Rapid Cognitive Assessment of Nursing Home Residents
A Comparison of the Brief Interview for Mental Status (BIMS) and Brief Cognitive Assessment Tool–Short Form (BCAT-SF)

Ryan A. Mace, BS; William E. Mansbach, PhD; and Kristen M. Clark, BS

ABSTRACT
In nursing homes, the ability to identify residents with cognitive impairment is critical; however, many providers fail to address symptoms of dementia due to insufficient time to assess cognition. In the current study, the authors compared two rapidly administered instruments, the Brief Interview for Mental Status (BIMS) and Brief Cognitive Assessment Tool–Short Form (BCAT-SF), in predicting specific cognitive stages. Two hundred twenty-five nursing home residents who were referred for neurocognitive evaluation and met inclusion criteria participated in the study. Both cognitive instruments were found to predict cognitive diagnoses generally, but only the BCAT-SF demonstrated sensitivity in differentiating among all cognitive levels. Receiver operating characteristic analyses did not confirm cutoff scores reported in the original BIMS study. However, sensitivity (0.91) and specificity (0.96) values for the BCAT-SF were robust for differentiating between dementia and non-dementia. The BCAT-SF appears to be a more sensitive instrument for predicting cognitive stages than the BIMS for nursing home residents.

Cognitive impairment is prevalent in American nursing homes, as approximately 70% of residents have some type of cognitive impairment, much of it severe (Centers for Medicare & Medicaid Services [CMS], 2013). Studies estimate that approximately 50% of all residents meet criteria for dementia (Magaziner et al., 2000). The ability to identify residents with cognitive impairment is critical for resident-centered care to be effective and for staff to meet residents’ needs (Singer & Luxenberg, 2003). Previous research has documented that inaccurate understanding of cognitive functioning can adversely affect management of medical conditions and resident life (Cohen-Mansfield, 2005; Cohen-Mansfield & Creedon, 2002). Since the 1980s, when the U.S. Congress legislated sweeping changes in the nursing home industry through the Omnibus Budget Reconciliation Act of 1987, a key provision and guiding principle has been the mandatory use of a standardized, comprehensive system to assist in assessment and care planning (Hawes et al., 1997). Recognizing the high prevalence and complex care needs of residents with dementia, this Mini-
Mace, Mansbach, & Clark

The mini Data Set (MDS) has included cognitive assessments and has since been an evolving process.

Despite the potential use of screening older adults who present with a subjective or objective memory complaint, many providers fail to address symptoms of dementia due to insufficient time to assess the cognition of their patients (Boise, Camicioli, Morgan, Rose, & Congleton, 1999; Østbye et al., 2005). For nursing home providers to use cognitive screening tools, they must be brief and easily administered in addition to being psychometrically robust (Mansbach & MacDougall, 2012). When MDS 3.0 (the most recent iteration) was introduced, the Brief Interview for Mental Status (BIMS; Chodosh et al., 2008) was selected as the mandated cognitive assessment tool. Some of the strengths of the BIMS are that it is based on an interview format, has acceptable reliability and convergent validity, can be rapidly administered (average time is approximately 3 minutes), and appears to be more accurate than observational methods of assessment (Chodosh et al., 2008).

However, the BIMS has some shortcomings that may limit its use as a nursing home cognitive measure. First, it is not designed to be sensitive to the full cognitive continuum. Second, the cutoff scores differentiating those residents with and without cognitive impairments are based, in part, on the Modified Mini-Mental State Examination (3MS; Teng & Chui, 1987), a measure that may not be sensitive to mild cognitive impairment (MCI) (McDowell, Kristjansson, Hill, & Hébert, 1997). Third, memory is assessed using a simple word list (i.e., three words), with no story recall component. There is empirical support for using story recall items in cognitive screening instruments, especially for identifying patients with MCI (Rabin et al., 2009). Story recall has syntactic and other contextual elements largely absent from word lists that may differentially affect recall (Foldi, 2011; Waters & Caplan, 2005) and learning (Kintsch, 1994). Fourth, despite the fact that deficits in executive functions are common features of dementia, the BIMS used in MDS 3.0 does not assess executive capacities. Fifth, the BIMS was normed on a Veterans Administration population in which approximately 95% of the sample were men. Most nursing homes have more women than men, and it is unclear if women would have different normative values.

For the current study, the authors selected the Brief Cognitive Assessment Tool–Short Form (BCAT-SF; Mansbach & MacDougall, 2012), as it appears to address many of these limitations. The BCAT-SF is an abridged instrument of the original Brief Cognitive Assessment Tool (BCAT; Mansbach, MacDougall, & Rosenzweig, 2012) and was designed to be used in situations in which time is limited (Mansbach & MacDougall, 2012). The BCAT-SF can be completed in 5 minutes or less. It contains six items representing the cognitive domains of contextual memory, executive functions, and attentional capacity. The BCAT-SF preserves the psychometric rigor of the full BCAT in predicting cognitive diagnosis (i.e., dementia versus non-dementia) and functional status (i.e., basic activities of daily living [ADLs] and instrumental ADLs [IADLs]) (Mansbach & MacDougall, 2012).

The primary purpose of the current study was to compare the utility of the BCAT-SF and BIMS in predicting cognitive diagnoses in nursing home patients. An additional goal was to cross-validate the BCAT-SF and BIMS in a nursing home population. In so doing, the authors attempted to determine which brief instrument is more sensitive in identifying levels of cognitive functioning.

METHOD

Participants

Three hundred seven individuals residing in a Maryland skilled nursing facility were referred for neurocognitive evaluation (October 2012 to April 2014). Approximately all patients were residents receiving short-term subacute rehabilitation following hospitalization. Two hundred twenty-five individuals met inclusion criteria for retrospective data analysis, which required a completed BCAT-SF and BIMS, proficiency in English, and age older than 59. Individuals with medical or psychiatric symptoms that would prevent completion of the tests (e.g., global aphasia, disabling depression) were excluded.

Participants ranged in age from 61 to 102, with a mean age of 80.44 (SD = 8.98 years). As shown in Table 1, 61% of participants were female, 72% were Caucasian, 49% were widowed, and 78% had ≥12 years of education. Of the sample, 56% met diagnostic criteria for dementia (i.e., mild, moderate, or severe), 28% met diagnostic criteria for MCI, and 10% did not have a dementia diagnosis. Using a Geriatric Depression Scale Short Form (GDS-SF) cutoff score of >5 for self-reported depression (Sheikh & Yesavage, 1986), 28% of the sample showed evidence of depression, whereas 72% were categorized as not having depression.

Procedure

The nursing home’s Medical Ethics Committee approved all procedures. All study participants were referred for neurocognitive evaluation by nursing home attending physicians. A licensed psychologist with specific exper-
tise in the evaluation of cognitive dysfunction conducted all evaluations. Evaluations consisted of: (a) a clinical interview with the patient; (b) comprehensive review of the medical chart, inclusive of hospital discharge summaries (if appropriate); (c) administration of the AD8 Dementia Screening Interview (AD8; Galvin et al., 2005), Kitchen Picture Test (KPT; Mansbach, MacDougall, Clark, & Mace, 2014), and GDS-SF; (d) staff interviews; and (e) interviews with family members. The AD8 and KPT were used in the clinical evaluation to help determine cognitive diagnosis and confirm convergent validity for the BCAT-SF and BIMS. The GDS was used to evaluate mood functioning as part of the clinical evaluation and confirm discriminate validity with respect to the BCAT-SF and BIMS.

Diagnoses were made by the licensed psychologist using the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2000) criteria to diagnose dementia, with Clinical Dementia Rating Scale (CDR; Morris, 1993) criteria for the staging of dementia as mild, moderate, or severe. The CDR 0.5 and 1.0 were combined, as they are both mild impairment categories, to better capture those participants who meet criteria for MCI. The diagnosis of MCI was further confirmed based on the Petersen et al. (1999) criteria. Cognitive diagnoses were made without knowledge of BIMS and BCAT-SF scores. Four social workers trained in administering the BIMS independently administered this screening tool and two trained mental health technicians administered the BCAT-SF. Test order was determined by assigning participants to one of two protocols. In Protocol A, the order of testing was: BCAT-SF, AD8, KPT, GDS-SF, and BIMS. In Protocol B, the order of testing was: BIMS, BCAT-SF, KPT, AD8, and GDS-SF. Those who administered the BIMS had no knowledge of the BCAT-SF scores. Those who administered the BCAT-SF had no knowledge of BIMS scores. Finally, the licensed psychologist making the clinical diagnoses had knowledge of the AD8, KPT, and GDS-SF scores to help make cognitive and mood diagnoses, but did not have knowledge of the BIMS and BCAT-SF test results.

**Measures**

**Brief Interview for Mental Status.** The 7-item BIMS (Chodosh at al., 2008) is an integral part of MDS 3.0 mandated for American nursing homes (Saliba et al., 2012). The BIMS focuses on two domains of cognitive function, memory and orientation, with scores ranging from 0 to 15. In the CMS RAI Version 3.0 Manual (CMS, 2010, p. 14), the following BIMS score ranges are recommended for assessing cognition: 13 to 15 = cognitively intact, 8 to 12 = moderately impaired, and 0 to 7 = severe impairment. Verbal memory on the BIMS is assessed using immediate and delayed free recall (i.e., 1- to 2-minute delay) of an orally presented word list (i.e., three words). The development study (N = 374) established interrater reliability between research assistant and facility nurse administrations of the BIMS within 72 hours of each other (r = 0.72) and found a correlation (r = 0.79, p < 0.01) with the 3MS (Teng & Chui, 2008).
1987). Chodosh et al. (2008) selected BIMS cutoff scores, based on benchmarks from the 100-point 3MS, for identifying severe impairment or probable dementia (3MS <48). A cutoff score of 10 (3MS <48) had a sensitivity of 0.94 and specificity of 0.79 for identifying severe impairment (Chodosh et al., 2008).

**Brief Cognitive Assessment Tool–Short Form.** The BCAT-SF (Mansbach & MacDougall, 2012) was designed as a shorter version of the full BCAT to be used when time for cognitive screening is constrained (e.g., primary care settings). The BCAT-SF can be easily administered in less than 5 minutes, with scores ranging from 0 to 21. To maintain content validity, Mansbach and MacDougall (2012) selected six items from the original BCAT that represent the cognitive domains of contextual memory, executive functioning, and attentional capacity. One hundred four older adults from assisted living facilities (87.5%, n = 91) and the independent section within continuing care retirement communities (12.5%, n = 13) participated in the development and validation of the BCAT-SF. Psychometric analyses yielded an internal consistency reliability of 0.86 (Cronbach’s alpha based on standardized items), a test–retest reliability of \( r = 0.98 \), and preliminary evidence of the construct validity of BCAT-SF score inferences through convergent, discriminant, and predictive validity analyses. Using a cutoff score of 16, the BCAT-SF differentiated between MCI and dementia with a sensitivity of 0.90 and a specificity of 0.81 (area under the curve = 0.93) (Mansbach & MacDougall, 2012). Similar to the BCAT, the BCAT-SF has been shown to predict both ADL and IADL functioning, as measured by the Lawton and Brody (1969) Physical Self-Maintenance Scale and Lawton IADL scale.

**Convergent Validity**

*AD8 Dementia Screening Interview.* The AD8 (Galvin et al., 2005) was designed to be a brief informant-based measure to differentiate individuals with and without dementia and is sensitive in detecting early signs of dementia. The 8-item screening interview was designed to assess memory, orientation, judgment, and function. Scores range from 0 to 8, with scores ≥2 predicting a likelihood of dementia in a sample of older adults with a 38% base rate of very mild dementia (i.e., CDR = 0.5). Galvin et al. (2006) investigated the psychometric properties of the scale through 225 patient–informant dyads. Evidence for interrater reliability through correspondence between independent raters (intraclass correlation coefficient = 0.82; 95% confidence interval [CI] [0.5 to 0.92]), intrarater reliability through the agreement of repeated administrations by a single rater (weighted kappa = 0.67; 95% CI [0.59 to 0.75]), and construct validity through correlations of individual AD8 items and other neuropsychological tests were found (Galvin et al., 2006). In the absence of an informant, Galvin, Roe, Coats, and Morris (2007) demonstrated that a self-reported AD8 can differentiate between individuals with and without dementia, albeit with greater use in individuals with mild impairment.

**Kitchen Picture Test (KPT).** The KPT (Mansbach, MacDougall, et al., 2014) was designed as a visually presented test of practical judgment. The KPT is an illustration of a kitchen scene in which three potentially dangerous situations are unfolding. Participants are asked to identify, rank the severity of, and provide solutions that would resolve these problems. Scores range from 0 to 8. In two independent studies (Study 1, \( N = 99 \) nursing home patients; Study 2, \( N = 163 \) nursing home and assisted living patients), psychometric analyses indicated strong evidence for reliability, construct validity, and predictive validity (Mansbach, MacDougall, et al., 2014). In addition to detecting problems in basic judgment, the KPT can significantly differentiate between individuals with and without dementia, with a suggested cutoff score of 7 (Mansbach, MacDougall, et al., 2014).

**Divergent Validity**

*Geriatric Depression Scale Short Form.* The 15-item GDS-SF (Sheikh & Yesavage, 1986) is the most commonly used scale to screen for depression in older adults. Scores range from 0 to 15, with higher scores indicating a greater likelihood of depression. Researchers have reported validity data for the GDS-SF from a variety of settings, including a nursing home setting (Gerety et al., 1994). de Craen, Heeren, and Gussekloo (2003) reported adequate sensitivity and specificity in adults older than 85. However, de Craen et al. (2003) also stated that the GDS-SF is most useful for older adults with minimal to no cognitive impairment. A reliability generalization study reported average GDS reliability (including all short forms) to be 0.85 (Kiefer & Reese, 2002). Chiang, Green, and Cox’s (2009) Rasch analysis from 177 older adult participants in a variety of residential locations (from retirement communities to skilled nursing facilities) found the GDS-SF to assess a unidimensional construct.

**Statistical Analysis**

All statistical analyses were performed using SPSS version 21.0. Descriptive statistics were used to report participant demographics, assessment tool scores, and
Clinical diagnoses. Internal consistency was estimated by Cronbach's alpha reliability coefficient. Convergent and divergent validity were addressed with Spearman correlation coefficients. A one-way multivariate analysis of variance (MANOVA) and the Games–Howell post hoc test were selected to investigate mean differences of BCAT-SF and BIMS scores between diagnostic/CDR categories (i.e., normal, MCI, mild dementia, moderate dementia, and severe dementia). A hierarchical multiple regression analysis was used to estimate the amount of variance in cognitive diagnoses explained by the BCAT-SF and BIMS scores. The predictive utility of the BIMS and BCAT-SF was also addressed with analyses of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Cutoff scores for identifying participants with dementia were selected that yielded the optimal balance between sensitivity and specificity, while emphasizing a higher sensitivity. Receiver operating characteristic (ROC) curves and area under the ROC were calculated from sensitivity and 1-specificity values for differentiating diagnoses of MCI and dementia.

**RESULTS**

In the current sample, the distribution of BIMS (range = 2 to 15) and BCAT-SF scores (range = 0 to 21) were negatively skewed. Mean BIMS and BCAT-SF scores were not significantly different for marital status, race, or sex. However, participants with less than 12 years of education scored lower on the BCAT-SF (median = 13, \( n = 41 \)) than those with 12 or more years of education (median = 15, \( n = 176 \)), \( U = 2799.0, z = -2.59, p = 0.01 \).

The BIMS (\( r_s = -0.27, p < 0.001 \)) and BCAT-SF (\( r_s = -0.38, p < 0.001 \)) were significantly correlated with age.

**Reliability**

In the current study, the internal consistency reliability, as estimated by Cronbach’s alpha based on standardized items, was 0.84 for the BCAT-SF and 0.77 for the BIMS. Both coefficients are within the acceptable range for tests that are used for clinical decision making (Nunnally & Bernstein, 1994).

**Validity**

Table 2 presents descriptive statistics used for validity analyses measured in the current study. As reported in Table 3, construct validity was demonstrated by significant correlations between the BIMS and BCAT-SF (\( r_s = 0.59, p < 0.001 \)). Further convergence validity was confirmed.

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Descriptive Statistics for Validity Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
<td>( n )</td>
</tr>
<tr>
<td>BCAT-SF</td>
<td>225</td>
</tr>
<tr>
<td>BIMS</td>
<td>225</td>
</tr>
<tr>
<td>AD8</td>
<td>195</td>
</tr>
<tr>
<td>KPT</td>
<td>204</td>
</tr>
<tr>
<td>GDS-SF</td>
<td>184</td>
</tr>
</tbody>
</table>

**Note.** BCAT-SF = Brief Cognitive Assessment Tool–Short Form (0 to 15 = dementia, 16 to 21 = no dementia); BIMS = Brief Interview for Mental Status (0 to 7 = severe impairment, 8 to 12 = moderately impaired, 13 to 15 = cognitively intact); AD8 = AD8 Dementia Screening Interview (0 to 1 = no dementia, 2 to 8 = dementia); KPT = Kitchen Picture Test (0 to 6 = dementia, 7 to 8 = no dementia); GDS-SF = Geriatric Depression Scale Short Form (0 to 5 = no depression, 6 to 15 = depression).

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Matrix of Spearman’s Rho Correlation Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Age</td>
</tr>
<tr>
<td>Age</td>
<td>—</td>
</tr>
<tr>
<td>BCAT–SF</td>
<td>—</td>
</tr>
<tr>
<td>BIMS</td>
<td>—0.27***</td>
</tr>
<tr>
<td>AD8</td>
<td>0.08</td>
</tr>
<tr>
<td>KPT</td>
<td>—0.35</td>
</tr>
<tr>
<td>GDS–SF</td>
<td>—0.00</td>
</tr>
</tbody>
</table>

**Note.** BCAT-SF = Brief Cognitive Assessment Tool–Short Form; BIMS = Brief Interview for Mental Status; AD8 = AD8 Dementia Screening Interview; KPT = Kitchen Picture Test; GDS-SF = Geriatric Depression Scale Short Form.

\( * p < 0.05, ** p < 0.01, *** p < 0.001 \) (two-tailed test).
by comparing the BCAT-SF and BIMS to two measures of cognitive impairment. The BIMS ($r_s = –0.20$, $p < 0.01$) and BCAT-SF ($r_s = –0.28$, $p < 0.001$) were significantly negatively correlated with scores on the AD8, which measures self-reported changes in cognition. Negative correlations are to be expected because items on the AD8 are reverse-scored relative to the BIMS and BCAT-SF. The BIMS ($r_s = 0.52$, $p < 0.001$) and BCAT-SF ($r_s = 0.68$, $p < 0.001$) were significantly correlated with the KPT, a measure of basic judgment that can differentiate between dementia and non-dementia. Table 3 also shows discriminant validity, which was supported by nonsignificant correlations between the BIMS ($r_s = –0.05$, $p = 0.50$), BCAT-SF ($r_s = –0.03$, $p = 0.66$), and GDS-SF, a measure of depression.

A one-way MANOVA was conducted to explore the mean differences of the BIMS and BCAT-SF scores between levels of cognitive functioning based on five categories: (a) normal, (b) MCI, (c) mild dementia, (d) moderate dementia, and (e) severe dementia. The cognitive assessment tool scores by levels of cognitive functioning are shown in Table 4. A statistically significant difference was found in BCAT-SF scores between the five cognitive categories ($F(4, 207) = 241.13, p < 0.001$). Likewise, a statistically significant difference was found in BIMS scores between the five cognitive categories ($F(4, 207) = 41.12, p < 0.001$). The effect size, calculated using partial eta squared, for the BCAT-SF and BIMS scores, was 0.82 and 0.44, respectively. Having met one-way MANOVA statistical significance, multiple comparisons were conducted for the BIMS and BCAT-SF scores among all five cognitive levels. Games–Howell post hoc tests revealed significant mean differences for BCAT-SF scores between all five CDR groups ($p < 0.001$ for all tests). However, mean BIMS scores did not significantly differ between normal and MCI ($p = 0.74$) or between mild dementia and moderate dementia ($p = 0.07$).

A hierarchical multiple regression analysis was also used to test the ability of the BIMS and BCAT-SF in significantly predicting diagnoses of normal, MCI, mild dementia, moderate dementia, and severe dementia. As shown in Table 5, the BIMS was entered at Step 1, explaining 42% of the variance in cognitive diagnosis ($R^2 = 0.42$, $F(1, 210) = 151.63, p < 0.001$). After entry of the BCAT-SF at Step 2, the total variance explained by the model as a whole was 82% ($R^2 = 0.82$, $F(2, 210) = 461.14, p < 0.001$).

### Table 4

<table>
<thead>
<tr>
<th>Diagnostic Category</th>
<th>BCAT-SF</th>
<th></th>
<th></th>
<th></th>
<th>BIMS</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Mean</td>
<td>SD</td>
<td>Range</td>
<td>Median</td>
<td>Mean</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td>No dementia</td>
<td>21</td>
<td>20.18</td>
<td>1.22</td>
<td>16 to 21</td>
<td>15</td>
<td>14.18</td>
<td>1.47</td>
<td>10 to 15</td>
</tr>
<tr>
<td>Mild cognitive impairment</td>
<td>18</td>
<td>18.02</td>
<td>1.96</td>
<td>10 to 21</td>
<td>14</td>
<td>13.73</td>
<td>1.59</td>
<td>8 to 15</td>
</tr>
<tr>
<td>Mild dementia</td>
<td>14</td>
<td>13.67</td>
<td>2.83</td>
<td>3 to 20</td>
<td>12</td>
<td>11.58</td>
<td>2.93</td>
<td>3 to 15</td>
</tr>
<tr>
<td>Moderate dementia</td>
<td>10</td>
<td>9.72</td>
<td>2.42</td>
<td>4 to 15</td>
<td>11</td>
<td>9.84</td>
<td>3.41</td>
<td>3 to 15</td>
</tr>
<tr>
<td>Severe dementia</td>
<td>4</td>
<td>3.93</td>
<td>2.67</td>
<td>0 to 10</td>
<td>6</td>
<td>6.63</td>
<td>3.81</td>
<td>2 to 14</td>
</tr>
</tbody>
</table>

Note: BCAT-SF = Brief Cognitive Assessment Tool–Short Form; BIMS = Brief Interview for Mental Status.

### Table 5

<table>
<thead>
<tr>
<th>Variable</th>
<th>Step 1</th>
<th></th>
<th></th>
<th>Step 2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE B</td>
<td>B</td>
<td>SE B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>BIMS</td>
<td>–0.21</td>
<td>0.02</td>
<td>–0.65**</td>
<td>–0.05</td>
<td>0.01</td>
<td>–0.15**</td>
</tr>
<tr>
<td>BCAT-SF</td>
<td>–0.17</td>
<td>0.01</td>
<td>–0.81**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R^2</td>
<td>0.42</td>
<td></td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delta R^2</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F for Delta R</td>
<td>151.63</td>
<td></td>
<td>461.14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: SE = standard error; BIMS = Brief Interview for Mental Status; BCAT-SF = Brief Cognitive Assessment Tool–Short Form.

*p < 0.01, **p < 0.001.
The BCAT-SF accounted for an additional 40% of the variance in diagnoses over and above BIMS scores.

Finally, the predictive utility of the BIMS and BCAT-SF for identifying dementia was addressed with analyses of sensitivity, specificity, PPV, and NPV. As shown in Table 6, a BIMS cutoff score of <13 (scores ≥13 indicate no dementia), yielded a sensitivity of 0.66 and a specificity of 0.87 (PPV = 0.88, NPV = 0.63). A BCAT-SF cutoff score of <16 (scores ≥16 indicate no dementia), yielded a sensitivity of 0.91 and a specificity of 0.96 (PPV = 0.97, NPV = 0.87). An ROC was calculated to differentiate diagnoses of MCI and dementia (i.e., mild, moderate, and severe). The resulting ROC curves are compared in the Figure. For identifying dementia, the area under the ROC curve for the BIMS was 0.84 ($p < 0.001$, 95% CI [0.78, 0.89]) and 0.97 for the BCAT-SF ($p < 0.001$, 95% CI [0.95, 0.99]).

### DISCUSSION

The current study compared the abilities of the BIMS and BCAT-SF, two rapidly administered assessment instruments, in predicting cognitive diagnoses and accurately differentiating among cognitive functioning levels (i.e., normal, MCI, mild dementia, moderate dementia, and severe dementia) in nursing home residents. Both instruments demonstrated robust reliability and construct validity. However, meaningful differences were found with respect to their abilities to differentiate among levels of cognitive functioning. The BIMS significantly predicted cognitive diagnosis in general, but did not significantly differentiate between residents with normal cognition and those who had MCI. Furthermore, it did not significantly differentiate between individuals with mild dementia and moderate dementia. The BCAT-SF was found to significantly predict cognitive diagnosis and demonstrated sensitivity in differentiating among all five cognitive levels.

Perhaps this superior sensitivity can be attributed to scale construction. The BCAT-SF was designed to briefly assess three cognitive domains critical to cognitive functioning: (a) attentional capacity, (b) contextual memory, and (c) executive functions. The memory component is especially broad for a rapid cognitive screening tool (four items) and emphasizes story recall to assess verbal memory. Previous research has demonstrated the predictive utility of story recall metrics in assessing cognitive functioning in patients with impaired executive skills (Mansbach, Mace, & Clark, 2014), a problem common to nursing home residents who often have moderate to severe cognitive impairment. Furthermore, in their analysis of the use of individual verbal memory metrics in predicting cognitive functioning, Mansbach, Mace, et al. (2014) found that the BCAT delayed story recognition and delayed story re-

### TABLE 6

Selected Cutoff Scores Based on Predictive Utility for Identifying Dementia

<table>
<thead>
<tr>
<th>Measure</th>
<th>Cutoff Score</th>
<th>AUC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCAT-SF</td>
<td>&lt;16</td>
<td>0.97</td>
<td>0.91 (0.84, 0.95)</td>
<td>0.96 (0.89, 0.99)</td>
<td>0.97 (0.93, 0.99)</td>
<td>0.87 (0.78, 0.93)</td>
</tr>
<tr>
<td>BIMS</td>
<td>&lt;13</td>
<td>0.84</td>
<td>0.66 (0.57, 0.74)</td>
<td>0.87 (0.78, 0.93)</td>
<td>0.88 (0.80, 0.94)</td>
<td>0.63 (0.54, 0.72)</td>
</tr>
</tbody>
</table>

Note. AUC = area under the curve; PPV = positive predictive value; NPV = negative predictive value; BCAT-SF = Brief Cognitive Assessment Tool–Short Form; BIMS = Brief Interview for Mental Status.

* Clinical Dementia Rating Scale cutoff scores were used to classify participants into dementia/no dementia (including mild cognitive impairment) diagnostic categories.

* 95% confidence interval in parentheses.
call items were stronger predictors of cognitive diagnosis than the BIMS delayed word list.

Furthermore, the BCAT-SF contains a cognitive set-shifting item, which in the original BCAT study was found to have the strongest factor loading for the executive functions principle component analysis cluster, and individually strongly predicted cognitive diagnosis (Mansbach et al., 2012). There is added utility to cognitive set-shifting items, as they have been found to be sensitive predictors of MCI to Alzheimer’s disease conversion (Gomar, Bobes-Bascaran, Conejero-Goldberg, Davies, & Goldberg, 2011). The BIMS used in nursing homes primarily relies on a simple word list to assess memory, and as Chodosh et al. (2008) and Saliba et al. (2012) point out, does not assess executive functions.

From a clinical perspective, cutoff scores are useful for estimating cognitive functioning. In the current study, the authors emphasized differentiating between patients with and without dementia. This decision was based on the recognition that MCI is a unique clinical syndrome (with subtypes) that may progress to dementia (Bruscoli & Lovestone, 2004; Petersen et al., 2001). Therefore, cognitive instruments that can differentiate between MCI and dementia may be of particular value from a treatment and symptom management perspective. For nursing home patients, the authors recommend a BCAT-SF cutoff score of 16 to differentiate between individuals with dementia and individuals with MCI/normal functioning (i.e., patients with scores ≥16 are likely to have MCI or normal cognition, and individuals with scores <16 are likely to have dementia). This cutoff score is the same value recommended in the original development and validation study (Mansbach & MacDougall, 2012); therefore, the current study appears to cross-validate the original findings.

The current study did not confirm cutoff scores reported in the original BIMS study; however, Chodosh et al. (2008) found that a cutoff score of 10 had a sensitivity of 0.94 and specificity of 0.79 for detecting severe impairment. The current authors found different values using this cutoff score, especially for sensitivity (0.42) and specificity (0.98). In the current study, a BIMS cutoff score of 13 had the best balance of sensitivity (0.66), specificity (0.87), PPV (0.88), and NPV (0.63) for severe impairment. The fact that the authors could not validate the original values might be explained by the fact that the current study was based largely on nursing home residents receiving rehabilitation services, whereas the BIMS sample may have had proportionately more long-term care residents. In this respect, these samples may tap into somewhat different populations. Another possible explanation is that Chodosh et al. (2008) BIMS ROC analyses were based, in part, on the 3MS, which appears to have poor sensitivity in detecting MCI (McDowell et al., 1997). Therefore, the BIMS may be more vulnerable to diagnostic errors.

LIMITATIONS

There are important caveats and study limitations that bear highlighting. First, the BIMS and BCAT-SF are screening or single instrument assessments. The actual diagnosis of MCI and dementia should only be made with confirmatory evidence. Second, the current study is based primarily on test scores from nursing home short-term rehabilitation patients. It is possible that long-term care nursing home residents may represent a somewhat different population. Because PPV and NPV values are dependent on the base rate of cognitive impairment, different populations may yield different PPV and NPV values. Third, although the original BCAT-SF validation study did not find an education bias, the current study revealed a modest education effect. Mild educational biases are not unusual for screening tools, and might be attributed to the higher proportion of post-secondary education in this sample compared to the original BCAT-SF study. Of course, the clinician should take the patient’s education into account. Fourth, although independent cognitive diagnoses were made using multiple clinical sources, a “gold standard” cognitive measure was not used, which is in contrast to the original BCAT-SF study where comprehensive neuropsychological evaluations were used to determine cognitive diagnoses. Finally, interrater reliability data for those administering the BIMS and BCAT-SF were not available in the current study. Future comparisons of the two measures should determine whether a possible lack of agreement between raters of the BIMS and BCAT-SF accounts for differences in predictive utility.

CONCLUSION

For nursing home providers to use and optimize cognitive screening tools, a balance of instrument administration brevity and psychometric rigor is key. The issue of insufficient time to complete assessments and plan effective interventions based on test results is often reported as a barrier (Boise et al., 1999; Østbye et al., 2005). The BCAT-SF, which can be completed in less than 5 minutes and effectively identify patients at different levels of cognitive functioning, appears to be a promising choice. Although the current study does not refute the validity of the BIMS as a cognitive screening tool, it suggests that it may not be
optimal in a nursing home population. Given the shortcomings of the BIMS, despite the fact that it is currently the mandated cognitive screening tool for MDS 3.0, providers may elect to also administer the BCAT-SF. In the current study, the BCAT-SF accounted for an additional 40% of the variance in cognitive diagnoses over and above BIMS scores.

REFERENCES
Galvin, J.E., Roe, C.M., Coats, M.A., & Morris, J.C. (2007). Patient's rating of cognitive ability using the AD8, a brief informant interview, as a self-rating tool to detect dementia. Archives of Neurology, 64, 725-730. doi:10.1001/archneur.64.5.725


