluation and a control group of nonreferred physicians. We established exclusion criteria for this study.

The MicroCog: Assessment of Cognitive Functioning, or MicroCog, is a commercially available screening measure for cognitive impairment. The Risk Management Foundation of the Harvard Medical Institutions funded the development of the MicroCog, and the instrument was designed to screen elderly physicians and other professionals for subtle changes in cognitive functioning in hopes of reducing malpractice liability. The instrument was purchased by Psychological Corporation and normed on the general population for use as a general neuropsychological screen. The MicroCog is a useful neuropsychological screen that can help determine whether a person should be referred for a full neuropsychological evaluation; it is not a substitute for a complete neuropsychological battery.

Typically, an individual's scores on psychological tests are interpreted based on how they compare to the scores of a representative group called a normative sample. Normative samples can be drawn from any number of populations, such as a general population, a physician population, or a population of graduate students. The normative sample for the current version of the MicroCog was based on a stratified sample of the general population of the United States divided into age groups and three groups based on years of education (<12 years, 12 years, and >12 years). Because it was originally developed for physicians, the MicroCog shows virtually no ceiling effect and thus can detect cognitive deficits in well-educated, higher functioning individuals. As a screening instrument, the MicroCog provides clinicians with a valid, reliable, and well-normed assessment tool.

The 18 subtest scores obtained from the MicroCog contribute to three levels of index scores. The first set of index scores are made up of five neuropsychological content domains: Attention/Mental Control, Memory, Reasoning/Calculation, Spatial Processing, and Reaction Time. The second level of index scores represents processing speed and accuracy as two separate scores. The two global cognitive functioning scores on the third level incorporate both speed-of-processing and accuracy scores. Figure 1 is a visual depiction of the three levels of MicroCog Index Scores.

**Method**

**Participants**

Data for the group of physicians who had been referred for competency evaluations (competency group) were obtained from CPEP, the Center for Personalized Education for Physicians program. This unique, not-for-profit program has provided more than 900 competency evaluations for physicians from all over the United States since 1990; the program, based in Denver, Colorado, depends on more than 200 clinical consultants from a variety of specialties, many of whom are faculty members at the local medical school, the University of Colorado Denver School of Medicine. Beginning in 1997, the MicroCog became a part of the standard battery of tests used in evaluation. The physician competency group consisted of physicians who completed competency evaluations at CPEP from January 1997 to January 2004. The data were deidentified, and demographic information and MicroCog results were provided to the authors for use in this study.

The control group consisted of physicians who were recruited from a western metropolitan area via three sampling strategies: (1) physicians and other health professionals (e.g., psychologists) gave business cards with information about the study to physicians who might be interested in participating, (2) we sent e-mails inviting participation in the study to physicians on the staff of a major city hospital and a medical school, and (3) we invited physicians who had served as paid consultants with CPEP to participate in this study (these consultants were not involved in the phase of a CPEP evaluation that used the MicroCog). Control group data were collected in 2004 and 2005.

We established exclusion criteria for both groups. Exclusion criteria for the competency group included having taken the MicroCog more than once, having been born outside the United States,
having been trained or educated outside the United States, having reported a primary language other than English, or having presented for assessment for reasons other than quality of care concerns.

For the control group, exclusion criteria included visual, hearing, or physical impairment that might interfere with the testing; having been referred for a competency evaluation; having been born outside the United States, having been trained or educated outside the United States, having a primary language other than English; having taken, seen, or having familiarity with the MicroCog; and being retired and no longer working in the medical field.

Control physicians completed the MicroCog in quiet locations at local libraries, universities, and private offices. After informed consent was obtained, the examiner explained the testing procedures and familiarized the examinee with the appropriate procedures as outlined in the MicroCog manual. The MicroCog was administered on a laptop computer. Most physicians completed the test in 45 to 60 minutes. Once testing was completed, control physicians either received $50 or donated $50 to a charity of their choice. A lunch was also provided. This protocol was approved by the University of Denver institutional review board and the Colorado Multiple Institutional Review Board.

Statistical analyses

We reviewed all variables for data entry errors and missing data before undertaking statistical analysis. Scores for one domain index were missing for one participant. The index score missing from this participant’s data was used in two analyses; thus, this case was deleted from those analyses. Entries were determined to be within an acceptable accuracy range based on a review of extreme values for each variable used. All outliers were viewed individually and checked for accuracy. Unless otherwise noted, the scores used for all analyses were corrected for age and education (higher than high school). We used the cognitive proficiency score as a cognitive summary measure because it includes weighted accuracy and speed information from all of the subtests in the five cognitive domains. Accuracy data from the subtests constitute the accuracy summary score, and speed data from the subtests make up the processing speed summary score.

An alpha level of .05 was used for statistical tests. We calculated statistics using SPSS Graduate Pack 13.0 for Windows (Chicago, Illinois, 2004). In all but one analysis, for the variables that were determined not to have equal variance, the variance was greater in the larger group. According to Glass and Hopkins, if the greater variance is in the larger group, the “true probability of a type-I error is always less than the nominal probability.” Thus, violation of homogeneity of variance was not a statistical concern in these analyses. In the comparison of age between the two groups, the variance was greater in the smaller group; thus, a more stringent alpha level of .01 was used for that specific analysis, and the results reported were from the “equal variances not assumed” category. Differences between the two physician groups were examined using t-tests or χ² as appropriate.

In a second analysis, to determine whether the physicians with scores suggesting cognitive problems in the competency evaluation group accounted for the difference between the two physician groups, scores for physicians whose performance suggested cognitive problems were removed from the competency group. This then provided a comparison of the two groups without physicians whose scores suggested cognitive problems.

### Results

#### Participants

Of 367 potential candidates in the competency group, a total of 94 (26%) physicians were excluded based on exclusion criteria. Therefore, we examined data for a total of 279 candidates referred for competency evaluation.

A total of 143 physicians initially responded to the invitation to participate in the control group. Seventy-three physicians responded but did not schedule an appointment, of whom 17 (11.9%) were ineligible based on exclusion criteria. The remaining 56 physicians did not schedule an appointment because they were curious about the study but did not want to participate, could not find time to take the test, had personal issues such as illness or family concerns, or for unknown reasons. Seventy-four (49.0%) physicians made an appointment and 68 (47.5%) completed testing. All physicians in the control group also completed a demographic questionnaire.

#### Statistical analyses

No statistically significant difference was found between the mean age of the two physician groups \(t(86.13) = -1.27, P = .206\). The effect size for this analysis was very small (eta squared = 0.005). See Table 1 for descriptive statistics.

There was a significant difference in the proportion of female physicians to male physicians between the two groups.

#### Table 1

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Control (N = 68)</th>
<th>Competency (N = 267)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>41 (60.3%)</td>
<td>223 (83.5%)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Female</td>
<td>27 (39.7%)</td>
<td>44 (16.5%)</td>
<td></td>
</tr>
<tr>
<td><strong>Specialty</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary care</td>
<td>29 (42.6%)</td>
<td>126 (47.2%)</td>
<td>.003</td>
</tr>
<tr>
<td>Surgical</td>
<td>16 (23.5%)</td>
<td>98 (36.7%)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>23 (33.9%)</td>
<td>43 (16.1%)</td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>31–76</td>
<td>30–80</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>49.4</td>
<td>51.5</td>
<td>.206</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>12.5</td>
<td>9.1</td>
<td></td>
</tr>
</tbody>
</table>
In addition, there was a significant difference in the proportions of physician specialties in the two groups [Pearson χ² (335) = 11.643, P = .003]. The effect sizes for these two tests were small (w = 0.23 and w = 0.19, respectively). Evaluating differences in performance between physician specialties was beyond the scope of this study.

A two-tailed, independent-samples t test was used to determine if there was a difference between control group females (n = 27, M = 106.52, SD = 8.437) and males (n = 41, M = 111.66, SD = 9.082) on the global cognitive proficiency score. A significant difference [t(66) = 2.198, P = .022] was found between the two groups, with males scoring higher (better) than females. The mean difference between the two groups was 5.14, and the magnitude of difference between these two groups was moderate (eta squared = 0.077). A different result was found when this same comparison was made between genders in the competency evaluation group. There was no significant difference [t(265) = −1.005, P = .316] between males (n = 223, M = 95.48, SD = 12.655) and females (n = 44, M = 97.57, SD = 12.149) on a global proficiency score. The mean difference between the two groups was 2.09, with the degree of difference being small (eta squared = 0.004).

The physician competency group scored significantly lower than the control group on three MicroCog summary measures: processing speed [t(161) = 7.274, P < .001], processing accuracy [t(333) = 4.727, P < .001], and cognitive proficiency [t(139) = 10.23, P < .001], which incorporates both accuracy and speed. Table 2 provides descriptive statistics, and Figure 2 depicts the percentage distribution of proficiency scores for both groups.

### Table 2

**Significant Differences Between Study (Competency) and Control Groups on Three Index Scores in a Comparative Study of Cognitive Performance of Physicians Referred for Competency Evaluations at the Center for Personalized Education for Physicians, January 1997 to January 2004**

<table>
<thead>
<tr>
<th>Index</th>
<th>Control Mean</th>
<th>Control Standard Deviation</th>
<th>Competency Mean</th>
<th>Competency Standard Deviation</th>
<th>Mean Difference</th>
<th>P</th>
<th>Eta squared (effect size) *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing speed</td>
<td>108.94</td>
<td>11.011</td>
<td>96.55</td>
<td>17.284</td>
<td>12.391</td>
<td>&lt;.001</td>
<td>0.137 (large)</td>
</tr>
<tr>
<td>Processing accuracy</td>
<td>105.94</td>
<td>10.635</td>
<td>98.93</td>
<td>10.991</td>
<td>7.012</td>
<td>&lt;.001</td>
<td>0.063 (moderate)</td>
</tr>
<tr>
<td>Cognitive proficiency</td>
<td>109.62</td>
<td>9.126</td>
<td>95.83</td>
<td>12.575</td>
<td>13.79</td>
<td>&lt;.001</td>
<td>0.239 (large)</td>
</tr>
</tbody>
</table>


**Discussion**

The results from this study show significant differences between a group of physicians referred for competency evaluations and a physician control group on a neuropsychological screening measure. First, the physician control group scored significantly higher than the competency group on three cognitive global functioning scores (processing accuracy, processing speed, and cognitive proficiency) and four of five specific neuropsychological domains (Attention/Mental Control, Memory, Spatial Abilities, Reasoning and Calculation, and "Physicians' Competence"

**[Pearson χ² (335) = 16.142, P < .001]**

In addition, there was a significant difference in the proportions of physician specialties in the two groups [Pearson χ² (335) = 11.643, P = .003]. The effect sizes for these two tests were small (w = 0.23 and w = 0.19, respectively). Evaluating differences in performance between physician specialties was beyond the scope of this study.

A two-tailed, independent-samples t test was used to determine if there was a difference between control group females (n = 27, M = 106.52, SD = 8.437) and males (n = 41, M = 111.66, SD = 9.082) on the global cognitive proficiency score. A significant difference [t(66) = 2.198, P = .022] was found between the two groups, with males scoring higher (better) than females. The mean difference between the two groups was 5.14, and the magnitude of difference between these two groups was moderate (eta squared = 0.077).

A different result was found when this same comparison was made between genders in the competency evaluation group. There was no significant difference [t(265) = −1.005, P = .316] between males (n = 223, M = 95.48, SD = 12.655) and females (n = 44, M = 97.57, SD = 12.149) on a global proficiency score. The mean difference between the two groups was 2.09, with the degree of difference being small (eta squared = 0.004).

The physician competency group scored significantly lower than the control group on three MicroCog summary measures: processing speed [t(161) = 7.274, P < .001], processing accuracy [t(333) = 4.727, P < .001], and cognitive proficiency [t(139) = 10.23, P < .001], which incorporates both accuracy and speed. Table 2 provides descriptive statistics, and Figure 2 depicts the percentage distribution of proficiency scores for both groups.

**Physician cognitive difficulty** (which suggests need for further, more complete, neuropsychological evaluation) was defined as a Cognitive Proficiency Score (based on the age- and education-adjusted normative data provided for the MicroCog) more than 1 standard deviation (SD) below the mean, or any two index scores more than 1 SD below the mean. There was a significant difference between the proportion of physicians meeting one of these criteria in the two groups [Pearson χ² (335) = 20.54, P < .001]. In the competency group, 24% (65 physicians) [95% CI (0.19–0.29)] of the physicians scored in the range suggesting cognitive difficulty, compared with the control group in which 0% had scores suggesting cognitive problems. The effect size for this test was small (w = 0.238).

The control group performed significantly higher than the scores for the age- and education-adjusted normative sample provided in the MicroCog on three summary scores: processing speed [M = 108.94, SD = 11.01, mean difference = 8.94, t(67) = 6.696, P < .001]; processing accuracy [M = 105.94, SD = 10.64, mean difference = 5.94, t(61) = 4.6607, P < .001]; and cognitive proficiency [M = 109.62, SD = 9.126, mean difference = 9.62, t(67) = 8.69, P < .001]. Respective effect sizes were moderate (d = 0.596), small (d = 0.396), and moderate (d = 0.641).

The control group performed significantly higher than the competency group on four neuropsychological domains: Attention/Mental Control [t(131) = 9.228, P < .001]; Reasoning/Calculation [t(333) = 3.824, P < .001]; Memory [t(132) = 5.865, P < .001]; and Spatial Abilities [t(133) = 7.551, P < .001]. There was no significant difference in reaction time between the control physicians and the competency physicians [t(160.08) = 1.312, P = .191] (Table 3).

When scores for physicians whose performance suggested cognitive problems were removed from the competency group, we still found a significant difference in cognitive proficiency [t(268) = 6.781, P < .001] between the control group [n = 68, M = 109.62 (73rd percentile), SD = 9.126] and the group of physicians referred for competency evaluations [n = 202, M = 101.1 (52nd percentile), SD = 8.897]. The mean difference between these two groups was 8.514, and the magnitude of this difference was large (eta squared = 0.146).

**Discussion**

The results from this study show significant differences between a group of physicians referred for competency evaluations and a physician control group on a neuropsychological screening measure. First, the physician control group scored significantly higher than the competency group on three cognitive global functioning scores (processing accuracy, processing speed, and cognitive proficiency) and four of five specific neuropsychological domains (Attention/Mental Control, Spatial Abilities, Reasoning and Calculation, and...
Memory). These findings are similar to results found in a study by Williams and colleagues.6

Second, a significantly greater proportion of physicians in the competency group had scores suggesting cognitive difficulty. The 24% of physicians with scores suggesting cognitive difficulty is similar to the percentage found in the Turnbull et al11 study. Our study suggests that there is significantly greater likelihood of neuropsychological problems in a group of physicians referred for competency evaluations than in a group of physicians who had not been referred.

Third, when physicians who scored in the range suggesting cognitive difficulty were removed from the competency evaluation sample, the control group still scored significantly higher on a cognitive proficiency measure than the competency group.

These findings, along with results of several other studies,4,5,16,17 suggest that, as a group, physicians referred for competency evaluations have more cognitive difficulties than physicians whose competency is not in question. Because approximately one quarter of physicians referred for competency evaluation demonstrate possible cognitive problems,4,5 physician competency evaluation programs should include neuropsychological screening, followed by referral for comprehensive evaluation and follow-up, if indicated. Cognitive difficulties in some physicians could adversely affect their ability to maintain a competent and safe practice and to complete a recommended remediation successfully.7

Neuropsychological difficulties might be associated with a progressive disease, in which case it might be necessary to monitor a physician’s abilities while implementing a plan for early retirement.18 If there is a reversible cognitive problem, as suggested by Hanna et al7 and Kapur,18 then the recovery process could be monitored and a gradual return to practice could be responsibly coordinated. In either case, neuropsychological screening in physician competency programs should be used routinely to inform the evaluating team, physicians, hospitals, and medical boards about the need for a more comprehensive neuropsychological evaluation. The comprehensive evaluation may then be used to confirm or refute evidence of impairment, to clarify the nature and degree of impairment, and to determine whether impairment adversely affects a physician’s ability to practice medicine.

When a physician’s neuropsychological evaluation reveals a specific cognitive deficit, that information should be considered in developing a remediation plan. As noted by Hanna and colleagues,7 some physicians may not have the capacity to learn because of cognitive deficits. In such cases, health concerns impairing neuropsychological functioning should be addressed first. If addressing health concerns does not improve a physician’s cognitive functioning, then remediation may not be an option. For physicians with the capacity to learn, compensatory strategies could be taught and evaluated for

![Figure 2](image-url) Percentage distribution of proficiency scores on the MicroCog Assessment of Cognitive Functioning test for study (competency evaluation) and control groups in a comparative study of cognitive performance of physicians referred for competency evaluations at the Center for Personalized Education for Physicians, January 1997 to January 2004.

Table 3

<table>
<thead>
<tr>
<th>Domain</th>
<th>Control</th>
<th>Competency</th>
<th>Mean difference</th>
<th>P</th>
<th>Eta squared (effect size)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention/mental control</td>
<td>110.07</td>
<td>97.82</td>
<td>12.25</td>
<td>&lt;.001</td>
<td>0.21 (large)</td>
</tr>
<tr>
<td>Reasoning/calculation</td>
<td>106.87</td>
<td>99.70</td>
<td>14.17</td>
<td>&lt;.001</td>
<td>0.04 moderate</td>
</tr>
<tr>
<td>Memory</td>
<td>110.41</td>
<td>101.43</td>
<td>8.99</td>
<td>&lt;.001</td>
<td>0.09 (moderate to large)</td>
</tr>
<tr>
<td>Spatial abilities</td>
<td>108.79</td>
<td>99.72</td>
<td>9.08</td>
<td>&lt;.001</td>
<td>0.15 (large)</td>
</tr>
<tr>
<td>Reaction time</td>
<td>105.93</td>
<td>104.34</td>
<td>1.59</td>
<td>0.191</td>
<td>0.005 (small)</td>
</tr>
</tbody>
</table>

effectiveness in overcoming cognitive shortfalls. A remediation plan that takes cognitive difficulties into account would alleviate physician frustration at trying to learn new material presented in a way that is ineffective.

Perry and Crean\textsuperscript{17} were concerned that, on many neuropsychological tests, physicians are being compared with normative groups that do not provide a fair representation of their intellectual and neuropsychological abilities. Because of extensive education and above-average general intelligence,\textsuperscript{19} it would be logical to assume that physicians' performance on neuropsychological tests would be higher than that of a general population normative sample. According to the results of this study, the control group of physicians performed significantly higher than the average education (>12 years)-adjusted score of the normative group on measures of processing speed, processing accuracy, and cognitive proficiency. Therefore, a physician who functioned in the low average range when compared with a population of people with more than 12 years of education might function at a below-average range if he or she were compared to a physician population. Understanding how this level of performance translates into functioning in the practice of medicine would help determine the normative group with which a physician should be compared. In addition, for comprehensive neuropsychological evaluation, more information is needed on how scores are interpreted and used to decide level of cognitive functioning and ability to maintain a competent practice.

It should also be noted that when assessing neuropsychological functioning in adults, age corrections are typically used because cognitive functioning tends to decline with age. Therefore, adjustments to performance scores are made to correct for the decline in performance with increasing age, as was done in this study. Turnbull and colleagues\textsuperscript{5} questioned this practice when assessing physician competency, suggesting that no matter the age, physicians need to demonstrate a certain level of cognitive ability to maintain a safe and proficient practice. This is an interesting point and is one that could be explored in future studies.

This study has limitations that need to be considered when interpreting the results. Our findings show important group differences, but they may not reflect individual scores. There were several individuals from the competency evaluation group who performed exceptionally well on the MicroCog.\textsuperscript{8} Additionally, the control sample may not have been representative of the broader physician population. The control group who were recruited to take the test could have comprised higher-functioning individuals than the general physician population, which could have accentuated the difference in the two physician groups. However, the physician control group's scores were similar to the scores of other physician control groups.\textsuperscript{6,8} The control sample was collected from a western metropolitan area and may not be representative of physicians across the United States. However, because many of the control group physicians grew up and were trained elsewhere in the United States, the sample may be more diverse than might be assumed with the small catchment area. Additionally, the exclusion of foreign-born and/or foreign-trained physicians limits the generalizability of this study, because this screening instrument has not been validated on any foreign-born group.

Although there was a significant difference between the performance of male and female physicians in the control group, that difference was not present in the physician competency group. In the control group, male physicians obtained higher proficiency score than female physicians. The control group, however, had a greater proportion of females than the group of physicians referred for competency evaluations. Thus, the difference in gender ratios between the two physician groups would not have artificially increased the difference between the two groups.

The instrument was a 60-minute neuropsychological screen and thus did not provide the depth of information that a full neuropsychological battery could provide. An impaired score on the MicroCog ought not to be viewed as a definitive result and should be followed up with a full neuropsychological battery to determine and define possible cognitive impairment. However, the shorter time commitment required for testing with the MicroCog did facilitate recruitment of practicing physicians and the ability to test in a variety of settings, making it feasible as a first step in evaluating cognition.

The two groups of physicians were tested under different personal circumstances. The referred physicians were involved in an evaluation that could ultimately have a significant impact on their career and their livelihood. The anxiety and pressure created by that scenario could have negatively impacted their performance. This group of physicians would also be highly motivated to perform well on the test. On the other hand, the control group physicians had little fear that the results could affect their practice. Even so, several of the control group physicians expressed anxiety and concern over their performance. In fact, a few control physicians indicated they wanted to be tested because of perceived changes in their own cognitive abilities. It is doubtful that the number and magnitude of statistically significant results in this study could be fully explained by these different testing scenarios.

It is important to understand that the evaluation of a physician's cognitive functioning is only one factor that could affect a physician's performance. Using a neuropsychological screen as a measure, 24% of the competency evaluation physicians scored in a range suggesting cognitive difficulties. Seventy-six percent scored in a range suggesting no cognitive difficulties. Physician competency is a much more complex issue than just testing physicians' levels of neuropsychological functioning. Though tests like the MicroCog could be used to screen for cognitive deficits that contribute to physician competency, it is not a panacea for the myriad issues that contribute to physician competency and performance. However, it is important that physicians, hospitals, physician competency evaluation programs, and medical boards begin discussions about the use of neuropsychological assessment and how neuropsychological difficulties impair physician competency.

More research is needed to identify the specific neuropsychological domains that impact particular areas of physician performance.\textsuperscript{4} Also, future research will benefit from a clear delineation of how cognitive impairment is defined for physicians. Another area of research that could contribute to our understanding of
physicians’ cognitive functioning is looking at what part wisdom—the accumulation of knowledge and experience—plays in compensating for age-related decline. Much could be gained from studying clearly capable practicing older physicians, their neuropsychological characteristics, and what dynamic efforts they employ to maintain a competent practice.

Being a physician is a high and difficult calling. We hope these results will be used to pursue competent clinical performance for all physicians and for identifying, evaluating, and rehabilitating physicians struggling with competency concerns.

Acknowledgments
The authors wish to thank Gerald Carpenter, PsyD, for assistance with physician recruitment for the control group; Kathy Green, PhD, for assistance with study design and statistical analysis; Thomas Paskus, PhD, for assistance with study design and statistical analysis; and Maria Riva, PhD, for assistance with study design.

Disclaimer

Colorado Mental Health Institute at Fort Logan, Denver, Colorado is not affiliated with this research.

“MicroCog” is a trademark of The Risk Management Foundation of the Harvard Medical Institutions, Inc.

References