An Empiric Approach to Level of Care Determinations: The Importance of Executive Measures

Donald R. Royall, Laura K. Chiodo, and Marsha J. Polk

Departments of Psychiatry, Medicine, and Pharmacology, South Texas Veterans’ Health System, Audie L. Murphy Division Geriatric Research, Education and Clinical Center (GRECC), and the University of Texas Health Science Center, San Antonio.

Background. The ability to predict the level of care received by elderly retirees was compared in a discriminant model and using a classification tree derived from cognitive and noncognitive variables.

Methods. Participants were 193 residents (mean age, 79.1 ± 5.1 years) of a single, 1500-bed, continuing care retirement community. They were given a battery of cognitive measures that included tests of general cognition, memory, and executive control function. A multivariate discriminant model of level of care was compared with a classification tree.

Results. Residents in congregate high-rises (n = 115) differed significantly from those in apartment settings (n = 78) with respect to age, Executive Interview (EXIT25), and the Executive Clock-Drawing Task (CLOX). Only age and executive control function measures (CLOX1, EXIT25, and Trail Making Test Part B [Trails B]) contributed independently to a discriminant model of level of care (Wilke’s λ = 0.92; F[4,170] = 3.48; p < .01). Sixty-three percent of participants were correctly classified. A classification tree derived from the same variable set was more accurate (75% correctly classified).

Conclusions. Executive control function appears to be most responsible for the effect of cognition on level of care. Young adult norms may be most relevant when the effects of cognitive impairment on functional status are assessed.

Clinical assessment of risk for institutionalization and level of care are important parts of geriatric clinical practice. Many studies have used multivariate linear statistics (e.g., multivariate analysis of covariance, multiple logistic regression, or discriminant models) to identify risk factors or independent predictors. However, the products of these approaches are difficult to apply to individual patients in routine clinical settings. Alternatively, groups of experts may develop consensus algorithms for making placement decisions. However, these are seldom directly developed from empirical data and may not be applicable to the particular patient population being examined.

Classification tree analysis (CTA) is a quasistatistical approach to clinical decision making that allows the development of empirically based decision-making tools (1). It is a “data mining” technique that predicts group membership in a categorical dependent variable (level of care) based on one or more predictor variables (2). It is often used to develop empirically based treatment or decision-making algorithms, and it is similar to discriminant function analysis because it allows an examination of the relative abilities of multiple variables to differentiate clinical groups. However, CTAs have several distinct advantages over discriminant modeling. First, they can be applied to both categorical and ordered variables. Discriminant analysis requires that the dependent variables have at least ordinal scaling. Second, CTA can explore the sequential hierarchical relationships among the predictive variables. In contrast, discriminant analyses explore simultaneous linear combinations of the predictor variables.

We have evaluated the relative ability of physical, sociodemographic, and cognitive measures to predict level of care needs among the residents of comprehensive care retirement communities (CCRCs). These are essentially closed systems in which a resident’s living setting is free to change in proportion to the services and supervision they require. Basing our investigation in this setting allows us to develop cross-sectional and longitudinal models of the variables that make the strongest independent contributions to level of care. In this article, we assess the potential of CTA (a) to facilitate the empirical development of level-of-care assessments, and (b) to demonstrate the relative importance of clinical variables to level of care decisions.

Methods

Participants

We tested 561 healthy elderly CCRC residents as part of the Air Force Villages’ Freedom House Study. The Air Force Villages is a 1500-bed, for-profit CCRC in San Antonio, Texas that is open to retired Air Force officers and their dependents. Participants represent a random sample of Air Force Villages residents older than 70 years receiving noninstitutionalized levels of care. We collected data for this analysis in the first year of the study (1994–1995). We obtained informed consent before the evaluations.

We offered a subset of Freedom House Study participants (n = 193) formal neuropsychological testing that included measures of memory, language, constructions, and executive control function (ECF). We randomly selected this subgroup from computer-generated logs of eligible Freedom
House Study participants. They were slightly older at baseline than the larger Freedom House Study cohort (mean age = 79.2 years vs 78.1 years, respectively) but did not differ significantly with regard to sex, education, baseline level of care, or Mini-Mental State Examination (MMSE) scores (3). We limited this analysis to the participants in this subset. We limited multivariate models to participants with complete data at baseline (minimum, \( n = 186 \)).

We assessed the relative contributions of various cognitive, physical, and sociodemographic variables to required level of care. The cognitive tests available to us included general measures of cognitive function, specific measures of ECF, and memory. To ensure its suitability for routine clinical settings, we included only easily administered “bedside” measures of cognition in these models.

### Cognitive Measures and Executive Measures

**CLOX: An executive clock-drawing task.**—The CLOX is a brief ECF measure based on a clock-drawing task (4). It is divided into two parts. CLOX1 is an unprompted task that is sensitive to executive control. CLOX2 is a copied version that is less dependent on executive skills. CLOX1, but not CLOX2, makes a significant independent contribution to the number of categories achieved on the Wisconsin Card Sorting Task (5). Unlike several comparable clock-drawing tasks, CLOX is a significant independent correlate of ECF (6). Each CLOX subtest is scored on a 15-point scale. Lower CLOX scores indicate impairment. Cut points of 10/15 (CLOX1) and 12/15 (CLOX2) represent the 5th percentiles for young adult controls (4).

**Controlled Oral Word Association.**—The Controlled Oral Word Association (COWA) (7) is a test of oral word production (word fluency). The patient is asked to say as many words as possible beginning with a certain letter of the alphabet. Reduced word fluency scores are associated with frontal lobe impairment, particularly in the left hemisphere (8–10).

**The Executive Interview (EXIT25).**—The EXIT25 provides a standardized clinical ECF assessment (11). It contains 25 items designed to elicit signs of frontal system disorder (e.g., imitation, intrusions, disinhibition, environmental dependency, perseveration, frontal release, and so forth). The EXIT25 requires 15 minutes to administer and can be given by nonmedical personnel. Interrater reliability is high (\( r = .90 \)). It correlates well with other measures of ECF including the Wisconsin Card Sorting Task: Categories (\( r = .54 \)), Trail Making Part B (\( r = .64 \)), Lezak’s Tinker Toy test (\( r = .57 \)), and the Test of Sustained Attention (time, \( r = .82 \); errors, \( r = .83 \)). EXIT25 scores are reported to correlate strongly and specifically with left frontal system disorders on magnetic resonance imaging (12). Scores range from 0 to 50. High scores indicate impairment. A cut point of 15/50 is recommended.

**Trail Making Test, Part B (Trails B).**—Trails B is a complex test of conceptualization, psychomotor speed, and attention (13). The patient draws lines to connect consecutively numbered and lettered circles, alternating between the two sequences. The time, in seconds, to complete the task is recorded. Performance on Trails B is specifically affected by frontal lobe lesions, and it is sensitive to neurodegenerative disorders and the effects of normal aging (14–16).

### Memory and General Cognition

**Mattis Dementia Rating Scale: Memory Subscale (DRS: MEM).**—The DRS assesses cognitive impairment in patients with suspected or known dementia (17,18). Five subtests—attention, initiation and perseveration, construction, conceptualization, and verbal and nonverbal memory—are used to derive a global measure of dementia. We used only the memory subscale in this model.

**The Mini-Mental State Examination (MMSE).**—The MMSE is a well-known and widely used test for screening cognitive impairment (3,19). Scores range from 0 to 30. Scores less than 24/30 reflect cognitive impairment. In these analyses, the MMSE is used as a proxy for posterior cortical disorder (12). It has no items that are specifically addressed to ECF (20), and it may underestimate cognitive impairment in the absence of posterior cortical disease (5,21).

### Functional Status Measures

**Level of care.**—The Air Force Villages CCRC provides services at three levels of care: independent living apartments (level 1) in which essentially no services are provided; congregate high-rises (level 2) in which laundry, housecleaning, meals, and medication supervision are provided; and skilled nursing units (level 3) in which residents are provided with assistance in their activities of daily living (ADLs), nursing care, and medications. We included only community-dwelling residents living at levels 1 or 2 in this analysis.

**Instrumental activities of daily living.**—Instrumental activities of daily living (IADLs) were assessed using the Older American Research and Service Center Instrument (22). This instrument is a multidimensional assessment tool that provides information in three domains that are likely to influence functional outcomes among CCRC residents: social resources, physical health, and IADLs. In addition, the Older American Research and Service Center Instrument services utilization section provides data on the quantity and quality of services provided to residents by their caregivers. For this model, we have abstracted IADLs, self-reported visual handicap (VISION), the receipt of personal care services (GETS HELP), and the use of prosthetic devices (including canes, hearing aids, or walkers) (PROSTHESSES) from the Older American Research and Service Center Instrument.

Residents were separated based on their level of care. Cross-group differences were tested by analysis of variance. We evaluated the relative ability of cognitive and other variables to distinguish levels of care using a forward stepwise discriminant model. The model included executive cognitive measures (e.g., CLOX1, COWA, EXIT25, and Trails B), other cognitive measures (e.g., CLOX2, DRS: MEM, and MMSE) and functional status variables (includ-
Table 1. Clinical Features of Cases, Stratified by Level of Care

<table>
<thead>
<tr>
<th>Variable</th>
<th>Apartments (N = 115)</th>
<th>High-Rises (N = 115)</th>
<th>Total (N = 193)</th>
<th>ANOVA df (1,191) p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>78.0 (±4.9)</td>
<td>79.8 (±5.1)</td>
<td>79.1 (±5.1)</td>
<td>&lt;.02</td>
</tr>
<tr>
<td>Education (y)</td>
<td>15.2 (±2.4)</td>
<td>15.0 (±2.5)</td>
<td>15.1 (±2.4)</td>
<td>NS</td>
</tr>
<tr>
<td>IADL total (MAX = 14.0)</td>
<td>13.2 (±1.4)</td>
<td>12.7 (±2.3)</td>
<td>12.9 (±2.0)</td>
<td>NS</td>
</tr>
<tr>
<td>ADL total (MAX = 14.0)</td>
<td>13.6 (±1.1)</td>
<td>13.5 (±1.3)</td>
<td>13.5 (±1.2)</td>
<td>NS</td>
</tr>
<tr>
<td>MD visits in previous 6 mo (n)</td>
<td>3.6 (±3.8)</td>
<td>4.5 (±5.0)</td>
<td>4.1 (±4.6)</td>
<td>NS</td>
</tr>
<tr>
<td>Vision (2 = good; 3 = fair)</td>
<td>2.1 (±0.7)</td>
<td>2.3 (±0.9)</td>
<td>2.2 (±0.8)</td>
<td>NS</td>
</tr>
<tr>
<td>Female (%)</td>
<td>57.7</td>
<td>60.9</td>
<td>59.6</td>
<td>NS</td>
</tr>
<tr>
<td>Living alone (%)</td>
<td>31.2</td>
<td>38.3</td>
<td>35.4</td>
<td>NS</td>
</tr>
<tr>
<td>Reporting (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glaucma</td>
<td>14.1</td>
<td>22.6</td>
<td>19.2</td>
<td>NS</td>
</tr>
<tr>
<td>Arthritis</td>
<td>59.0</td>
<td>57.4</td>
<td>58.0</td>
<td>NS</td>
</tr>
<tr>
<td>HTN</td>
<td>47.4</td>
<td>43.5</td>
<td>45.1</td>
<td>NS</td>
</tr>
<tr>
<td>CAD</td>
<td>23.1</td>
<td>25.2</td>
<td>24.4</td>
<td>NS</td>
</tr>
<tr>
<td>CVA</td>
<td>5.1</td>
<td>8.8</td>
<td>7.3</td>
<td>NS</td>
</tr>
<tr>
<td>AODM</td>
<td>5.1</td>
<td>6.1</td>
<td>5.7</td>
<td>NS</td>
</tr>
<tr>
<td>Receiving help (%)</td>
<td>5.1</td>
<td>11.4</td>
<td>8.9</td>
<td>NS</td>
</tr>
<tr>
<td>Using prostheses (%)</td>
<td>14.1</td>
<td>13.9</td>
<td>14.0</td>
<td>NS</td>
</tr>
</tbody>
</table>

Note: ANOVA = analysis of variance; ADLs = activities of daily living; AODM = adult-onset diabetes mellitus; CAD = coronary artery disease; CVA = cerebrovascular disease; IADLs = instrumental activities of daily living; HTN = hypertension; MAX = maximum; MD = physician; NS = not significant.

Table 2. Cognitive Test Scores by Level of Care

<table>
<thead>
<tr>
<th>Executive measures</th>
<th>Apartments (N = 78)</th>
<th>High-Rises (N = 115)</th>
<th>Total (N = 193)</th>
<th>ANOVA df (1,191) p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOX1</td>
<td>10.8 (±3.1)</td>
<td>9.7 (±3.3)</td>
<td>10.2 (±3.3)</td>
<td>.02</td>
</tr>
<tr>
<td>COWA</td>
<td>32.1 (±14.1)</td>
<td>29.9 (±12.6)</td>
<td>30.8 (±13.2)</td>
<td>NS</td>
</tr>
<tr>
<td>EXIT25</td>
<td>13.3 (±5.4)</td>
<td>15.3 (±5.2)</td>
<td>14.5 (±5.4)</td>
<td>.01</td>
</tr>
<tr>
<td>Trails B (s)</td>
<td>131.7 (±102.7)</td>
<td>128.5 (±71.2)</td>
<td>129.8 (±85.2)</td>
<td>NS</td>
</tr>
</tbody>
</table>

General measures

| CLOX2                   | 13.5 (±1.7)         | 13.0 (±2.3)         | 13.2 (±2.1)     | NS                |
| DRS/MEM                 | 38.1 (±10.6)        | 38.3 (±7.5)         | 38.2 (±8.9)     | NS                |
| MMSE                    | 27.7 (±2.1)         | 27.3 (±2.5)         | 27.5 (±2.4)     | NS                |

Note: ANOVA = analysis of variance; CLOX1, CLOX2 = executive clock-drawing tasks; COWA = Controlled Oral Word Association; EXIT25 = executive interview; DRS/MEM = Mattis Dementia Rating Scale: Memory Subscale; MMSE = Mini-Mental State Examination; NS = not significant.

predicted to need services in the high-rise setting; and 21% of high-rise residents were predicted to be living in the apartments.

Figure 1 presents the relative importance of each predictor variable to the resulting CTA. Age made the most important contribution to the model. Cognitive measures appeared to be more relevant to level of care than were sociodemographic characteristics and functional status. The executive measures contributed more than did tests of memory or general cognitive function. Yet even among the executive measures, differences existed with regard to the contributions made to level of care. As in the discriminant model, Trails B and COWA contributed less to this determination than either CLOX1 or the EXIT25. CLOX1 contributed nearly as much as age itself.

Figure 2 presents the final CTA, which might also serve as an algorithm for the determination of level of care in this sample. Each decision point is represented by a box. The box contains the critical threshold for the predictor variable chosen by the CTA. Over this box is the number of participants still available for discrimination. A positive response takes the investigator to the left. Eventually, terminal nodes (circles) are reached. Each gives a classification for the participant.

Figure 1. A classification tree for determining level of care in a high-rise setting.

Discriminant Analysis

The discriminant model of level of care was significant [Wilke’s λ = 0.92; F(df(4,170)) = 3.48; p < .01]. However, only age and executive measures (EXIT25, CLOX1, and Trails B) entered. Age (p < .01), EXIT25 (p = .01), and CLOX1 (p = .05) made significant independent contributions. Functional status variables, including marital status, living alone, the receipt of assistance with personal care, the use of prosthetic devices, self-reported visual handicap, IADLs, and ADLs, did not contribute significantly to level of care independently of age and ECF. Resubstitution correctly classified 63% of participants.

Classification Tree Analysis

The CTA was more accurate: resubstitution correctly classified 75% of participants; 24% of apartment residents were

ing age, marital status, living alone, the receipt of assistance with personal care, the use of prosthetic devices, self-reported visual handicap, IADLs, and ADLs). Finally, we constructed a classification tree from the same set of variables. We performed CTAs using STATISTICA 5.5 (Tulsa, OK).

RESULTS

Table 1 presents the mean demographic characteristics for the total sample (n = 193) and both levels of care. Residents in congregate high-rises (n = 115) differed significantly from those in apartment settings (n = 78) only with respect to age (by analysis of variance: df [1,191], p < .02).

Table 2 presents the cognitive test scores by level of care. Residents in congregate high-rises differed significantly from those in apartment settings only with respect to executive measures (EXIT25 and CLOX1); by analysis of variance: df [1,191], both p < .02). This is consistent with our findings in the larger Freedom House Study sample (23).

Discriminant Analysis

The discriminant model of level of care was significant [Wilke’s λ = 0.92; F(df(4,170)) = 3.48; p < .01]. However, only age and executive measures (EXIT25, CLOX1, and Trails B) entered. Age (p < .01), EXIT25 (p = .01), and CLOX1 (p = .05) made significant independent contributions. Functional status variables, including marital status, living alone, the receipt of assistance with personal care, the use of prosthetic devices, self-reported visual handicap, IADLs, and ADLs, did not contribute significantly to level of care independently of age and ECF. Resubstitution correctly classified 63% of participants.

Classification Tree Analysis

The CTA was more accurate: resubstitution correctly classified 75% of participants; 24% of apartment residents were

predicted to need services in the high-rise setting; and 21% of high-rise residents were predicted to be living in the apartments.

Discussion

This study suggests that CTAs may offer a convenient approach to the development of empirically based level of care decision-making algorithms. However, it is not our intention to offer this particular model for clinical applications. It is a cross-sectional model derived from a single CCRC and may not be applicable outside of that setting. A 75% successful classification rate may not be high enough for some clinical purposes, and higher rates might have been possible had the algorithm been offered different variables. Furthermore, the model can only re-
construct how the participants were actually distributed across levels of care at the time of their assessment. This does not imply that they were ideally placed. In fact, some of the “misclassifications” may reflect true inefficiencies in the participants’ baseline distribution within this CCRC. We have shown that longitudinal changes in cognition are particularly strongly associated with changes in functional status (23). Classification trees derived from longitudinal data might prove to be more accurate in predicting future caregiving needs. These results should be considered provisional until they can be replicated.

Nevertheless, the current algorithm classified participants more accurately than did a linear discriminant model derived from the same data set. Furthermore, by allowing para-professionals to collect and interpret the data, this algorithm might be more easily applied to individual patients. Thirty-nine participants (20.2%) could be classified by the CTA based on only two examinations (i.e., the EXIT25 and vision screening), whereas no participant would be exposed to more than three examinations. In contrast, all participants would have needed three examinations (i.e., EXIT25, CLOX, and Trails B) to allow us to run the discriminant model. Furthermore, test data would have to be entered into a calculator or computer to run against the discriminant function before individual residents could be classified. Finally, a discriminant model’s output would be in the form of classification probabilities, which would require further clinical interpretation. Thus, discriminant modeling is inherently less well suited than CTA for “in the field” clinical determinations.

These data also provide new insights into the variables that best account for the cross-sectional variance in level of care. Age and ECF measures emerged as the major determinants in both the discriminant model and our CTA. Furthermore, all three of the executive measures selected by the algorithm (i.e., CLOX1, EXIT25, and Trails B) are nonverbal. Executive functions control the sequencing and execution of complex goal-directed activities. These include cooking, dressing, shopping, and doing housework. It should not be surprising then that ECF impairment is emerging as a robust determinant of functional status and disability. CLOX1 and the EXIT25 have previously been shown to be strong independent correlates of cross-sectional IADL performance and level of care and longitudinal predictors of change in functional status (11,23–26).

We recently performed factor analysis on the ECF measures in our data set (27) and associated them with cross-sectional measures of functional status. Three factors emerged, but only one, “procedural control,” which represented 28% of the variance in our battery, was associated with functional status. Both CLOX (r = .53) and the EXIT25 (r = -.70) loaded most strongly on this factor. In contrast, 24.2% of variance in our ECF measures was explained by an “abstraction” factor that was dominated by subtests from the Wisconsin Card Sorting Task (WCST) (28). This factor was not associated with functional status. Similarly, we recently reported that longitudinal change in the WCST performance is not related to change in IADL independently of change in the EXIT25 (29). The current algorithm empirically confirms a specific relationship between level of care and a nonverbal subset of executive measures.

Finally, we would like to draw attention to the EXIT25 and CLOX1 thresholds that were empirically selected by CTA to predict level of care. An EXIT25 score of 10/50 and a CLOX1 score of 10/15 both coincide with the fifth percentiles for these instruments in young adult samples (4). This suggests that the performance of seniors may need to be judged in comparison to young adult norms if level of care is to be accurately modeled. In contrast, most dementia case finding is currently predicated on age-adjusted norms. Both mild cognitive impairment (30) and dementia defined by the National Institute for
Neurological, Communicative Disorders and Stroke (31) are explicitly defined in relation to age-adjusted norms. However, evidence-based reviews have not substantiated either dementia case definitions or dementia screening processes that are predicated on this approach (32,33).

Our data suggest that much lower thresholds for cognitive impairment are more accurate in predicting functional outcomes, including level of care. For example, an EXIT25 score of 15/50 (its recommended cut point) perfectly distinguishes elderly patients who can and cannot be taught to competently manage respiratory inhalers (34). The same threshold best distinguishes elderly retirees who can give informed consent for inpatient psychiatric treatment (35), adhere to medical recommendations (36), or accurately assess their financial decision-making skills relative to performance-based assessments (37). In the latter study, a CLOX score of 10/15 (both its recommended cut point and the fifth percentile for young adults) best performed that discrimination. This is consistent with recent reports of a significant independent effect of CLOX1 failure (relative to young adult norms) on simulated driving performance (personal communication with Barbara Freund, MD, Eastern Virginia Medical School, April 6, 2001) and longitudinal institutionalization risk among elderly retirees (38).

The significance of this is that the age-specific prevalence of ECF impairment will greatly increase if defined in relation to young adult norms rather than age-adjusted norms. In the Air Force Villages, 38% of septuagenarians display EXIT25 and CLOX1 impairment at their respective cut points (23). Community samples show even higher prevalence rates. A recent San Luis Valley Health and Aging Study (39) showed that 33% of 1298 community-dwelling persons older than 60 years (our sample frame is older) had ECF impairment. As in the Freedom House Study, ECF was a significant independent predictor of IADL in age-adjusted and MMSE-adjusted models.

In the study by Grigsby and colleagues (39), 45.4% of (n = 746) elderly Hispanic persons were ECF impaired. Similarly, we observed a 56% prevalence of CLOX1 failure among 1202 community-dwelling Hispanic persons in the Hispanic Established Population for Epidemiological Studies in the Elderly (40). Thus, the community prevalence of ECF impairment is quite high if compared with young adult norms, but that practice may be justified by the empiric validity of such impairment to predict functional status, decision-making, medical adherence, self-medication capacity, and, in the current study, level of care.

A CTA can provide empirically based decision-making algorithms based on easily acquired clinical variables. In this population, a decision tree derived from physical,
sociodemographic, and bedside cognitive measures produced a more accurate model of cross-sectional level of care than did a conventional discriminant model. Regardless of the method used, both models show that cognitive function contributes significantly to the level of care received by elderly retirees, independently of age and other variables. Of the cognitive domains we tested, ECF appears to be most responsible for the effect of cognition on level of care. This model also provides empirical support for the use of young adult norms when the effects of cognitive impairment on functional status are assessed.

ACKNOWLEDGMENTS

The authors thank the Air Force Villages for their valuable cooperation and support.

Requests for reprints and detailed information regarding the EXIT25 or the CLOX should be addressed to Dr. Donald Royall, Department of Psychiatry, The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78248-7792.

Address correspondence to Marsha J. Polk, MMEd, Department of Psychiatry, MC 7792, UTHSCC, 7703 Floyd Curl Drive, San Antonio, TX 78229-3900. E-mail: polk@uthscsa.edu

REFERENCES


Received March 19, 2004
Accepted May 7, 2004

Decision Editor: John E. Morley, MB, BCh