3 Editor's Note

ORIGINAL ARTICLES

4 Association of Depression With Agitation in Elderly Nursing Home Residents
Oscar Heeren, MD, Luda Borin, MD, Allen Raskin, PhD, Ann L. Gruber-Baldini, PhD, A. Srikumar Menon, MD, PhD, Bruce Kaup, MD, David Loreck, MD, Paul E. Ruskin, MD, Sheryl Zimmerman, PhD, and Jay Magaziner, PhD

8 Correlates of Behavioral Disturbances and Pattern of Psychotropic Medication Use in Five Skilled Nursing Facilities
Fadi H. Ramadan, MD, Bruce J. Naughton, MD, and Roger Prior, ScD

15 Health Service Utilization by Alzheimer's Disease Patients: A 2-Year Follow-up of Primary Versus Subspecialty Care
Peter M. Aupperle, MD, MPH, Edward R. MacPhee, BA, Andrew C. Coyne, PhD, Jonathan Blume, BS, and Betty Sanchez, MS

18 Smoking and Cognitive Performance in the Community Elderly: A Longitudinal Study
Wei-Ta Chen, MD, Pei-Ning Wang, MD, Shu-Jue Wang, MD, Jong-Ling Fuh, MD, Ker-Neng Lin, PhD, and Hsiu-Chih Liu, MD

23 Construct Validity of the 15-Item Geriatric Depression Scale in Older Medical Inpatients
R. Antonelli Incalzi, MD, M. Cesari, C. Pedone, and P. U. Carbonin

29 Treatment of Delirium in Older Adults With Quetiapine
Kye Y. Kim, MD, Geoffrey M. Bader, MD, Victor Kotlyar, MD, and Debra Gropper, PharmD

32 Recorded Delirium in a National Sample of Elderly Inpatients: Potential Implications for Recognition
Helen C. Kales, MD, Barbara A. Kamholz, MD, Stephanie G. Visnic, BA, and Frederic C. Blow, PhD

39 Long-Term Effects of Donepezil on P300 Auditory Event-Related Potentials in Patients With Alzheimer's Disease
Eiichi Katada, MD, PhD, Koichi Sato, MD, PhD, Akira Sawaki, MD, Yasuaki Dohi, MD, PhD, Ryuzo Ueda, MD, PhD, and Kohei Ojika, MD, PhD

44 Age-Related Impairments in Comprehending Affective Prosody With Comparison to Brain-Damaged Subjects
Diana M. Orbelo, PhD, Julie A. Testa, PhD, and Elliott D. Ross, MD

53 Effects of Cognitive Challenge on Gait Variability in Patients With Parkinson's Disease
Jeffrey M. Hausdorff, PhD, Jackov Balash, MD, PhD, and Nir Giladi, MD

59 Familial Dementia With Lewy Bodies With an Atypical Clinical Presentation
Lauren T. Bonner, MD, Debby W. Tsuang, MD, MSc, Monique M. Cherrier, PhD, Charisma J. Eugenio, BS, Jennifer Q. Du, RN, Ellen J. Steinbart, RN, Pornprot Limprasert, MD, PhD, Albert R. La Spada, MD, PhD, Benjamin Seltzer, MD, Thomas D. Bird, MD, and James B. Leverenz, MD
This issue of the Journal marks another transition in its journey. We bid adieu to B.C. Decker, Inc., which has published this journal for nearly a decade, and welcome our new publisher, Sage Publications, which has recently acquired JGPN. I want to personally express my gratitude to the staff at Decker for their support during the past 5 years I have worked with them. It has been gratifying to see the quality and impact of the Journal increase over this time, and their stewardship has been critical to this success. I am truly excited to welcome our new publisher—Sage is a distinguished publishing house which has already shown its commitment to helping further the mission of the Journal. I look forward to working with them to provide the very best to our readership.

Alan M. Mellow, MD, PhD
Editor-in-Chief
Association of Depression With Agitation in Elderly Nursing Home Residents

Oscar Heeren, MD, Luda Borin, MD, Allen Raskin, PhD, Ann L. Gruber-Baldini, PhD, A. Srikumar Menon, MD, PhD, Bruce Kaup, MD, David Loreck, MD, Paul E. Ruskin, MD, Sheryl Zimmerman, PhD, and Jay Magaziner, PhD

ABSTRACT

Agitation is a serious problem for elderly individuals with dementia. It is often the major reason for admission to a restrictive environment such as a nursing home or hospital. The objectives of the current study were to (1) identify the components of agitation embedded in the Psychogeriatric Dependency Rating Scale (PGDRS) and (2) find correlates of these factors in demographic and clinical variables. The study sample comprised 2285 subjects who were admitted to 59 nursing homes across Maryland. The factor analysis of the PGDRS confirmed that agitation is made up of a number of different elements ranging from physical and/or verbal aggression to wandering. Correlates of these elements varied, as did possible treatments. For example, physical and/or verbal aggression often accompanied severe depression, suggesting that treating the depression may alleviate this problem. However, wandering and psychotic behavior may be less amenable to existing treatments as these behaviors were associated with severe cognitive impairment. (J Geriatr Psychiatry Neurol 2003; 16:4–7).
from September 1992 through March 1995, aged 65 years and older, who had not been a resident in a nursing home in the previous year were eligible for the study. Informed consent was obtained from either the resident or a significant other. The Institutional Review Board of the University of Maryland approved the consent procedure. There were a total of 3283 eligible subjects, of whom 2285 were enrolled in the study. There were 998 who did not participate. The major reason for not participating was because the patients or their significant others refused consent. A more detailed account of subject selection and data collection procedures may be found elsewhere.15

Procedure
Data were collected from multiple sources, including interviews with residents, nursing staff, and significant others; medical records; and hospital discharge summaries. The current article reports on a secondary, post hoc analysis, using data from a National Institute on Aging–sponsored study conducted by the Department of Epidemiology and Preventive Medicine at the University of Maryland School of Medicine. The latter, larger study was an epidemiologic study to determine the prevalence of dementia in persons admitted to nursing homes in Maryland.15

Measures
The choice of instruments selected for this study was based on two factors. First, as noted above, this was a secondary analysis, and we were limited to the assessment measures included in the larger study. Second, however, we did have a wide range of instruments and demographic variables from which to choose. The instruments and variables selected were those that had shown promise in prior studies as predictors of agitated behavior.

Cognitive Deficit
Cognitive deficits were assessed using the Mini-Mental State Examination (MMSE).16 The MMSE was administered to every patient who did not refuse, who could communicate, and for whom testing could be done within 2 months of admission (n = 1446; 63% of the sample). Residents were classified by severity of cognitive impairment based on MMSE scores: 1 = severe (< 10), 2 = moderate (10–19), 3 = mild (20–26), and 4 = absent (27–30). This is a widely used screen for dementia that has shown high inter-rater reliability and validity.16

Depression
Depressive symptoms were ascertained with a modified version of the Cornell Scale for Depression in Dementia (CSDD),17 administered to a nurse informant. The CSDD is an observer-rated scale designed to rate depression in residents with dementia. It contains 19 items, and total scores range from 0 to 38. Based on their scale scores, residents were categorized as follows: 0 = no depression (scores 1–7), 1 = mild to moderate depression (scores 8–20), and 3 = severe depression (scores 21+). These categorizations follow the recommendations of Alexopoulos and associates.17 This scale has been validated as a depression assessment instrument in both demented and nondemented persons.17

Agitation
Behavioral functioning was assessed using the behavior subscale of the PGDRS, which assessed the presence of aggression, passive hostility, attention seeking, wandering, restlessness, and fearfulness. The PGDRS behavior subscale included 17 items rated on a frequency of occurrence scale as 0 = none, 1 = one to two times a week, and 2 = more often. The original reference to this scale cited evidence of high inter-rater reliability.14

Data Analysis
Prior factor-analytic studies of scales assessing agitation in elderly nursing home residents have identified at least three independent factors.6,7,8 Based on these results, a factor analysis of the PGDRS behavior subscale was undertaken. To obtain relatively independent factors, a normal varimax rotation, which provides an orthogonal solution, was performed on all factors with eigenvalues of 1 or greater. Five factors met this criterion and were extracted. We also looked at the four-, three-, and two-factor solutions but decided that the five-factor solution best described the components imbedded in the PGDRS behavior subscale. This was a judgment on our part, but subsequent findings seemed to justify this decision. An exact factor scoring method was used that includes all items in a factor, but the items are assigned weights based on their loading on the factor. This approach maintains the orthogonality among factors. Separate analyses of variance were computed using each of the five factors as the dependent variable and race, gender, CSDD score, and MMSE score as the independent variables. Range tests (Scheffé) were also performed to assess where the significant differences were when the independent variable contained three or more groups. An analysis of variance was also performed using the total score of the PGDRS behavior subscale as the dependent variable and the same variables noted above as the independent variables.

RESULTS
The mean age of the study population was 81.45 years, with a median age of 81. Female residents accounted for 70.9% of the sample. In terms of race, 80.4% were white and 19.4% were nonwhite. The majority were widowed (62.2%); 23.9% were still married. Fifty-one percent had at least high school education. Based on the 4-point scale developed for the MMSE, 19.8% of the sample had severe cognitive impairment, 34.3% had moderate cognitive impairment, 27.3% had mild cognitive impairment, and 18.6% had no cognitive impairment. Based on the CSDD, 21.7% of the
study population was classified with either mild or moderate depression. Only 0.09% had severe depression.

As noted, five factors were extracted from the PGDRS behavior subscale. The factor loadings of the scale items are listed in Table 1. Based on items with the highest loading for each factor (0.40 or higher), factors were assigned the following names: 1 = physical aggression/verbal abuse (5 items), 2 = wandering (3 items), 3 = intrusiveness (2 items), 4 = psychotic behavior (2 items), and 5 = manic/destructiveness (2 items). Factor 1, physical aggression/verbal abuse, describes hitting, biting and scratching, that is, physical abuse against others. Factor 5, manic/destructiveness, describes behavior inflicted on one’s belongings, that is, destroying and damaging clothes and belongings. Factor 5 differs from factor 4, psychotic behavior, in that the latter includes visual and auditory hallucinations and incoherent speech, behaviors that are frankly psychotic.

The F-ratios for the PGDRS behavior subscale total score and for factors 1 to 4 were significant at the .01 level (Table 2). Cognitive impairment significantly predicted outcome for the PGDRS total score and for factor scores 1, 2, and 4 (see Table 2). Depression measured by the CSDD significantly predicted the PGDRS total score and factor scores 1, 3, and 4 (see Table 2). Gender significantly predicted factors scores 1 and 2 (see Table 2). Race significantly predicted outcome only for factors score 1 (see Table 2).

There were also eight significant interactions among the independent variables (see Table 2). For the PGDRS total score, the only significant interaction was between cognitive deficit and depression. For factor 1, the significant interactions were between cognitive deficit and depression, between cognitive deficit and gender, between depression and gender, between depression and race, and between gender and race. For factor 3, the only significant interaction was between depression and race. There were no significant interactions for factor 2, 4, or 5.

DISCUSSION

The factor analysis of the PGDRS confirmed that agitation is a construct made up of a number of different elements ranging from physical and/or verbal aggression to wandering. This finding is consistent with the results of other factor analyses, including the work of Bogner and associates,† who view agitation as a general or unitary construct with three underlying factors. This is an important point to keep in mind because the predisposing factors and treatment of these elements or underlying factors can be quite different. For example, we found that different variables predicted physical and/or verbal aggression (factor 1), wandering (factor 2), and intrusiveness (factor 3).

Moderate cognitive deficits and severe depression were associated with physical aggression. In particular, high levels of physical aggression were present in severely depressed white females and males with moderate cognitive impairment. In this context, bear in mind that whites comprised 80% of the sample and women 70%. Wandering was a problem in the severely cognitively impaired and more so in males than females. Attention seeking occurred most often in residents with severe depression, especially white residents with moderate cognitive deficits. Finally, as one might expect, psychotic behaviors were seen in residents with severe cognitive impairment. These findings are summarized in Table 3.

The results indicate that physical and/or verbal aggression often accompanies severe depression in nursing home residents, suggesting that treating depression may alleviate this problem. Similarly, treating depression in attention-seeking or intrusive patients may also alleviate this problem. However, wandering and psychotic behavior may be less amenable to treatment as these behaviors were associated with severe cognitive impairment, for which there is currently no effective treatment. As noted above, the sample studied was primarily white and female. This may explain, in part, why many of the results seem targeted to this cohort of nursing home residents.

These results are consistent with prior findings that identified cognitive deficits and depression as precursors

---

Table 1. Item Loading on the Five-Factor Solution of the Behavior Subscale of the Psychogeriatric Dependency Rating Scale

<table>
<thead>
<tr>
<th>Item Content</th>
<th>Factor Loading*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Interferes with and disturbs activities of other residents or staff</td>
<td>.51</td>
</tr>
<tr>
<td>Gets you to do things for him/her that he/she can do</td>
<td>.13</td>
</tr>
<tr>
<td>Wanders around the floor/nursing home</td>
<td>.13</td>
</tr>
<tr>
<td>Shows behavior that are objectionable (spit, undress, urinate)</td>
<td>.32</td>
</tr>
<tr>
<td>Demands conversation</td>
<td>.17</td>
</tr>
<tr>
<td>Has difficulty understanding what you want</td>
<td>.12</td>
</tr>
<tr>
<td>Makes noises that are objectionable</td>
<td>.48</td>
</tr>
<tr>
<td>Hits, bites, scratches</td>
<td>.75</td>
</tr>
<tr>
<td>Deliberately refuses to obey commands</td>
<td>.66</td>
</tr>
<tr>
<td>Verbal abuse and threats</td>
<td>.80</td>
</tr>
<tr>
<td>Restlessly paces</td>
<td>.12</td>
</tr>
<tr>
<td>Self-destructive acts</td>
<td>.29</td>
</tr>
<tr>
<td>Destroys and damages clothes and belongings</td>
<td>.16</td>
</tr>
<tr>
<td>Inappropriate elation</td>
<td>.14</td>
</tr>
<tr>
<td>Thinks people are trying to harm him/her</td>
<td>.61</td>
</tr>
<tr>
<td>Sees and hears things that are not there</td>
<td>.11</td>
</tr>
<tr>
<td>Speech incoherent</td>
<td>.13</td>
</tr>
</tbody>
</table>

*Factor names: 1 = physical aggression/verbal abuse ; 2 = wandering; 3 = intrusiveness; 4 = psychotic behavior; 5 = manic/destructiveness.

Numbers in bold indicate items that were used in that factor.
and correlates of agitation. The present study extended these findings by pinpointing the levels of cognitive deficits and depression associated with various components of agitation, such as physical aggression. This information is valuable for identifying specific underlying psychiatric problems that may be amenable to treatment as a way of managing specific aspects of agitated behavior.

Acknowledgments
We would like to acknowledge the cooperation of the facilities, residents, and families participating in the University of Maryland Long-Term Care Project.

References

and correlates of agitation. The present study extended these findings by pinpointing the levels of cognitive deficits and depression associated with various components of agitation, such as physical aggression. This information is valuable for identifying specific underlying psychiatric problems that may be amenable to treatment as a way of managing specific aspects of agitated behavior.

Acknowledgments
We would like to acknowledge the cooperation of the facilities, residents, and families participating in the University of Maryland Long-Term Care Project.

References

and correlates of agitation. The present study extended these findings by pinpointing the levels of cognitive deficits and depression associated with various components of agitation, such as physical aggression. This information is valuable for identifying specific underlying psychiatric problems that may be amenable to treatment as a way of managing specific aspects of agitated behavior.

Acknowledgments
We would like to acknowledge the cooperation of the facilities, residents, and families participating in the University of Maryland Long-Term Care Project.

References

and correlates of agitation. The present study extended these findings by pinpointing the levels of cognitive deficits and depression associated with various components of agitation, such as physical aggression. This information is valuable for identifying specific underlying psychiatric problems that may be amenable to treatment as a way of managing specific aspects of agitated behavior.

Acknowledgments
We would like to acknowledge the cooperation of the facilities, residents, and families participating in the University of Maryland Long-Term Care Project.

References

and correlates of agitation. The present study extended these findings by pinpointing the levels of cognitive deficits and depression associated with various components of agitation, such as physical aggression. This information is valuable for identifying specific underlying psychiatric problems that may be amenable to treatment as a way of managing specific aspects of agitated behavior.

Acknowledgments
We would like to acknowledge the cooperation of the facilities, residents, and families participating in the University of Maryland Long-Term Care Project.

References
Correlates of Behavioral Disturbances and Pattern of Psychotropic Medication Use in Five Skilled Nursing Facilities

Fadi H. Ramadan, MD, Bruce J. Naughton, MD, and Roger Prior, ScD

ABSTRACT

There are several treatment options for behavioral disturbances (BDs) in dementia. However, the choice of a specific psychotropic agent is directed by personal preferences and local community practice patterns. We examined the relationship between common clusters of BDs and the use of different classes of psychotropic agents in our community. A cross-sectional study of 430 long-term care residents from 5 nursing homes was undertaken. The Behavior Measurement Scale (BMS) was used to measure the frequency of BDs grouped in 4 categories. Residents with > 4 BD episodes in at least one category during a 2-week observation period were the behavior group and were considered to have clinically significant BDs. A sample of patients who had < 4 BDs in all BMS categories during the same observation period defined the nonbehavior group. A BD cluster was defined as > 4 BDs occurring in one or more BMS categories during the 2-week observation. Data on functional status, comorbidity, use of benzodiazepines, antidepressants, and neuroleptic agents were collected with chart review. The chi-square test was used to examine the correlation between variables. Clinically significant BDs were identified in 27.2% (117/430) of the residents in the sample. Five of 15 behavior clusters accounted for 73% of all clinically significant BDs. The 5 clusters were verbally nonaggressive behaviors (cluster 1, 20.5%), behaviors from all 4 categories (cluster 2, 17.9%), verbally and physically nonaggressive behaviors (cluster 3, 14.5%), physically nonaggressive behaviors (cluster 4, 12.8%), and verbally aggressive and nonaggressive behaviors (cluster 5, 7.7%). Cluster 5 had a negative correlation with functional impairment ($P = .009$). There was a significant correlation between cluster 2 and benzodiazepine use ($P = .014$). No other significant correlation was found between any of the 5 clusters and demographic variables, comorbidity status, and use of antidepressant or neuroleptic medications. Residents in the behavior group had higher impairment in self-feeding ($P = .036$) and bathing ($P < .001$) and were more likely to be treated with benzodiazepines ($P = .004$) and neuroleptic agents ($P = .009$) than residents in the nonbehavior group ($n = 116$). The higher use of neuroleptics and benzodiazepines in the behavior group compared with the nonbehavior group indicates that BDs are being identified for treatment, but the medications used may not be efficacious. The lack of association between specific classes of psychotropic medications and distinct behavior clusters indicates that clinicians are not using a standardized approach to target the neurochemical abnormalities that may underlie certain behavior clusters. Some behavior clusters correlate with impairment in specific activities of daily living categories such as bathing and feeding, making room for nonpharmacologic interventions. (*J Geriatr Psychiatry Neurol* 2003; 16:8–14).
The necessity for managing patients with significant behavioral problems in nursing homes is increasing, owing in part to the growth of the “very old” or “frail elderly”. With advanced age, physical and mental health problems increase significantly. Both pharmacologic and non-pharmacologic interventions have been effective for certain behavioral disturbances (BDs); however, consensus on specific treatment strategies remains elusive. Even for instances in which there is consensus, treatment often varies with local community practices and clinician biases.

Research on agitated behaviors in dementia is characterized by inconsistent and imprecise definition of BD, unavailability of reliable and uniform clinically applicable measurement tools, and absence of consensus guidelines for treatment. The 7th International Psychogeriatric Association Congress (1995) issued a number of recommendations for research priorities. These recommendations emphasized the need for determining the incidence and characteristics of BDs and developing treatment guidelines based on etiology. Longitudinal evaluation of therapeutic responses of BDs was identified as another critical area for research. Reliable and accurate measurement of agitated behaviors is crucial for tracking illness progression and monitoring the effects of treatment.

Despite improved understanding of BDs and their underlying etiologies, traditional neuroleptic medications are often used to treat BDs. However, the efficacy of neuroleptic medications may be limited, and adverse effects are common. Although other agents, such as selective serotonin reuptake inhibitors (SSRIs), valproic acid, and benzodiazepines, are also being used for BDs, it is not clear how clinicians select a specific agent for a particular BD.

It has been previously shown that psychiatric symptoms and BD may co-occur in dementia. However, treatment of these psychiatric symptoms may have either no effect on BD or a beneficial effect that is independent of the medication’s effect on psychosis. A recent study by Sultzer et al found that treatment of delusions in concert with BDs did not predict greater behavioral improvement when compared with treatment of subjects without signs of psychosis. On the other hand, treatment of depressive symptoms with trazodone resulted in greater behavioral improvement. These findings support the distinction between psychoses and BDs and can be used to improve the prescribing patterns of psychotropic medications, specifically a more targeted approach to the use of neuroleptic agents.

The objective of our study was to examine the pattern of psychotropic medication use in five skilled nursing facilities (SNFs) in relation to the five most frequently encountered clusters of BD during a 2-week sampling period. Our goal was to determine whether pharmacologic interventions were consistent with current recommendations and to identify potential areas for improvement. An example would be to determine the extent to which SSRIs are used for verbal agitation, a behavior that is often associated with depressive symptoms and reduced central serotonin levels.

**METHODS**

**Participants**

Four hundred and thirty-one residents from five skilled nursing care units in five SNFs were included in the study. The physician staffing, medical directorship, and hospital affiliation differed among these SNFs. In addition, the study SNFs represented different populations that included inner city and suburban communities. Because of the high prevalence of dementia in SNFs and the frequent association of psychiatric symptoms and BDs with cognitive impairment, only residents who had the diagnosis of dementia and resided on one of the long-term care units during the data collection period were eligible for the study. We excluded residents who were comatose or in a vegetative state and residents from subacute rehabilitation units.

Residents with > 4 BD episodes in at least one Behavior Measurement Scale (BMS) category during a 2-week observation period were defined as the behavior group and considered to have clinically significant BDs. A random sample (n = 116) of residents who did not meet the definition for clinically significant BDs represented the comparison or nonbehavior group.

**Measures**

Demographic data were collected on the study sample. The BMS was used for behavior measurement. The BMS measures the frequency of 33 behaviors grouped in four categories: (1) verbally aggressive, (2) verbally nonaggressive, (3) physically aggressive, and (4) physically nonaggressive. The BMS, derived from the Cohen-Mansfield Agitation Inventory (CMAI), allows the rating of the same behaviors included in the CMAI, daily, and per nursing shift. The BMS represents a prospective version of the CMAI.

Although there are other scales available for behavior measurement, such as the Neuropsychiatric Inventory (NPI) and the Manchester and Oxford Universities Scale for the Psychopathological Assessment of Dementia (MOUSEPAD), most of them include psychiatric components, such as delusions and paranoia. In addition, most of these scales are time consuming and some require the skills of research assistants for completion. On the other hand, the BMS is a simple scale that rates BDs only, and it can be completed by any member of the nursing team.

The 33 CMAI behavior items on the BMS are shown in the Appendix. In each category, behaviors were rated either present (occurred at least once during a nursing shift) or absent (never occurred) at the end of nursing shifts. Behavior was recorded when the corresponding category (from 1 to 4) was
circled on the BMS chart. The inter-rater reliability and validity of the BMS have been previously reported.

The Katz Index of Activities of Daily Living was used to measure functional status. The research coordinators completed this measurement in collaboration with the head nurse on each nursing unit. A score of > 10 on this index defined significant functional impairment. The level of comorbidity was determined based on chart review. Significant comorbidity was defined as having two or more medical illnesses, such as congestive heart failure, chronic obstructive lung disease, and cancer. Data on the use of antidepressants, neuroleptic agents, and benzodiazepines were collected.

The diagnosis of dementia was obtained from the SNF medical records. We did not collect data on neuroimaging or mental state examination results since studies have found only modest differences in the prevalence of psychiatric symptoms and BDs in different types of dementia (Alzheimer’s disease, vascular, etc.) and at different stages of the illness.

Procedures
Institutional Review Board approval for the use of the BMS was obtained. Prior to data collection, in-service training sessions for all members of the nursing staff were conducted in each of the five nursing homes. The training sessions consisted of a 1-hour group presentation in each nursing facility on the purpose of the study, the value of consistent and complete data, and appropriate use of the BMS. It was emphasized that certified nurse assistants (CNAs) should perform the ratings.

Fifteen residents from each unit were rated during each 2-week observation period. The sequence was repeated until all residents satisfying the inclusion criteria from all units were rated. The research coordinators reviewed the data collection sheets from each unit twice a week and reported any problems to the head nurse. Residents with missing ratings (no items circled on the BMS on a nursing shift) were observed for an extended period (beyond 2 weeks) to complete the 2-week data.

The BMS sheets were collected at the end of the study period. All participants had 14 observations completed. A 2-week score in each behavior category was obtained for each nursing home resident by adding the daily positive ratings. A score of 4 or more per the 2-week study period per category (corresponding to 2 behavior episodes or more per category per week) was considered to be clinically significant. This cutoff score is based on the Manchester and Oxford Universities Scale for the Psychopathological Assessment of Dementia (MOUSEPAD), which sets the frequency of moderate symptoms at one or more episodes per week.

A data collection sheet was created for the demographic variables, functional status, comorbidity status, and medication use. The sheet included information on age, sex, total Katz index scores, scores on individual activities of daily living listed on the Katz index, and number of comorbid illnesses. Data on medication use were reported using the number 1 (medication class used) or zero (medication class not used). The same data were collected for the nonbehavior group.

Analyses
The whole study sample was used to calculate the frequencies of the 4 behavior categories and examine their distribution in relation to morning and evening nursing shifts. We used the cutoff score of > 4 episodes/2-week period/category/resident, to identify behavior clusters defined as BDs occurring in one or more categories. Correlation between the 5 most commonly encountered behavior clusters and demographic variables, functional status, comorbidity status, and medication use was examined. Data on demographic variables, clinical characteristics, frequency of behaviors, and medication use were entered into a desktop computer and analyzed using the SPSS 9.0 program. Functional status and comorbidity scores were dichotomized to > 10 or < 10, and > 2 or < 2, respectively. The chi-square test (Fisher exact test) was used to study correlation between variables in the group with clinically significant BDs. All possible combinations of clusters were tested for any significant differences in demographic variables. The chi-square test was also used to examine the differences in clinical characteristics or medication use between the behavior and nonbehavior groups.

RESULTS
Clinically significant BDs were observed for 27% (117/431) of the study sample. Of those with significant BDs, 58% (68/117) had BDs in two or more categories. A comparison between the behavior and the nonbehavior groups on demographic data, Katz scores, and comorbidity level is shown Table 1. The only statistically significant differences found were higher impairment in self-feeding (P = .036) and bathing (P < .001) for the behavior group. The differences between behavior frequencies on the morning versus the evening shifts were not statistically significant. There were 15 behavior clusters identified in the 117 patients who had clinically significant BDs (Table 2). Five clusters comprised 73% of the behaviors. These were verbally nonaggressive behaviors (cluster 1, 20.5%), behaviors from all 4 categories (cluster 2, 17.9%), verbally and physically nonaggressive behaviors (cluster 3, 14.5%), physically nonaggressive behaviors (cluster 4, 12.8%), and verbally aggressive and nonaggressive behaviors (cluster 5, 7.7%).

No statistically significant correlations were found between the 5 behavior clusters and demographic variables or comorbidity level. There was a negative correlation between cluster 5 (verbally aggressive and nonaggressive behaviors) and functional impairment (P = .009).
Analysis of medication use showed that 27% (32/117) of patients with clinically significant BDs received a neuroleptic medication, and 15% (17/117) received a benzodiazepine (see Table 1). Individuals in the behavior group were treated more frequently with a benzodiazepine (17/117) \( (P = .004) \) or a neuroleptic (32/117) \( (P = .009) \) compared with those in the nonbehavior group (4/116 and 16/116, respectively). The use of neuroleptic medications and antidepressants did not differ between the 5 clusters of BD. Patients in cluster 2 (behaviors from all categories) received benzodiazepines more frequently \( (P = .014) \) than patients in other clusters.

### DISCUSSION

This study reports on the prevalence and treatment pattern of BDs in five nursing homes in one community. Patients with clinically significant BDs were more dependent in bathing and eating compared with those without clinically significant BDs. Higher use of neuroleptics and benzodiazepines was found among patients with clinically significant BDs compared with patients without clinically significant BDs. Among patients with BDs, those with BDs in all four categories (cluster 2) were more often treated with benzodiazepines. All remaining clusters did not correlate with specific classes of psychotropic medications.

The finding that verbal nonaggression is the most prevalent behavior in nursing homes is consistent with previous reports. A majority of patients (72%) with clinically significant BDs had verbal nonaggression either alone or in combination with other BDs. Physically aggressive behaviors were the least prevalent. It has previously been noted that the four behavior categories are not independent and that BDs from different categories co-occur with varying frequencies.

Patients with clinically significant BDs had higher impairment in self-feeding and bathing compared with patients without significant BDs. Our results are consistent with previous reports that examined correlates of agitated behaviors. In a sample of 24 cognitively impaired nursing home residents, Cohen-Mansfield et al found that eating, bathing, and toileting were associated with more agitation when nursing staff members initiated these activities. Other studies have found significant correlation between verbal agitation (screaming, cursing) and routine care activities such as bathing and feeding. These findings have implications for the use of nonpharmacologic interventions to reduce triggers for BDs.

The finding of increased benzodiazepine use among patients with the widest range of BDs indicates that BDs are being identified for treatment; however, the treatment may not be efficacious. The results also support the evidence that benzodiazepines may induce paradoxical agitation in the elderly. Recent reports do not recommend the use of benzodiazepines for treatment of agitation in older adults. In addition to paradoxical agitation, these medications have been associated with falls, impaired cognition, and daytime sedation. They may also induce delirium, leading to a variety of agitated behaviors.

The significantly higher use of neuroleptic medications for patients with clinically significant BDs suggests that these medications are of limited efficacy for certain BDs. A meta-analysis of neuroleptic agents for the treatment of agitated behaviors found only a modest response to these medications. Placebo response rate ranged between 40% and 60%. In contrast to the studies of psychoses in dementia, a recent study of BDs in dementia found no difference in outcome between treatment with haloperidol,
trazodone, behavior management, or placebo. A study on withdrawal of neuroleptics in SNFs found short-term improvement in 55% of patients and no change in another 22%. Overall, there is a paucity of placebo-controlled trials to show the effectiveness of neuroleptic medications in the treatment of BDs.

There are several limitations to this study. First, the BMS simplifies the rating procedure and does not permit rating of individual BDs within each category. Second, the results were based only on frequency of behaviors. There were no data on the severity of specific BDs, such as physically aggressive behaviors, and their correlation with any of the variables tested. Third, the study did not address medication doses or differentiate between various medications within the same class. Fourth, the large number of correlations examined may increase the likelihood of significant findings occurring by chance. Last, the pattern of medication use found in this study may differ from that in other communities or regions in the United States.

In conclusion, this study found no pattern of medication use to suggest targeted treatment of specific clusters of BDs. There does not appear to be a specific treatment pattern of BDs in our community in which psychotropic medications are chosen based on distinct characteristics of a particular behavior cluster. We anticipate that the results of the present study can be used in the development of treatment guidelines for BD in dementia.

Acknowledgments
The authors would like to acknowledge the Departments of Nursing at the Deaconess, Waterfront, Degraff, Millard-Fillmore, and Greenfield skilled nursing facilities for their participation in this research project. We acknowledge the efforts of Angela Vacanti, Maureen Fassl, and Eileen Green in ensuring proper data collection and providing expert data entry. The authors are grateful to Dr. Joseph Mylotte, who reviewed the final version of the manuscript and provided expert opinion.

References


Appendix 1

BMS

Circle the behavior/s exhibited by the resident (categories 1-4) daily during the 7-3, 3-11, and 11-7 nurses shifts. Circle 0 if no behavior is present.

Name: ____________________________

Unit: ____________________________

Month/Year: ________________________

Key: 1 = Verbally Aggressive

2 = Verbally Non-Aggressive

3 = Physically Aggressive

4 = Physically Non-Aggressive

| 7 - 3 Shift | DAY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
|-------------|-----|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Behavior    | Category | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|             | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
|             | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
|             | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |

| 3 - 11 Shift | DAY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
|-------------|-----|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Behavior    | Category | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|             | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
|             | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
|             | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |

| 11 - 7 Shift | DAY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
|-------------|-----|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Behavior    | Category | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|             | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
|             | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
|             | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |

A. Behavioral Intervention
B. Medication Changes
C. Side Effects of Medication

WHCC-N6398 © 1998, SUNY at Buffalo. All Rights Reserved.

F.R.

BMS Key

1 - Verbally Aggressive Behaviors
   -Threatening
   -Cursing
   -Verbal sexual advances
   -Verbal aggression

2 - Verbally Nonaggressive behaviors
   -Constant requests for attention
   -Repetitious sentences/questions
   -Screaming
   -Making strange noises
   -Complaining
   -Demanding
   -Crying
   -Talkativeness

3 - Physically Aggressive Behaviors
   -Hitting
   -Pushing
   -Scratching
   -Splitting
   -Kicking
   -Grabbing
   -Throwing things

4 - Physically Nonaggressive Behaviors
   -Pacing
   -Disrobing
   -Restlessness
   -Smearing feces
   -Sexual exhibition/sexual activities
   -Negativism
   -Uncooperative with care
   -Refuses to eat
   -Aimless wandering
   -Refuses to take medications
   -Trying to leave unit
Health Service Utilization by Alzheimer’s Disease Patients: A 2-Year Follow-up of Primary Versus Subspecialty Care

Peter M. Aupperle, MD, MPH, Edward R. MacPhee, BA, Andrew C. Coyne, PhD, Jonathan Blume, BS, and Betty Sanchez, MS

ABSTRACT

All dementia patients and their caregivers who had received a university-based comprehensive evaluation and a diagnosis of Alzheimer’s disease during 1997 (N = 80) were identified. Of the original cohort, 48.8% (n = 39) were able to be contacted approximately 2 years after their initial assessment, and the caregivers were the informants for this follow-up. Two subgroups were defined: 22 patients were being seen only by their primary care physicians (MED), while 17 patients were also being treated by a geriatric psychiatry faculty member (GERO). There were statistically significant differences between the 2 groups (MED versus GERO, respectively) at follow-up in terms of (1) institutionalization (30.0% versus 4.6%, P < .05), (2) CDR (2.3 versus 1.5, P < .005), and (3) prescription of donepezil at follow-up (45.5% versus 76.5%, P = .05). These differences are being assessed in a larger scale prospective study. (J Geriatr Psychiatry Neurol 2003; 16:15–17)

Keywords: Alzheimer’s disease; primary care; health service utilization

The cost of caring for patients with dementia is twice as high as that of the average Medicare patient. With approximately 4 million Americans currently diagnosed with Alzheimer’s disease (AD) accounting for 10% to 30% of nursing home admissions, this amounts to nearly 70 billion dollars annually. Therefore, considering the cost and the growing number of AD patients, it is imperative to address psychosocial issues and delay institutionalization. Current practice guidelines stress 3 treatments to accomplish this goal: cognitive enhancers, specifically cholinesterase inhibitors, to provide symptomatic treatment of cognitive deficits; psychotropics for behavioral complications; and psychoeducation for caregivers.

However, many patients rely on primary care physicians to treat dementia, and it is often difficult for them to provide the 3 areas of treatment suggested by the guidelines. They are under time limitations and are sometimes not fully informed about diagnosing patients with dementia. As a result, they may fail to appropriately assess the severity of the patient’s cognitive deficits and often lack a well-developed plan for managing the long-term psychosocial problems of the disease. The constraints on primary care physicians may also lead to differences in care for their patients. This was previously demonstrated in a pilot 1-year follow-up study that looked at patients with AD who were treated by their primary care physician only or in combination with a geriatric psychiatrist. It noted significant differences in treatments and outcomes between the 2 groups. There existed a greater rate of hospitalization and cognitive decline in the cohort treated only by a primary care physician. In addition, the percentage of patients that maintained donepezil treatment, the only cognitive enhancer available at the time, at the 1-year follow-up was significantly lower in this cohort.

Given the results of the above study, a 2-year follow-up was conducted to reevaluate the aforementioned patients and to expand our understanding of the potential differential outcomes. Specifically, the study was broadened to include rates of institutionalization and psy-
chotic prescriptions. In addition, donepezil usage, hospitalization, and cognitive status were reassessed at a 2-year interval.

METHODS

All dementia patients and their caregivers who had received a comprehensive neuropsychiatric evaluation (including neuropsychological testing) and a diagnosis of AD at a university-based diagnostic clinic during 1997 (N = 80) were identified. From the original cohort, 57 were contacted after 1 year, and 39 of those remaining could be contacted at the 2-year follow-up. Of these 39 patients, 22 had opted to be seen only by their primary care physician (MED) for their AD, while 17 continued to be treated at the university-based clinic by a geriatric psychiatry faculty member (GERO) in collaboration with a case manager (eg, geriatric social worker or geriatric nurse). The comprehensive case management provided to the GERO cohort included AD education, review of caregiver coping skills, behavioral management, community resources, long-term care planning, and legal/financial planning.

All baseline data were collected from the initial comprehensive evaluation, and demographic data and assessment of physical impairment with the Cumulative Illness Rating Scale (CIRS) were collected via standardized chart reviews. Data collected at follow-up were obtained by one of the authors via telephone contact with the patient’s primary caregiver. This information included assessments of use of health services by the patient (institutionalization, hospitalizations for nonpsychiatric problems, usage of home health aides, emergency room visits, primary care doctor visits, and the usage of dementia day programs), prescriptions, and cognition (Clinical Dementia Rating Scale [CDR]). Nonparametric and correlational analyses of the baseline and follow-up data were performed using the SAS Language for Personal Computers. The protocol was approved by the University of Medicine and Dentistry of New Jersey Institutional Review Board, and informed consent was obtained from the caregiver after a complete description of the study.

RESULTS

The cohort at the 2-year follow-up (n = 39) was a subset of the original cohort of patients diagnosed with AD at baseline (n = 80). To assess whether this subset was not demographically or clinically disparate from the original cohort, the 2-year follow-up cohort was compared with the 41 patients that could no longer be assessed at year 2. The cohorts were not statistically different in terms of age (78.8 vs 78.4 years), gender, marital status, living arrangement, and relationship of caregiver. In terms of cognition and medical comorbidity, there were also no statistical differences: CDR = 1.7 versus 1.5 (t = 0.03, df = 50, P = .97), CIRS = 2.3 versus 1.8 (t = 0.57, df = 54, P = .57). The demographic characteristics of the MED (n = 22) and GERO (n = 17) groups at year 2 were compared, and there were no statistical differences between the 2 cohorts. Analysis of variance also revealed no significant differences in the CIRS scores of the groups.

Chi-square analyses revealed a significantly greater rate of institutionalization among MED patients, specifically with respect to placement in a nursing home or the combination of admission to an assisted living facility or nursing home (P < .05). Hospitalization rates, usage of home health aides, emergency room visits, primary care doctor visits, and the usage of dementia day programs were statistically similar in both groups. Chi-square analyses also revealed a smaller percentage of MED patients receiving donepezil (P = .05). Behavioral complications were also taken into account by determining if the patient was on a psychotropic medication (an antidepressant, antipsychotic, or antianxiety prescription), and no significant differences were found between the MED and GERO cohorts (Table 1). The baseline CDRs for each cohort were statistically equivalent (P = .53). Analyses of variance showed a significantly greater CDR in the MED cohort relative to the GERO cohort at year 2 (2.3 vs 1.5, F = 11.71, P < .005). This difference was evidenced by a divergence in the CDR scores between the 2 cohorts.

DISCUSSION

In this 2-year follow-up study, 3 important themes emerged in comparing the treatment of AD by primary care physicians versus geriatric psychiatrists: (1) a significantly greater rate of institutionalization for those being treated only by primary care physicians; (2) less usage of donepezil (45.5% vs 76.5%), consistent with the results of the 1-year follow-up; and (3) greater cognitive decline via CDR scores, also consistent with the previous study.

Data on the usage of psychotropic medications were evaluated to assess for behavioral complications between the 2 cohorts, and no significant differences existed. Hospitalization rates (38.7% vs 14.8%) and usage of home health aides (45.2% vs 18.5%) were significantly different at the 1-year follow-up but not at the 2-year follow-up.

There exist several areas of concern regarding the results of this study. Since only 39 of the original cohort of 80 could be contacted at the 2-year follow-up, there was a selection bias. However, no significant baseline differences were noted between these cohorts, leading to the conclusion that comparisons between the remaining subjects are useful and that any bias due to attrition was minimal. It should also be noted that in this retrospective study, the cohorts were not randomized, the interviewers were not blinded, and the informant was the caregiver.

In summary, significant differences exist at the 2-year follow-up in this pilot study of 2 different models of care. Primary care intervention trials would be useful in assessing differences in outcomes after an educational
intervention to determine if a collaborative care model is efficacious. Specifically, one would randomize equivalent clinical sites and provide specialized intensive education and consultation for the physicians at the “intervention” sites, while providing only minimal general education at the “control” sites. One would then assess the clinical outcomes and health service utilization. In addition, the assessment of the cost/benefit implications (both direct and indirect) of such an intervention would be necessary. A large-scale prospective study with this format is currently in progress.

Such research is akin to the large body of data demonstrating success in enhancing the treatment of depression in general adults in the primary care setting. Given that the majority of older adults receive treatment for cognitive or behavioral changes from their primary care physicians, the goal of the above investigations is to enhance the outcome in this setting. It is hoped that in the future, all AD patients will have access to state-of-the-art care, with a resultant improvement in quality of life for themselves and their caregivers.

**References**


**Table 1. Two-Year Outcome: Primary Care Physicians (MED) Versus Geriatric Psychiatrist Care (GERO)**

<table>
<thead>
<tr>
<th>Service usage</th>
<th>MED (n = 22)</th>
<th>GERO (n = 17)</th>
<th>χ²</th>
<th>P</th>
<th>OR</th>
<th>1/OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospitalization</td>
<td>5 (22.7%)</td>
<td>2 (11.8%)</td>
<td>0.78</td>
<td>.38</td>
<td>0.45</td>
<td>2.21</td>
</tr>
<tr>
<td>Assisted living</td>
<td>4 (18.2%)</td>
<td>1 (5.9%)</td>
<td>1.30</td>
<td>.26</td>
<td>0.28</td>
<td>3.56</td>
</tr>
<tr>
<td>Nursing home</td>
<td>5 (22.7%)</td>
<td>0 (0.0%)</td>
<td>4.06</td>
<td>.04</td>
<td>0.00</td>
<td>Undefined</td>
</tr>
<tr>
<td>Assisted living/nursing home</td>
<td>9 (40.9%)</td>
<td>1 (5.9%)</td>
<td>5.30</td>
<td>.01</td>
<td>0.09</td>
<td>11.08</td>
</tr>
<tr>
<td>Medication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Donepezil</td>
<td>10 (45.5%)</td>
<td>13 (76.5%)</td>
<td>3.81</td>
<td>.05</td>
<td>3.90</td>
<td>0.26</td>
</tr>
<tr>
<td>Antidepressant</td>
<td>8 (36.4%)</td>
<td>8 (47.1%)</td>
<td>0.15</td>
<td>.70</td>
<td>1.56</td>
<td>0.64</td>
</tr>
<tr>
<td>Antianxiety</td>
<td>1 (4.5%)</td>
<td>1 (5.9%)</td>
<td>0.04</td>
<td>.85</td>
<td>1.31</td>
<td>0.76</td>
</tr>
<tr>
<td>Antipsychotic</td>
<td>3 (13.6%)</td>
<td>3 (17.6%)</td>
<td>0.12</td>
<td>.73</td>
<td>1.36</td>
<td>0.74</td>
</tr>
</tbody>
</table>

OR = odds ratio.
Smoking and Cognitive Performance in the Community Elderly: A Longitudinal Study

Wei-Ta Chen, MD, Pei-Ning Wang, MD, Shuu-Jiun Wang, MD, Jong-Ling Fuh, MD, Ker-Neng Lin, PhD, and Hsiu-Chih Liu, MD

ABSTRACT

This prospective study investigated the association between smoking and cognitive performance in a community of nondemented elderly subjects aged 65 or older. All subjects were categorized as current smokers, former smokers, or never smokers. The lifetime cigarette exposure was computed. At baseline, we found the abstainers from smoking had better cognitive performances; however, the differences were not significant after adjusting for age, education, hypertension, diabetes, and vascular events. The lifetime cigarette exposure was not predictive of the cognitive status. At a 3-year follow-up, neither the smoking status nor the lifetime cigarette exposure predicted the declination of cognition. (J Geriatr Psychiatry Neurol 2003; 16:18–22)

Keywords: smoking; elderly; cognitive performance

Previous researchers have shown that nicotine administration had cognitive-enhancing effects in normal adults.¹ Nicotine administration would activate brain areas pertaining to attention, working memory, motivation, mood, and emotion.²,³ Similar effects also exist in patients with Alzheimer’s disease (AD); nicotine injection or skin patches significantly improve patients’ attention, learning, and memory.⁴ Even though these biological evidences suggest a protective effect of smoking against AD or cognitive decline, the results from epidemiological studies did not show consistent benefits of smoking to cognition. Notably, smoking may adversely increase the risk of cognitive impairment through atherosclerotic and hemodynamic processes.⁵

Recently, several large-scale longitudinal studies have suggested that smoking might be a risk factor for AD or cognitive decline.⁶-⁸ In addition, some longitudinal studies did not indicate any relationship between smoking and cognitive function.¹⁰-¹²

To date, the impact of long-term cigarette smoking on cognition has yet to be clarified. The purpose of this article is to use cohort data from the Kinmen Neurological Disorder Survey (KINDS)¹³,¹⁴ to examine whether cigarette smoking is predictive of cognitive status (cross-sectional relationship) or cognitive decline (longitudinal relationship).

METHODS

Study Setting and Design

The general methodology of the KINDS has been detailed elsewhere.¹³-¹⁴ The KINDS was a community-based, 2-phase epidemiological study of several neurological disorders including dementia, Parkinson’s disease, essential tremor, stroke, and headache. The target population was registered residents aged 65 years and older living in Kinmen, a rural islet located west of Taiwan and near the coast of mainland China. The population in Kinmen was quite stable; the rate of immigration and emigration was very low. Most of the population lived on farms and were illiterate since compulsory education was not introduced until the 1950s.

A total of 798 nondemented subjects (424 women, 374 men) received a test of cognitive functioning and completed the first phase of the KINDS in 1993. In the follow-up phase of the KINDS in 1996, 671 subjects (84.1%, 357 women and 314 men) successfully repeated the cognitive testing. Data including demographics as well as medical histories were collected, and the smoking histories of each subject at baseline were assessed. Subjects with dementia identified during the first phase were excluded from data analysis because their recall of medical and smok-
ing histories was unreliable and their cognitive function declined beyond that associated with normal aging.

**Smoking History Assessment**
The assessment of smoking history started with a trigger question: “Did you ever smoke during your lifetime?” When the answer was “yes,” further detailed information would be collected. Those who answered “yes” were asked at what age they began smoking, whether they were still smoking, at what age they quit smoking (or the number of years since they quit smoking), and how many cigarettes on average they had smoked or still smoked per day.

**Outcome Variables: Cognitive Measures**
The cognitive performances were assessed by the Chinese version of the Cognitive Abilities Screening Instrument (CASI).\(^{15,16}\) The CASI is a comprehensive cognitive test validated against clinically diagnosed dementia\(^{17}\) and has a higher sensitivity/specificity than the Mini-Mental State Examination (MMSE).\(^{17}\) The CASI was originally designed for cross-cultural studies and is globally recognized nowadays. Pilot studies conducted in Japan and in the United States have demonstrated its cross-cultural applicability and usefulness in screening for dementia, in monitoring disease progression, and in providing profiles of cognitive impairment.\(^{15}\) The CASI has a score range of 0 (worst) to 100 (best).\(^{15}\) It takes about 20 minutes to administer and includes 9 domain subscales: attention, concentration, orientation, long-term memory, short-term memory, abstraction and judgment, language abilities, constructional praxis, and category fluency. The MMSE score could be derived from CASI items according to the CASI manual.\(^{15}\)

The CASI has shown good reliability and validity in the Chinese illiterate population of Kinmen.\(^{16}\) Educational disparity influences the performance in the CASI test; people with varied educational achievements score differently on the CASI test.\(^{18}\) Our recent study\(^{18}\) obtained cutoff scores of the Chinese version of the CASI in dementia diagnosis for different educational groups: schooling year = 0: 49/50 (sensitivity = 0.83; specificity = 0.85); schooling year = 1–5: 67/68 (sensitivity = 0.83; specificity = 0.91); schooling year ≥ 6: 79/80 (sensitivity = 0.89; specificity = 0.90).

**Potential Confounders**
The potential confounders considered in the present study included age, educational attainment, hypertension, diabetes, and vascular events. Data of these confounders were based on demographics and medical histories collected at baseline. Notably, educational attainment was recorded as schooling years. Vascular events was an artificially composed variable indicating the presence of one of the following self-reported histories: stroke, transient ischemic attack, angina pectoris, and myocardial infarction. These variables were considered confounders because they are associated with either smoking or cognitive function and possibly mediate the relationship between them both.\(^{19,20}\)

**Statistical Analyses**
In this study, the CASI score was the measure of cognitive performance; the CASI score change between both phases (follow-up CASI score minus baseline CASI score) was computed as a measure of cognitive decline. Data analyses included mainly 2 parts. First, all subjects were classified by the smoking status at baseline as current smokers, former smokers, or never smokers. We examined the cross-sectional relationship between cognitive performance and smoking status. Subsequently, we tested the longitudinal relationship between smoking status and cognitive decline. The differences were analyzed by the \(\chi^2\) test for categorical variables and one-way analysis of variance (ANOVA) for continuous variables. Crude correlations of the smoking status to baseline cognitive performance or cognitive decline were further examined by a multiple linear regression model, adjusting for potential confounders. In this model, age and education were entered as continuous variables; hypertension, diabetes, and vascular events were considered as categorical variables.

For the second part of data analyses, the number of pack years of each subject was computed as a measure of lifetime cigarette exposure by the formula (daily number of cigarettes smoked/20 cigarettes per pack) \(\times\) years smoked. We tested the associations between lifetime cigarette exposure amount and baseline cognitive performance, or the cognitive decline, by Pearson’s correlation. Crude associations were adjusted for potential confounders by linear regression analysis.

Additional analyses focused on the influence of quitting smoking on cognitive performance. Using the same statistical strategy as the second part, we checked the relationship between quitting duration, which was recorded as the number of years since smoking cessation, and the cognitive status or decline.

All statistical analyses were carried out using SPSS for Windows (SPSS Inc., Chicago, release 10.0, 1999), and the statistical levels quoted (\(P\) values) were two-tailed.

**RESULTS**

**Subjects Selection**
The preliminary results of baseline KINDS revealed that most men (78%) had smoked; however, few women (only 10%) were current or former smokers. The enormous gender disparity also existed in educational achievement: 65% of men had received formal education, although only a few years of schooling, but only 11% women were educated. Because the women smokers were few, and considering the education effect on their cognitive performance, we decided to use data from only male subjects to eliminate the confounding effect of gender differences. The data...
analyses thus finally focused on 314 nondemented male subjects who completed the CASI tests in both phases and had no missing data in all variables of interest. A comparison of the included with the excluded men disclosed similar age, educational attainment, and prevalence of hypertension, diabetes, and vascular events.

**Smoking Status and Cognitive Performance**

Of the 314 included subjects, 68 (21%) were never smokers, 51 (16%) were former smokers, and 195 (63%) were current smokers. The duration of smoking was 24.0 ± 17.4 years for former smokers and 40.8 ± 19.0 years for current smokers. The distributions of the demographics, hypertension, diabetes, vascular events, and the CASI scores (baseline, follow-up, and decline) are summarized in Table 1. Those who continued to smoke were older (P = .025) and had fewer schooling years (P = .008). The proportion of the subjects with hypertension, diabetes, and vascular events did not differ across smoking status categories. At baseline, the former smokers had significantly higher CASI scores than other smokers did (P = .013); however, the difference disappeared after adjustment for age, education, hypertension, diabetes, and vascular events (see Table 2, model 1). The CASI score changes between both phases were comparable regardless of smoking status.

After grouping the former and the current smokers (ie, the "ever smokers") for further comparison with those that never smoked, these 2 groups had similar mean age (72.3 ± 6.2 vs 72.0 ± 6.0 years, P = .665), but the ever smokers had fewer schooling years (3.1 vs 4.4 years, P = .012). The baseline CASI score (78.4 ± 12.0 vs 80.9 ± 9.7, P = .116) and the CASI score change (–0.6 ± 11.1 vs –2.0 ± 9.0, P = .323) of ever smokers and never smokers were comparable. Since the former and current smokers had significantly fewer schooling years than the never smokers did, this may act to minimize any differences related to smoking status. Thus, we divided all the subjects into 2 groups by their education (illiterate vs literate) for subanalysis. In the illiterate subgroup, never and ever smokers had a comparable mean CASI score (74.9 ± 10.0 vs 69.9 ± 11.7,

### Table 1. Demographics, Hypertension, Diabetes, Vascular Events, and Cognitive Measures in All of the Subjects and Subgroups Categorized by Smoking Status

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Total Subjects (n = 314)</th>
<th>Never Smokers (n = 68)</th>
<th>Former Smokers (n = 51)</th>
<th>Current Smokers (n = 195)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years (SD)*</td>
<td>72.0 (6.0)</td>
<td>72.3 (6.2)</td>
<td>70.0 (4.3)</td>
<td>72.5 (6.3)</td>
</tr>
<tr>
<td>Education in years (SD)*</td>
<td>3.4 (3.6)</td>
<td>3.9 (3.8)</td>
<td>2.9 (3.4)</td>
<td></td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>26</td>
<td>25</td>
<td>35</td>
<td>23</td>
</tr>
<tr>
<td>Diabetes mellitus (%)</td>
<td>8</td>
<td>15</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Vascular events (%)</td>
<td>9</td>
<td>10</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Baseline CASI score (SD)*</td>
<td>79.0 (11.6)</td>
<td>80.9 (9.7)</td>
<td>82.0 (10.3)</td>
<td>77.5 (12.3)</td>
</tr>
<tr>
<td>Follow-up CASI score (SD)</td>
<td>78.1 (14.5)</td>
<td>78.9 (13.2)</td>
<td>81.3 (14.3)</td>
<td>76.9 (15.0)</td>
</tr>
<tr>
<td>CASI score change (SD)</td>
<td>-0.9 (10.7)</td>
<td>-2.0 (9.0)</td>
<td>-0.7 (10.4)</td>
<td>-0.6 (11.4)</td>
</tr>
<tr>
<td>Baseline MMSE-CE score (SD)*</td>
<td>24.8 (4.0)</td>
<td>25.3 (3.7)</td>
<td>25.8 (3.1)</td>
<td>24.3 (4.2)</td>
</tr>
<tr>
<td>Follow-up MMSE-CE score (SD)</td>
<td>24.5 (4.6)</td>
<td>24.6 (4.4)</td>
<td>25.3 (4.4)</td>
<td>24.2 (4.7)</td>
</tr>
<tr>
<td>MMSE-CE score change (SD)</td>
<td>-0.3 (4.0)</td>
<td>-0.7 (3.4)</td>
<td>-0.5 (3.4)</td>
<td>-0.1 (4.2)</td>
</tr>
</tbody>
</table>

CASI = Cognitive Abilities Screening Instrument; MMSE-CE = CASI-estimated Mini-Mental State Examination.

### Table 2. Results of Linear Regression Analyses to Test Smoking Status (Model 1) and Quitting Duration (Model 2) as a Predictor of Baseline CASI Score Controlling for Potential Confounders

<table>
<thead>
<tr>
<th>Smoker Status</th>
<th>β Coefficient</th>
<th>SD</th>
<th>P Value</th>
<th>β Coefficient</th>
<th>SD</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current smokers</td>
<td>-1.12</td>
<td>1.40</td>
<td>.424</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Former smokers</td>
<td>0.69</td>
<td>1.85</td>
<td>.707</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Never smokers</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Quitting duration</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
<td>0.02</td>
<td>.753</td>
</tr>
<tr>
<td>Age</td>
<td>-0.33</td>
<td>0.09</td>
<td>&lt; .001</td>
<td>-0.35</td>
<td>0.09</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Education</td>
<td>1.55</td>
<td>0.16</td>
<td>&lt; .001</td>
<td>1.58</td>
<td>0.15</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Hypertension</td>
<td>0.95</td>
<td>1.34</td>
<td>.476</td>
<td>1.09</td>
<td>1.33</td>
<td>.411</td>
</tr>
<tr>
<td>Diabetes</td>
<td>-0.04</td>
<td>2.13</td>
<td>.984</td>
<td>-0.04</td>
<td>2.11</td>
<td>.986</td>
</tr>
<tr>
<td>Vascular event</td>
<td>-3.59</td>
<td>1.95</td>
<td>.067</td>
<td>-3.70</td>
<td>1.95</td>
<td>.059</td>
</tr>
</tbody>
</table>

CASI = Cognitive Abilities Screening Instrument. Adjusted R² = 0.30 for both model 1 and model 2.
In the literate subgroup, both smokers did not score differently (82.5 ± 9.0 vs 83.8 ± 8.7, P = .348).

The Influence of Lifetime Exposure and Duration of Quitting

The lifetime cigarette exposure ranged from 0 to 140 pack years, with a mean of 27.9 ± 28.5 pack years. Age and education had no association with the lifetime exposure. Likewise, the baseline CASI score or the CASI score change did not correlate with lifetime exposure. The lack of association remained unchanged in analyses confined to the ever smokers or the current smokers alone.

The quitting duration among the former smokers averaged 7.7 ± 10.5 years. It was slightly associated with education (γ = 0.13, P = .02) but not with age. A significant correlation existed between the quitting duration and the CASI scores at both phases (γ = 0.44, P = .002 for baseline; γ = 0.12, P = .038 for follow-up). Nonetheless, the crude associations disappeared after adjusting for all the potential confounders (Table 2, model 2). Likewise, the CASI score change was not associated with the quitting duration. A similar lack of correlation occurred in further analyses pooling data from the current and the former smokers, in which the quitting duration of the latter was regarded as zero (year).

DISCUSSION

The present study did not demonstrate any consistent effects of smoking histories (including the current smoking status, lifetime cigarette exposure, and number of years since smoking cessation) on cognitive status and cognitive decline. The crude data showed a better cognitive performance by those who once smoked but quit and a positive association between the years after quitting and the cognitive status. Lifetime cigarette exposure, and number of years since smoking cessation) on cognitive status and cognitive decline. The crude data showed a better cognitive performance by those who once smoked but quit and a positive association between the years after quitting and the cognitive status. However, these findings were not significant after adjustment for age, education, hypertension, diabetes, and vascular events.

Our results were in line with those obtained from a Cleveland cohort study and a Boston cohort study. These 2 studies did not disclose any substantial impact of smoking on cognition, although they focused on some specific cognitive fields rather than the global cognitive function. One study from the Medical Research Council (MRC), also revealed negative association between smoking and cognition, but interpretation of the study results was limited to individuals with hypertension. Similarly, recent laboratory studies of brain function via quantitative electroencephalography (EEG) or event-related potentials (ERP) did not evidence any marked differences between elderly lifelong nonsmokers and smokers.

Contrary to these negative findings, some longitudinal studies suggested smoking as an isolated predictor of cognition. A French cohort study showed some benefits of smoking on attentional and visuospatial functioning, but dose effect was unavailable. A Dutch cohort study disclosed harmful effects of smoking on cognition, but the participants with dementia were not excluded at the baseline. Moreover, it might be of concern that the data collection for cognitive outcomes was administered in different situations: research center at baseline versus home at follow-up. Another cohort study conducted in London indicated that smoking might be a prospective risk factor for incidental cognitive impairment. However, this finding was still limited by a short follow-up interval (1 year), less comprehensive cognitive measures, and absence of dose effect analyses. The Honolulu-Asia Aging Study (HAAS) used the CASI to measure cognition and disclosed positive association between smoking during middle age and later risk of cognitive impairment. Nevertheless, congruent findings in dose effect are lacking. In conclusion, many possible unmeasured confounders (eg, genetic susceptibility) of the relation between smoking and cognition remain to be considered before these “positive” studies can be validly interpreted.

Our study results reflect the possibility that harmful effects after long-term cigarette exposure in atherosclerosis and hemodynamic change might “counterbalance” the beneficial effects mediated through nicotine receptors. However, current smokers did not have a higher prevalence of vascular risks in our study (Table 1). This result might be due to confounding from the “healthy smokers effect.” The rate of cognitive decline in nondemented aging may change with different education attainment. To date, the role of education in predicting cognition is still a matter of discussion. The rate of cognitive decline in nondemented aging may change with different education attainment. Furthermore, the age of our study subjects were on average 72 years at baseline and 75 years at the end of the 3-year cohort study. Age also may play a role in the speed of cognitive decline. Further studies with a longer follow-up interval in subjects with older age are necessary. Finally, some potential confounders that may distort the relationship between smoking and cognition should be controlled, such as head injury, medication use, and so forth.

We examined the predictive value of smoking on global cognitive function and found little evidence of any significant association. Further studies with longer follow-up interval and using more cognitive measures are mandatory to confirm our results since the impact of smoking history on late-life cognition is enormously complex, involving at least lifelong trends in behavior, morbidity, and mortality.
Acknowledgments
This study was supported by grants from the National Science Council, NSC86-2314-B075-070.

References
Construct Validity of the 15-Item Geriatric Depression Scale in Older Medical Inpatients

R. Antonelli Incalzi, MD, M. Cesari, C. Pedone, P. U. Carbonin

ABSTRACT

The construct validity of the 15-item Geriatric Depression Scale (sfGDS) has been assessed in selected populations. The aim of this study was to assess the appropriateness of applying the sfGDS to unselected older inpatients. The main component analysis of sfGDS was performed in 2032 medical inpatients (mean age = 76.3 ± 8.4). sfGDS did not qualify as a unidimensional test. Three factors explained 47.7% of variance and explored the following dimensions: positive attitude toward life, distressing thoughts/negative judgment about the own condition, and inactivity/reduced self-esteem. The internal homogeneity was poor (Cronbach’s α = .46). A higher fraction of variance was explained in patients independent in all or dependent in ≥ 1 activity of daily living (ADL). In older medical inpatients, sfGDS is not a single construct, which prevents the univocal interpretation of the final score. The higher fraction of explained variance in patients with comparable ADL performance probably reflects the dependency of affective from physical status. (J Geriatr Psychiatr Neurol 2003; 16:23–28)

Keywords: screening of depression; elderly; medical inpatients; short form GDS

The short form of the Geriatric Depression Scale (sfGDS) is currently used for clinical and research purposes.1 Its appropriateness as a screening tool for home-dwelling elderly people has been questioned by Alden et al.2 Burke et al3 reported that the sfGDS is an effective screening instrument in cognitively intact but not in mildly demented patients. The sfGDS was very effective in identifying depressed subjects in the Greek but not in the Israeli elderly population.4,5 The only study that validated the sfGDS among inpatients was performed in a selected population that included depressed, demented, and thought-disordered people and not in the broad hospitalized geriatric population, which is more commonly screened for depression.6

The present study aimed at verifying whether the final score of sfGDS can be considered really representative of the outcome depression and to what extent impairment of selected psychological dimensions can be recognized by analyzing the performance of older medical inpatients.

METHOD

The present study uses data from a large collaborative observational study group, the Italian Group of Pharmacoepidemiology in the Elderly (Gruppo Italiano di Farmacoepidemiologia nell’anziano [GIFA]). The GIFA is a multicenter study involving wards of geriatrics and internal medicine in community or university hospitals scattered over the whole Italian territory. The main objective of the GIFA study is to survey drug consumption, incidence and type of adverse drug reactions, and quality of hospital care. For the present study, we used data collected by 24 centers in the last survey performed (May-June and September-October 1998) because the sfGDS was not used in previous surveys.

Procedures were approved by the Catholic University Ethical Committee as well as by the Steering Committee of CNR–Aging Project. A detailed description of proce-
dures is available elsewhere. Briefly, patients underwent a multidimensional assessment covering several domains: sociodemographic characteristics, smoking and alcohol intake history, medical problems, pharmacological therapy and adverse drug reactions, complete blood count, cognitive and affective status, and functional capabilities. Each domain was assessed according to a standardized and validated method, for example, functional capabilities by activities of daily living (ADL) and instrumental activities of daily living (IADL) scales, comorbidity by the Charlson’s index, and cognitive status by the Abbreviated Mental Test (AMT). Diseases and drugs were codified by the International Classification of Diseases, 9th revision (ICD-9), and the Anatomical and Therapeutical Classification, respectively. A detailed description of the GIFA protocol is available elsewhere.

Affective status was assessed using the sfGDS. Five items explore a positive attitude toward life (happy most of the time, satisfied with life, wonderful to be alive, full of energy, in good spirit) and thus can be defined as positive items. The 10 negative items generically assess dissatisfaction with life (life is empty, feel pretty worthless, feel helpless, often get bored, worry about the future, situation is hopeless, others are better off) or point at selected personal problems (problems with memory, prefer to stay home, dropped activities/interests). Every item contributes 1 point to the final score, which ranges from 0 to 15. A score greater than 5 is commonly considered as indicative of a depressive trait.

The present analysis refers to patients older than 60 years of age consecutively admitted to the participating centers. By protocol, a study physician administered the sfGDF on the day before the planned discharge because a great proportion of patients were too seriously ill to be interviewed on admission. Patients who died during hospital stay (n = 138) were excluded from the study. A total of 363 patients were excluded because of 1 or more of the following reasons: aphasia, illiteracy, deafness, unconsciousness, unwillingness to cooperate, or severe dementia. Due to these exclusion criteria, the analysis was conducted on 2032 patients.

**Statistical Analysis**

Statistical analyses were performed using SPSS software. Two complementary methods, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy and Bartlett’s test of sphericity, were used to evaluate the appropriateness of factor analysis. The aim of this preliminary phase was to evaluate whether items are reciprocally related so that a factor model can be developed. Weakly correlated variables cannot share common factors, thus making factor analysis inappropriate. Factors were extracted by the principal component analysis. From the basic component matrix, a rotated component matrix was obtained through a varimax with Kaiser normalization. The rotation phase allows the identification of factors summarizing sets of closely related variables and then being more meaningful and easily interpretable than those obtained in the extraction phase. The varimax method pursues this objective by minimizing the number of items having loadings on a factor. The strength of the relationship of items to individual factors is directly proportional to the magnitude of the corresponding correlation coefficient in the rotated factor matrix. This allows the identification of items having loadings on a given factor. The optimal number of factors to be selected in the final model was chosen by the scree pattern method.

The Cronbach’s α was measured to assess the degree of internal homogeneity of the sfGDS. Factor analysis was repeated on patients matched for cognitive performance (AMT score less than 7, n = 435) or physical impairment (dependency in at least 1 ADL, n = 780). Sensitivity, specificity, and diagnostic accuracy of the sfGDS versus a diagnosis of major depression made according to the Diagnostic and Statistical Manual of Mental Disorders (3rd ed., revised) criteria (n = 128) were computed.

**RESULTS**

Table 1 shows the general characteristics of patients. Most of them had been admitted to wards of geriatrics. Their mean age was 76 years, and 37.2% were aged 80 or older. The overall educational level was low: only 21.6% had an 8-year formal education or more. Cerebrovascular, cardiovascular, respiratory, and metabolic diseases were the most common diagnoses. Dependency, as reflected by need of assistance in at least 1 ADL, affected 32.1% of patients. Impairment in 1 or more IADL was recorded in 69.7% of patients. Twenty percent of patients scored greater than 7 on the AMT; this figure rose to 28.9% in patients older than 80 years.

Table 2 shows results from the main component analysis on the whole population. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.865; this value is consistent with the correlations between individual pairs of variables being explained by the other variables. The value of the Bartlett’s test of sphericity excludes that the correlation matrix for the 14 items is an identity matrix. Results from these 2 tests show that the main component analysis is appropriate to data. A 3-factor model was developed and could explain 47.7% of variance in GDS performance. The ratio between variances explained by the first and the second factor was 2.59, far from the cutoff of 3.5, which is considered to characterize a scale as unidimensional. Factor 1 had large loadings on 4 out of the 5 positive items, factor 2 on 5 negative items, and factor 3 on 1 positive and 4 negative items. Distressing thoughts and negative judgment about the own condition were the main areas related to factor 2, whereas inactivity and reduced self-esteem were related to factor 3. The item “worry about the future” remained unrelated to any factor.
Comparable results were obtained in the 2 strata with AMT < 7 or AMT > 6. The cumulative variance explained in these samples was 47% and 47.4%, and the best model was a 3-factor one with loadings of individual items fairly comparable to those observed in the general model.

In subjects dependent in at least 1 ADL, a more complex model was developed and is summarized in Table 3. It includes 5 main components explaining 28.9%, 11.7%, 7.9%, 7.3%, and 6.7% of the variance, respectively. Thus, the cumulative variance explained was 62.5%, but the relationship between factors and items was quite complex: while positive and most negative items had loadings with positive and negative factors, respectively, factor 3 was related to loss of usual activity (prefer to stay home, dropped activities/interests); factors 3 and 4 were weakly related to items assessing self-esteem or memory and overall conditions in comparison with those of the general population.

Table 4 shows the model for subjects independent in all ADLs. It included four factors and could explain 53.6% of the variance. The ratio of variance explained by the first factor to that explained by the second factor was 2.6. The loadings of factors on the items did not reflect an easily interpretable pattern of psychological dimensions.

The Cronbach’s α for the general model was .46, which indicates a low level of homogeneity among items of the sfGDS. Values for partial models ranged between .40 and .48. Sensitivity and specificity of sfGDS > 5 versus a first-listed or secondary diagnosis of major depression were 78.9% and 61.7%, respectively.

DISCUSSION

Our findings show that in broad hospitalized geriatric populations, the sfGDS explores 3 different psychological areas, has good sensitivity but suboptimal specificity versus a diagnosis of depression, can explain less than 50% of total variance, and lacks unidimensionality. The latter 2 findings limit the interpretation of the sfGDS because the final score may reflect in variable proportions factors unrelated to the state of mood. After stratification by presence of physical impairment, the sfGDS could explain a larger proportion of variance, probably because of the strong relationship between physical and affective status in the elderly. Indeed, physical disability was found to predict the onset of depression in subjects older than 64 years of age.

None of the previous studies assessed the construct validity of sfGDS on an unselected hospitalized geriatric population. Furthermore, the size of previously studied

| Table 1. General Characteristics of the Patients |
|-----------------|--------|--------|
| Demographics and functional status | n % | M ± SE |
| Age | 76.3 ± 0.18 |
| Gender (female) | 976 48 |
| Education (years) | 5.26 ± 0.09 |
| Living alone | 348 17.3 |
| Abbreviated Mental Test | 8.07 ± 0.04 |
| Geriatric Depression Scale-short form | 5.43 ± 0.08 |
| Length of stay (days) | 16.9 ± 3.7 |
| Activities of daily living (impaired in 1 or more) | 780 38.3 |
| Instrumental activities of daily living (impaired in 1 or more) | 1485 73.1 |
| Comorbid conditions | |
| Charlson’s index | 1.8 ± 0.04 |
| Chronic obstructive pulmonary disease | 350 17.2 |
| Cerebrovascular disease | 351 17.3 |
| Coronary artery disease | 595 29.3 |
| Congestive heart failure | 568 28 |
| Dementia | 86 4.2 |
| Depression | 113 5.6 |
| Diabetes | 419 20.6 |
| Hypertension | 862 42.4 |
| Acute myocardial infarction | 142 7 |
| Renal disease | 162 8 |
| Cancer | 163 8 |
| Stroke | 130 6.4 |
| Medication | |
| Antidepressant agents | 191 9.4 |
| Digoxin | 486 23.9 |
| Angiotensin-converting enzyme inhibitors | 199 9.8 |
| Diuretics | 499 24.6 |
| Calcium channel blockers | 364 17.9 |
| β-blockers | 54 2.7 |
| Corticosteroids | 202 9.9 |
| Admission ward | |
| Geriatrics | 1549 76.2 |
| Internal medicine | 483 23.8 |

| Table 2. Construct Validity of the Geriatric Depression Scale-Short Form in Older Medical Inpatients |
|--------------------------|--------|--------|--------|
| Component | 1 | 2 | 3 |
| Are you basically satisfied with your life? | 0.686 | -0.286 | -0.11 |
| Have you dropped many of your activities and interests? | -0.223 | 0.069 | 0.61 |
| Do you feel that your life is empty? | -0.434 | 0.578 | 0.13 |
| Do you often feel bored? | -0.355 | 0.489 | 0.239 |
| Are you in good spirits most of the time? | 0.729 | -0.147 | -0.138 |
| Are you afraid that something bad is going to happen to you? | -0.014 | 0.358 | 0.341 |
| Do you feel happy most of the time? | 0.775 | -0.097 | -0.199 |
| Do you often feel helpless? | -0.25 | 0.724 | 0.113 |
| Do you prefer to stay at home rather than going out and doing new things? | -0.02 | 0.197 | 0.58 |
| Do you feel you have more problems with memory than most? | 0.153 | 0.509 | 0.199 |
| Do you think it is wonderful to be alive now? | 0.731 | -0.016 | -0.051 |
| Do you feel pretty worthless the way you are now? | -0.266 | 0.282 | 0.417 |
| Do you feel full of energy? | 0.443 | 0.309 | -0.589 |
| Do you feel your situation is hopeless? | -0.222 | 0.72 | 0.054 |
| Do you think that most people are better off than you are? | 0.087 | 0.365 | 0.499 |
| Variance explained by individual components (%) | 28.903 | 11.154 | 7.681 |
| Total variance explained (cumulative variance, %) | 47.738 |

Loadings of identified components on individual items and explained variance are reported.
samples ranged between 72 and 285 subjects. The present findings add to the knowledge on sfGDS by presenting results generalizable to the older medical inpatients. The main limit of sfGDS lies in its doubtful meaning: its internal structure does not guarantee that the final score of sfGDS measures what it is assumed to measure, that is, the state of mood. Examples of construct validity for the Mini-Mental State Examination (MMSE), the test most commonly used for screening dementia, help to clarify this issue: explaining a cumulative variance of 68% and being unidimensional, MMSE really expresses what it measures in community-dwelling older adults. The 56.1% explained variance and the clustering of individual items into well-defined dimensions of neurocognitive ability indicate that MMSE can effectively screen for cognitive impairment in older nursing home residents, but the final score should be interpreted with some caution because of the lack of unidimensionality.

The complex psychological status of older patients experiencing both an acute illness and the hospital stay likely limits the possibility of a reliable screening of depression. Although we studied our patients immediately before discharge, it is unlikely that this limitation was completely overcome. We ignore whether the same limitation affects the other scales commonly used for screening depression in older medical inpatients because their construct validity has been assessed in samples that are small and/or poorly representative of this broad population.
tia in the general geriatric population and between suicidal ideas and nonsuicidal death in medical older inpatients are examples of such a possibility.27,28

Multidimensional assessment is the cornerstone of the current approach to the elderly patient.29 Assessing social, medical, neuropsychological, and functional status allows for complete information on the health status of the patient and the planning of interventions aimed at improving that status or preventing further worsening. Multidimensional assessment is based on diagnostic instruments that should require little time and be easy to use, reproducible, sensible to small changes, and related to major outcomes such as disability and mortality. Several instruments exploring functional and mental status meet these requirements, while the same is not true of tests screening older medical inpatients for depression. The prominent role and the poor specificity of somatic symptoms in geriatric depression could partly explain this finding.30 Furthermore, visual problems, illiteracy, and, mainly, lack of motivation frequently prevent older medical patients from completing the self-rated depression scale.31 Theoretically, lack of motivation could also affect the quality of self-rating. The important association between depression and cognitive impairment in older inpatients might further affect the quality of self-rating.32 Accordingly, the observed limitations of sfGDS might be common to other instruments based on self-rating. A comparison between self-rated and observer-rated depressive status in older medical patients could help clarify this issue.

In conclusion, analysis of construct validity of sfGDS disclosed a well-defined and logical set of component factors but failed to demonstrate a structure that could guarantee reliability of results. It is likely that selecting out the 15 items from the parent GDS according to findings obtained from a home-dwelling older population resulted in the development of a scale only partially fitting the needs of older medical inpatients.3 Thus, future efforts should be finalized to develop an instrument specifically designed for this very large population to assess both prevalence and implications of depression as well as its responsiveness to therapeutic interventions.

CLINICAL IMPLICATIONS

- sfGDS cannot be considered an optimal instrument for screening depression in unselected older medical inpatients.
- Cognitive performance, as assessed by the AMT, does not affect the internal structure of sfGDS.
- sfGDS performs better in patients matched for level of physical capabilities.

LIMITATIONS

- The lack of a symptom-rating instrument prevented us from distinguishing the effects of symptoms per se from those of disability on the internal structure of sfGDS.
- The data are taken from Italian hospitals, and the conclusions cannot be generalized to other settings.
- Men outnumbered women in our population. SFGDS might have a different internal structure in the presence of a female to male ratio similar to that of the older general population.

References


Delirium is a syndrome of disturbed consciousness, cognition, and perception that develops over a short period of time and tends to fluctuate during the course of the day, and it is caused by 1 or more physical conditions. Although delirium is significantly common among people of all ages, older adults are more prone to develop this syndrome because they are at higher risk for underlying brain disease. The prevalence of delirium has been estimated from 10% to 40% among hospitalized elderly patients on medical and surgical wards. Also, hospital mortality estimates range from 10% to 65%.

Delirium is traditionally considered a transient syndrome that ends in recovery after several days to weeks in most cases. This notion tends to lead most clinicians’ psychopharmacological interventions to a brief one as long as the underlying causes of delirium are identified and resolved. Interestingly, Levkoff and her colleagues in their prospective study demonstrated that many patients still had 1 or more symptoms as long as 6 months after hospital discharge. Older adults with delirium can present with a wide variety of neuropsychiatric symptoms. The principal treatment of delirium is still the diagnosis and treatment of the underlying physical conditions contributing to delirium. However, psychopharmacological intervention is a major component of all interventions for delirium. Typically, high-potency neuroleptic agents such as haloperidol have been used as first-line treatment for neuropsychiatric symptoms of patients with delirium. However, they are frequently associated with extrapyramidal symptoms (EPS), particularly in older adults. Recently, there have been several reports of use of the atypical antipsychotic agents, risperidone and olanzapine.

Quetiapine is an atypical neuroleptic agent and dibenzothiazepine derivative structurally related to clozapine and olanzapine. Quetiapine is well tolerated and associated with improvement in psychotic symptoms, although it demonstrates lack of EPS and minimal sedative, hypotensive, and anticholinergic side effects in the dose range used in older adults. This study was designed to determine the efficacy and safety of open-label quetiapine treatment in patients with delirium over a period of 3 months.

METHODS

All patients were enrolled from the acute medical units of Salem Veterans Affairs Medical Center, a 271-bed, general and teaching hospital for the University of Virginia, School of Medicine. The hospital's Human Studies Subcommittee...
and Research and Development Committee approved the study protocol. Informed consent was required for participation in this study. All patients were referred to Consultation-Liaison Psychiatry Service for evaluation of mental status changes. Patients who had known histories of psychotic disorders and were treated with neuroleptic agents within the previous 4 weeks prior to the enrollment were excluded from this study. They were required to be at least 60 years old. Patients provided written informed consent before admission to the study. When they did not have a capacity to consent due to significant changes in mental status, the consent was sought from their next of kin. All patients enrolled in this study met Diagnostic and Statistical Manual of Mental Disorders (4th ed.) criteria for delirium. They were administered the Mini-Mental State Examination (MMSE), Delirium Rating Scale (DRS), Clock Drawing Test (CDT), and Clinical Global Impression (CGI). At their admission to the study (T0), they were administered the MMSE, the DRS, and the CGI. When an attending psychiatrist determined that patients were in need of antipsychotic treatment based on clinical grounds, they were given a starting dose of quetiapine 25 mg twice a day. The dosages were then increased by 25 mg every 2 days until patients were maximally stabilized (T1). When patients required an adjunctive psychotropic therapy for acute symptoms, they were given oxazepam or lorazepam by mouth or by intramuscular injection as needed. At stabilization, patients were again given the above-mentioned measures and discharged on the stabilizing dose of quetiapine. All patients had follow-up visits at the first month of therapy (T1) and third month of therapy (T3). At the first-month visit, the quetiapine was tapered off by 25 mg every 3 days when they were considered stable. At each follow-up visit, they were given the same measurements. The efficacy of quetiapine was evaluated using responder analyses and paired t tests. Side effects were assessed with clinically oriented, open-ended questions.

**RESULTS**

Eleven of the 12 patients completed the study to T3. One patient was started on quetiapine and followed at T1, but he died of acute myocardial infarction a few days later. This patient had preexisting cardiac conditions, and his death was unrelated to the study drug. All 12 patients were considered evaluable (Table 1). There were no dropouts due to side effects. The age (mean ± SD) of patients was 74 ± 7 years, and all of them were male. None of the 12 patients had prior history of psychiatric treatment. Two patients had a diagnosis of early dementia, but they had no behavioral symptoms prior to the development of delirium. Only 1 patient required a one-time dose of lorazepam (1 mg) administered intramuscularly. The mean ± SD dose of quetiapine at T1 was 93.75 ± 23.31 mg/day. The mean duration for stabilization at T1 was 5.91 ± 2.22 days. None of the 12 patients developed EPS, and rates of other side effects were considered minimal (sedation in 2 patients and vivid dreams in 1).

### Table 1. Clinical Characteristics of Patients

<table>
<thead>
<tr>
<th>Patient Number</th>
<th>Age (years)</th>
<th>Gender</th>
<th>Medical Diagnosis</th>
<th>Dose (mg) at T1</th>
<th>Time to Stabilization (days)</th>
<th>Adverse Reactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>80</td>
<td>Male</td>
<td>DM, CRI*, hypertension, hypothyroidism, dementia</td>
<td>50 mg bid</td>
<td>6</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>64</td>
<td>Male</td>
<td>UTI, delirium tremens*</td>
<td>75 mg bid</td>
<td>8</td>
<td>None</td>
</tr>
<tr>
<td>3</td>
<td>68</td>
<td>Male</td>
<td>S/P laryngectomy, hypertension, hypothyroidism, UTI*</td>
<td>50 mg bid</td>
<td>5</td>
<td>None</td>
</tr>
<tr>
<td>4</td>
<td>70</td>
<td>Male</td>
<td>DM, S/P hip hemiarthroplasty*, BPH</td>
<td>50 mg bid</td>
<td>10</td>
<td>None</td>
</tr>
<tr>
<td>5</td>
<td>82</td>
<td>Male</td>
<td>CRI*, CHF, hypertension, DM, BPH</td>
<td>25 mg qam</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>6</td>
<td>79</td>
<td>Male</td>
<td>GI bleeding*, DM, chronic anemia, dementia</td>
<td>75 mg ghs</td>
<td>4</td>
<td>None</td>
</tr>
<tr>
<td>7</td>
<td>75</td>
<td>Male</td>
<td>CHF*, hypertension, BPH, sleep apnea</td>
<td>50 mg bid</td>
<td>6</td>
<td>None</td>
</tr>
<tr>
<td>8</td>
<td>67</td>
<td>Male</td>
<td>S/P coronary bypass*, hypertension, DM, chronic anemia</td>
<td>50 mg bid</td>
<td>8</td>
<td>None</td>
</tr>
<tr>
<td>9</td>
<td>71</td>
<td>Male</td>
<td>MI*, cardiomyopathy, hypertension, DM</td>
<td>25 mg bid</td>
<td>4</td>
<td>None</td>
</tr>
<tr>
<td>10</td>
<td>77</td>
<td>Male</td>
<td>UTI, dehydration*</td>
<td>25 mg bid</td>
<td>3</td>
<td>None</td>
</tr>
<tr>
<td>11</td>
<td>74</td>
<td>Male</td>
<td>Pneumonia*, PE, duodenal ulcer, esophageal stricture</td>
<td>25 mg qam, 50 mg ghs</td>
<td>4</td>
<td>None</td>
</tr>
<tr>
<td>12</td>
<td>88</td>
<td>Male</td>
<td>GI bleeding*, glaucoma, BPH</td>
<td>100 mg ghs</td>
<td>5</td>
<td>Drowsiness</td>
</tr>
</tbody>
</table>

DM = diabetes mellitus; CRI = chronic renal insufficiency; UTI = urinary tract infection; BPH = benign prostatic hypertrophy; GI = gastrointestinal; MI = myocardial infarction; PE = pulmonary embolism.

*Main cause of delirium.
haloperidol. In addition, it was of clinical interest that to what has been reported for olanzapine and patients developed EPS. Peak response times were comparable to what has been reported for olanzapine and haloperidol. In addition, it was of clinical interest that the main measure in this study, the DRS, continued to improve across time periods. This finding might have some implication for the duration of intervention, especially pharmacotherapeutic treatment. In our study, we found that quetiapine was a safe and effective treatment in hospitalized older patients with delirium. However, it will be difficult to absolutely determine whether the neuropsychiatric symptoms improved because of quetiapine, resolution of the underlying medical conditions, or a combination of the two. Larger controlled studies are needed to further explore these preliminary findings and conclusions.

**DISCUSSION**

Since this was a small, open-label study with no control group, our findings are certainly limited for generalization. This study was also limited in that all subjects were male. All patients had multiple physical disorders. They were also taking various drugs for their medical conditions. However, all patients were free of previous psychiatric diagnosis and were not on any type of psychotropic agents for at least 4 weeks prior to admission to this study; thus, it was an appropriate group of patients for a psychotropic drug study.

Quetiapine was well tolerated by all patients in this study. Vivid dreams and sedation were reported, but no patients developed EPS. Peak response times were comparable to what has been reported for olanzapine and haloperidol. In addition, it was of clinical interest that the main measure in this study, the DRS, continued to improve across time periods. This finding might have some

<table>
<thead>
<tr>
<th>Measure</th>
<th>$T_0$</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$T_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>14.50</td>
<td>21.17</td>
<td>24.27</td>
<td>25.18</td>
</tr>
<tr>
<td>SD</td>
<td>5.90</td>
<td>4.55</td>
<td>4.86</td>
<td>4.45</td>
</tr>
<tr>
<td>N</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>$t^a$</td>
<td>4.07</td>
<td>3.96</td>
<td>1.99</td>
<td></td>
</tr>
<tr>
<td>$t^b$</td>
<td></td>
<td>1.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t^c$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>.002</td>
<td>.003</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>DRS</td>
<td>18.25</td>
<td>8.00</td>
<td>2.27</td>
<td>0.63</td>
</tr>
<tr>
<td>SD</td>
<td>6.05</td>
<td>2.34</td>
<td>2.28</td>
<td>1.21</td>
</tr>
<tr>
<td>N</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>$t^a$</td>
<td>5.60</td>
<td>7.88</td>
<td>2.57</td>
<td></td>
</tr>
<tr>
<td>$t^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t^c$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>.0002</td>
<td>.0001</td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>CDT</td>
<td>3.25</td>
<td>6.25</td>
<td>7.91</td>
<td>8.09</td>
</tr>
<tr>
<td>SD</td>
<td>2.77</td>
<td>4.07</td>
<td>3.45</td>
<td>3.21</td>
</tr>
<tr>
<td>N</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>$t^a$</td>
<td>3.35</td>
<td>3.46</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>$t^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t^c$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>.007</td>
<td>.006</td>
<td>ns</td>
<td></td>
</tr>
<tr>
<td>CGI-S</td>
<td>3.00</td>
<td>1.91</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>0.43</td>
<td>0.83</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>$t^a$</td>
<td>-6.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t^c$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>.0004</td>
<td>.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MMSE = Mini-Mental State Examination; DRS = Delirium Rating Scale; CDT = Clock Drawing Test; CGI-S = Clinical Global Impressions–Severity.

a. Paired $t$ test between $T_0$ and $T_1$
b. Paired $t$ test between $T_1$ and $T_2$
c. Paired $t$ test between $T_2$ and $T_3$

**References**


Delirium has been estimated to occur in 14% to 56% of hospitalized elderly patients. Prior studies have shown that delirium may be associated with prolonged hospital lengths of stay, increased mortality, functional and cognitive decline, and increased rates of nursing home placement. Inouye et al have noted that delirium may serve as a “window” to the hospital care of elderly patients as it is frequently iatrogenic, closely linked to the processes of care, and, as such, serves as a useful outcome measure. The above-mentioned rates of delirium and negative associated outcomes generally have been found in prospective studies in which investigators were actively assessing for the presence of delirium. Other studies have consistently noted a different problem associated with delirium in clinical settings: that of its underrecognition and misidentification. It has been estimated that clinicians fail to recognize delirium one third to two thirds of the time. Various reasons have been suggested for the underrecognition of delirium: the use of different diagnostic terms; varying clinical presentations of delirium due to fluctuation or hypoactive forms of the syndrome; the attribution of cognitive changes to dementia, depression, or senescence; the lack of experience of health practitioners in testing for attentional disturbances; the absence of baseline cognitive assessments; and the failure to appreciate the significance of delirium as a marker for severe illness and mortality.

Studies involving record review at single institutions have found low rates of provider recording of delirium. Also, although “delirium” is the gold standard term in clinical and research use, providers may still use less precise, vague descriptions of symptoms such as “confusion” or “agitation” or alternative diagnostic terms to describe delirium such as “acute brain syndrome,” “toxic psychosis,” and “metabolic encephalopathy.” A study comparing the prospective diagnosis of delirium by a psychiatrist with evidence for delirium in the medical record found that only 17% of the delirious patients could be recognized as delirious by record review (International Classification of Dis-
eases, 9th revision [ICD-9] codes, admission notes, and progress notes). Identifying cases through chart review at the Ann Arbor Veterans Medical Center, Kamholz et al found delirium in approximately 20% of hospitalized elderly patients using the Confusion Assessment Method. The records of these patients with delirium were then searched for any provider notations indicative of the recognition of delirium. Only 45% of the patients with delirium on medical units had any such notations made by physicians in the chart. Far fewer (18%) surgical patients with delirium had any such notations recorded.

The recognition of delirium may be particularly affected by both the fluctuation of delirium symptoms and the appearance of hypoactive or “quiet” delirium. Physicians may see inpatients during morning rounds when behavior may be less disturbing than in the afternoon or night. Nursing documentation of delirium is more frequent, likely due to nurses’ closer contact with patients throughout the day, enabling them to spot delirium despite fluctuation. In terms of motoric subtypes, O’Keefe and Lavan have noted that it is the patients with agitated delirium who are most likely to attract medical and nursing attention, while patients with “quiet” delirium may appear to be compliant “model” patients.

Delirium is a syndrome associated with a multitude of possible risk factors including advanced age, chronic cognitive impairment, severe medical comorbidity, and functional impairment. It may be these underlying risk factors (as opposed to the delirium itself) that may account for the poor outcomes associated with delirium. These risk factors or comorbidities may also affect detection. Elie et al found that elderly emergency room patients with detected delirium were more likely to have a primary neurologic diagnosis than those with undetected delirium. The authors speculated that neurologic disorders may have led emergency room physicians to pay more attention to mental status. Notably, however, another study noted that delirium was less likely to be recognized by nurses and physicians in hospitalized elderly patients with dementia.

Therefore, there are factors including providers’ choice of diagnostic terms, behavioral manifestations of delirium, and associated comorbidities that may affect the detection and recording of delirium. No studies to our knowledge have examined the rates of recorded delirium and associated characteristics within a large health care system. As Gustafson et al have noted, if delirium is “not properly diagnosed and documented, subsequent lack of appropriate treatment and care is expected.” The Veterans Affairs (VA) health care system, the largest single provider of health care in the United States, serves a large clinical population of elderly veterans. Prior research has indicated that the VA patient population has poorer health status and more medical conditions than non-VA patient populations, related in part to older age of the veteran patients. Thus, the veteran patient population might be considered at higher risk for developing delirium. A detailed examination of the rates and characteristics of recorded delirium and related diagnoses in the VA system is highly relevant to the study of delirium; further data regarding the demographics of recognized delirium in a large health care system may give clues as to the types of delirium that may go undetected as well as suggesting the possible extent of underrecognition.

Given our awareness of past prospective prevalence estimates, as well as prior studies demonstrating poor physician recognition and recording of delirium in the medical record, we hypothesized that the rate of recorded delirium in the VA system would be lower than the prevalence noted in prospective studies. We also suspected that patients recorded as having delirium would be older and more likely to have dementia than elderly patients without recorded delirium.

METHODS

Data were obtained retrospectively from 2 sources, both national computerized abstracts of VA administrative data. The Patient Treatment File contains abstracts of discharge records for each episode of care in all VA medical centers and extended care facilities across the United States by fiscal year. The abstracts include patient demographic characteristics such as age, race, and gender and episode of care summary information such as admission date, discharge date, and discharge diagnoses. The latter include primary diagnoses responsible for length of stay and as many as 9 accompanying secondary diagnoses. The Annual Patient Census files contain analogous abstracts of records to date for all patients residing in VA hospitals and extended care facilities at the end of each fiscal year.

The sample for our study was defined as all veterans discharged from a VA medical center or extended care facility in FY96 (October 1, 1995, to September 30, 1996) who were 60 years of age or older at discharge and who had any one of the following (International Classification of Diseases, 9th revision [ICD-9CM]) delirium diagnoses recorded for at least 1 episode of care: acute delirium (293.0), subacute delirium (293.1), drug-induced delirium (292.0, 292.81), alcohol use or withdrawal delirium (291.0), senile dementia with delirium (290.3), arteriosclerotic dementia with delirium (290.41), and presenile dementia with delirium (290.11). As alternative terms for delirium might also be used by physicians such as “confusion,” “encephalopathy,” “organic brain syndrome,” or “organic psychosis,” ICD-9CM was also searched for these and any like terms. The FY96 entire sample was then examined for any of these related diagnoses recorded, which included acute encephalopathy (348.3); toxic encephalopathy (349.82); hypertensive encephalopathy (437.2); hepatic encephalopathy (572.2); drug-induced paranoid/hallucinatory state (292.1); pathological drug intoxication resulting in brief psychotic states (292.2); unspecified drug-induced mental disorder, subcoded as organic psychosis not otherwise specified (292.9); organic delusional syndrome, subcoded...
as transient organic psychotic condition, paranoid type (293.81); organic hallucinosis syndrome, subcoded as transient organic psychotic condition, hallucinatory type (293.82); organic affective syndrome, subcoded as transient organic psychotic condition, depressive type (293.83); other specified transient organic mental disorder (293.89); unspecified transient organic mental disorder, subcoded as organic psychosis (293.9); other specified organic brain syndromes (chronic) (294.8, includes mixed paranoid and affective organic psychotic states); and unspecified organic brain syndrome (chronic), subcoded as organic psychosis (294.9).

The overall rate of delirium diagnoses within the VA inpatient system for FY96 was calculated, as well as the rates of various delirium types and related confusional diagnoses recorded. Demographic variables included age at discharge from index hospitalization, gender, and race. Because of small numbers, Hispanic white, Hispanic black, American Indian, and Asian patients were grouped together as “other.” Length of stay at index hospitalization was computed as the number of days from admission to discharge, less days on leave. Other variables examined included index location, discharge location, index mortality, and medical comorbidity. The percentage of patients with dementia in each group was calculated by searching for the following dementia codes: 331.0 (Alzheimer’s disease), 331.1 (Pick’s disease), 290.0 (senile dementia, uncomplicated), 290.10-3 (presenile dementia), 290.20-1 (senile dementia), 290.3 (senile dementia with delirium), 290.40-3 (arteriosclerotic dementia), 291.2 (alcoholic dementia), 292.82 (drug-induced dementia), and 294.1 (dementia in conditions classified elsewhere).

Patients without recorded delirium and without related confusional diagnoses constituted the “no recorded confusion” group. As the results for recorded cases of delirium and encephalopathy cases were similar in their characteristics, these cases were added together to form the “recorded delirium” group. The organic psychoses and organic brain syndrome cases were significantly different enough in their characteristics from the “recorded delirium” cases that they were grouped into a third group, “other confusional diagnoses.” There were 3 patients with both recorded delirium and other confusional diagnoses; these patients were grouped under recorded delirium. The 3 groups, no recorded confusion, recorded delirium, and other recorded confusional diagnoses, were then compared for the demographic and utilization variables; categorical differences were assessed using chi-square tests of independence, and continuous variables were compared by 3-group 1-way univariate analyses of variance (ANOVA).

RESULTS

There were a total of 267,947 patients older than the age of 60 admitted to VA inpatient units nationally in 1996. Of these, a total of 3978, or 1.5%, of elderly patients had recorded delirium or encephalopathy diagnoses that constituted the recorded delirium group. An additional 6680, or 2.5%, had organic psychosis diagnoses that made up the other recorded confusional diagnoses group. Therefore, 4% (10,658 of the 267,947) of elderly patients admitted in 1996 had a delirium or a related confusional diagnosis recorded.

Table 1 lists demographics, utilization rates, and mortality for the 3 groups: (1) no recorded confusion, (2) recorded delirium, and (3) other recorded confusional diagnoses. Patients with no recorded confusion were significantly younger than those with either recorded delirium or other recorded confusional diagnoses. Patients with other recorded confusional diagnoses were also significantly older than the recorded delirium group. There were no gender differences found among the 3 groups. Significant racial differences were found among the 3 study groups. The other recorded confusional diagnoses group had a significantly higher percentage of African Americans (19%) than did the recorded delirium or no recorded confusion groups.

In terms of health care utilization patterns, the groups with recorded delirium and other recorded confusional diagnoses were significantly more likely to be admitted to psychiatry and nursing home units (and less likely to be admitted to medical or surgical units) than were patients with no recorded confusion. Patients with other recorded confusional diagnoses had the highest rates of admission to psychiatry and nursing home units. The groups with recorded delirium and other recorded confusional diagnoses were also significantly more likely to be discharged to nursing homes as compared to those with no recorded confusion. The group with other recorded confusional diagnoses had the highest rate of nursing home discharges (23.2%) as compared to patients with no recorded confusion (5.6%) or recorded delirium (17.1%).

Patients with recorded delirium had slightly shorter lengths of stay than did those with no recorded confusion, but this difference was not significant. Patients with other recorded confusional diagnoses had significantly longer lengths of stay on index admission than either the no recorded confusion or recorded delirium groups. Conversely, those with recorded delirium had the highest mortality rates during index hospitalization (13.3%); these mortality rates were significantly higher than the no recorded confusion (5.2%) or other recorded confusional diagnoses (8.7%) groups.

Types of Delirium:
Related Confusional Diagnoses Recorded
Table 2 lists the types of delirium noted among patients recorded as having delirium or encephalopathy diagnoses. Among the recorded delirium group, the highest rates were found for dementia with delirium (30.1%) and alcohol intoxication/withdrawal delirium (18.7%). Among patients having dementia with delirium recorded, demen-
Dementia or Other Medical Comorbidities

Thirty percent ($n = 1197$) of those with recorded delirium also had a recorded dementia diagnosis, significantly higher than the other 2 groups. Among hospitalized elderly patients with no recorded confusion ($n = 257,289$), 5.4% ($n = 13,828$) had a recorded dementia diagnosis. The lowest percentage of recorded dementia (4.9%, $n = 326$) was found among patients with other recorded confusional diagnoses.

Other medical comorbidities recorded were analyzed for all patient groups. The top 3 comorbidities recorded did not differ between groups and were (1) diseases of the circulatory system (390-459.9), (2) infectious diseases (001-139.8), and (3) endocrine, nutritional and metabolic, and immunity disorders (240-279.9).

**DISCUSSION**

The current study was undertaken to obtain a snapshot of the face of recorded delirium and related diagnoses within a large national health care system, the VA health care system, to see what characteristics might be associated with recorded delirium. The overall rate of the recorded diagnosis of delirium and very closely related diagnoses among all hospitalized elderly veterans was found to be only 1.5%. More common were other confusional diagnoses, recorded in another 2.5% of all elderly inpatients.

Among hospitalized elderly patients with no recorded confusion ($n = 257,289$), 5.4% ($n = 13,828$) had a recorded dementia diagnosis. The lowest percentage of recorded dementia (4.9%, $n = 326$) was found among patients with other recorded confusional diagnoses.

Other medical comorbidities recorded were analyzed for all patient groups. The top 3 comorbidities recorded did not differ between groups and were (1) diseases of the circulatory system (390-459.9), (2) infectious diseases (001-139.8), and (3) endocrine, nutritional and metabolic, and immunity disorders (240-279.9).

**Table 1. Demographic and Utilization Variables for the Study Groups**

<table>
<thead>
<tr>
<th></th>
<th>Group 1: No Recorded Confusion ($n = 257,289$)</th>
<th>Group 2: Recorded Delirium ($n = 3978$)</th>
<th>Group 3: Other Recorded Confusional Diagnoses ($n = 6680$)</th>
<th>Significance Test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years, mean [SD])</td>
<td>71.4 (6.8)</td>
<td>72.0 ($a$)</td>
<td>75.5 ($a,b$)</td>
<td>$F_{(2,267,944)}^\ast = 1186.22$; $\chi^2_{(4)} = 108.25$</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>198,520 (77.2)</td>
<td>3,086 (77.6)</td>
<td>5,088 (76.2)</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>40,458 (15.7)</td>
<td>605 (15.2)</td>
<td>1,276 (19.1)</td>
<td></td>
</tr>
<tr>
<td>All other</td>
<td>18,311 (7.1)</td>
<td>235 (5.9)</td>
<td>316 (4.7)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>0 (0.0)</td>
<td>52 (1.3)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>5,356 (2.1)</td>
<td>70 (1.8)</td>
<td>137 (2.1)</td>
<td>$\chi^2_{(4)} = 4706.47$</td>
</tr>
<tr>
<td><strong>Utilization</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index location</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical/surgery</td>
<td>233,474 (90.7)</td>
<td>3,238 (81.4)</td>
<td>4,444 (66.5)</td>
<td></td>
</tr>
<tr>
<td>Psychiatry</td>
<td>11,644 (4.5)</td>
<td>464 (11.7)</td>
<td>1,028 (15.4)</td>
<td></td>
</tr>
<tr>
<td>Nursing home</td>
<td>12,171 (4.7)</td>
<td>276 (6.9)</td>
<td>1,208 (18.1)</td>
<td></td>
</tr>
<tr>
<td>Discharge location</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community</td>
<td>214,717 (83.4)</td>
<td>2,474 (62.2)</td>
<td>3,893 (58.3)</td>
<td></td>
</tr>
<tr>
<td>Nursing home</td>
<td>14,344 (5.6)</td>
<td>681 (17.1)</td>
<td>1,549 (23.2)</td>
<td></td>
</tr>
<tr>
<td>Death (index)</td>
<td>13,347 (5.2)</td>
<td>530 (13.3)</td>
<td>579 (8.7)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>14,881 (5.8)</td>
<td>293 (7.4)</td>
<td>659 (9.9)</td>
<td></td>
</tr>
<tr>
<td>Index length of stay (days, mean [SD])</td>
<td>44.4 (288.5)</td>
<td>42.2 (171.9)</td>
<td>150.5 ($a,b$) (485.7)</td>
<td>$F_{(2,267,944)}^\ast = 426.02$</td>
</tr>
</tbody>
</table>

Significant post hoc (Tukey-Kramer’s) comparisons ($P < .05$): $\ast$significantly greater than group 1; $\ast\ast$significantly greater than group 2.

$P < .0001$.

**Table 2. Types of Delirium Recorded ($n = 3978$)**

<table>
<thead>
<tr>
<th>Diagnosis (ICD-9CM Code)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dementia with delirium (290.3, 290.41, 290.11, or dementia code plus delirium or encephalopathy code)</td>
<td>1197</td>
<td>30.1</td>
</tr>
<tr>
<td>Alcohol intoxication/withdrawal delirium (291.0)</td>
<td>743</td>
<td>18.7</td>
</tr>
<tr>
<td>Acute/subacute delirium (293.0, 293.1)</td>
<td>701</td>
<td>17.6</td>
</tr>
<tr>
<td>Hepatic encephalopathy (572.2)</td>
<td>560</td>
<td>14.1</td>
</tr>
<tr>
<td>Drug-induced delirium (292.0, 292.81)</td>
<td>492</td>
<td>12.4</td>
</tr>
<tr>
<td>Acute encephalopathy (348.3)</td>
<td>172</td>
<td>4.3</td>
</tr>
<tr>
<td>Hypertensive encephalopathy (437.2)</td>
<td>89</td>
<td>2.2</td>
</tr>
<tr>
<td>Toxic encephalopathy (349.82)</td>
<td>17</td>
<td>0.4</td>
</tr>
<tr>
<td>Combination diagnoses (349.83, 290.2/293.81, 290.2/293.83, 572.2/347.2, 572.2/292.0, 572.2/349.83, 290.2)</td>
<td>7</td>
<td>0.2</td>
</tr>
</tbody>
</table>
tions, and underreporting (and thus underrecording) of delirium, has been estimated to occur in 14% of hospitalized elderly patients.1 Our results are likely both a reflection of the consistent underrecognition of true prevalence. In previous prospective studies, hospitalized veterans, 4%, is most certainly an underestimation of recorded delirium and related diagnoses found in the VA national health care system.

The finding of 53% of study patients with delirium/other related confusional diagnoses being assigned a diagnosis of chronic organic brain syndrome is puzzling. Interestingly, while patients with other recorded confusional diagnoses were older and had longer lengths of stay than did the patients with delirium diagnoses, their mortality rates were significantly lower. Based on their demographic and utilization profile, we suspected that patients with other recorded confusional diagnoses would have higher rates of recorded dementia than the other 2 groups, but surprisingly, only 5% of these patients had such diagnoses recorded. These findings lead us to suspect that the group of patients who received other recorded confusional diagnoses is likely heterogeneous, containing patients with chronic psychiatric diagnoses as well as undiagnosed dementia. Another indication of this heterogeneity is possibly found in the large standard deviations in length of stay found for all groups (but largest for the group with other recorded confusional diagnoses); we suspect this reflects the tremendous variability in hospital and nursing home stays of this elderly patient population.

The racial differences found between groups were particularly striking: almost 20% of patients receiving other recorded confusional diagnoses or organic psychoses were African American. While race was missing in 52 cases in the recorded delirium group, even had all 52 patients been African American, the percentage of African Americans would still have been only 16.5% in this group. One explanation for the racial differences found may pertain to medical comorbidities with high prevalence rates in African Americans such as chronic renal failure, as encephalopathy associated with renal dialysis is one diagnosis coded as chronic organic brain syndrome. While African Americans account for about 13% of the population, they account for 29% of new cases of end-stage renal disease and 37% of those undergoing dialysis.34 However, based on the data available to us, we were unable to show different patterns of medical comorbidity between groups. Another explanation might be found in the findings of prior studies including our own, that is, that elderly African American patients are significantly more likely to receive clinical diagnoses of psychotic disorders than are Caucasian patients.35-37 In a previous national study of elderly veterans, we found that elderly African Americans were significantly more likely to receive psychotic, cognitive, and substance abuse disorder diagnoses and significantly less likely to receive mood disorder diagnoses.38 In that study, we hypothesized that differential rates of recognition/diagnosis of depression by race might pertain to both patient (differential symptom presentation) and provider (conscious or unconscious bias) factors. This could be the case with delirium as well. All of the diagnoses contained in the other recorded confusional diagnoses category include organic psychosis in either the coding or subcoding in ICD-9CM; it is possible that in some cases, psychotic symptoms in the context of a delirium syndrome are coded preferentially as organic psychosis in African Americans as opposed to actual delirium diagnoses.

In terms of patients given actual delirium or encephalopathy diagnoses, the most common type of delirium recorded was delirium with dementia; 30% of the delirium group also had dementia recorded. This finding could reflect both the increased risk for delirium in dementia, as dementia itself is the most common risk factor for delirium in the elderly with the possibility of almost any physical illness precipitating delirium in the elderly when there exists an underlying cerebral disease such as Alzheimer's or vascular dementia.38,39 The finding may

### Table 3. Other Recorded Confusional Diagnoses (n = 6680)

<table>
<thead>
<tr>
<th>Diagnosis (ICD-9CM Code)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic organic brain syndrome (294.8)</td>
<td>5616</td>
<td>84.1</td>
</tr>
<tr>
<td>Chronic organic psychosis (294.9)</td>
<td>367</td>
<td>5.5</td>
</tr>
<tr>
<td>Specified transient organic psychosis (293.8)</td>
<td>327</td>
<td>4.9</td>
</tr>
<tr>
<td>Unspecified transient organic psychosis (293.9)</td>
<td>189</td>
<td>2.8</td>
</tr>
<tr>
<td>Drug-induced paranoid/hallucinatory state (292.1)</td>
<td>88</td>
<td>1.3</td>
</tr>
<tr>
<td>Combination diagnoses (292.1/292.9, 293.9/292.9, 293.9/293.83, 294.8/293.83, 294.8/293.89, 294.8/293.81, 294.8/293.82, 294.8/293.9, 294.8/292.9, 294.8/294.9, 294.8/292.1, 294.8/292.82, 294.8/293.9)</td>
<td>48</td>
<td>0.7</td>
</tr>
<tr>
<td>Organic psychosis not otherwise specified related to drugs (293.2)</td>
<td>39</td>
<td>0.6</td>
</tr>
<tr>
<td>Pathological drug intoxication (brief psychotic state)</td>
<td>6</td>
<td>0.1</td>
</tr>
</tbody>
</table>
also point toward the greater detection of delirium in dementia. Elie et al.\(^4\) found that patients in their study with detected delirium were more likely to have a primary neurologic diagnosis than were those with undetected delirium. The authors hypothesized that neurologic disorders may lead physicians to pay more attention to mental status; however, Fick and Foreman\(^27\) found that delirium was less likely to be recognized by nurses and physicians in hospitalized elderly patients with dementia.

The second highest recorded delirium rate among patients with delirium/encephalopathy was for alcohol intoxication/withdrawal delirium. Age is a potential predictor of alcohol withdrawal severity and complications; Kraemer et al.\(^40\) found that 14% of elderly patients admitted to an alcohol detoxification unit had delirium during alcohol withdrawal. Prior studies have found that elderly patients had significantly more alcohol withdrawal symptoms for a longer duration than did younger patients.\(^31,42\)

Delirium has been phenomenologically divided into 3 different motoric subtypes according to the presence of psychomotor and behavioral symptoms: hyperactive, hypoactive, and mixed.\(^38\) Prior studies have suggested that the hyperactive delirium subtype may be associated with better outcomes\(^43\) including lower mortality rates\(^44\) and etiologically related to drug intoxication/withdrawal.\(^45\) Hypoactive delirium has been associated both with poorer outcomes\(^21\) including higher mortality rates\(^43,44\) as well as greater underrecognition.\(^46,47\) Delirium in the elderly is often quiet or hypoactive.\(^4\) Those with agitated, hyperactive delirium may be most likely to attract provider attention, while those who are quietly delirious may appear to be “model patients.”\(^21\) Thus, delirium presenting quietly and without disruptive behaviors in elderly patients often may be missed by physicians.\(^8\) Notably, among the 3 most common confusional diagnoses assigned in our study (chronic organic brain syndrome, dementia with delirium, alcohol intoxication/withdrawal delirium), 1 is related to organic psychosis and the other to alcohol withdrawal. While we cannot say for certain what the phenomenological subtypes were in our sample with recorded delirium/confusional diagnoses, we speculate that many of these patients may have had hyperactive delirium with behavioral/physical manifestations (hallucinations, withdrawal symptoms) increasing their visibility to clinicians.

There are several limitations to our study. These include the fact that all patients were veterans and most patients were men. Thus, our results cannot be generalized to nonveteran populations containing equal numbers of men and women. In addition, our study was retrospectively performed, using information from an administrative database. Consequently, we could not prospectively assess actual delirium recognition rates. It is notable that the rate of delirium/related confusional diagnoses in FY96 was only 4% of all elderly patients hospitalized on VA inpatient units. We suspect this to be a gross underestimate of the true prevalence of delirium within the health care system. Some patients in the group with no recorded delirium very likely had actual episodes of delirium that were either unrecognized, unrecorded, or both.

Despite the caveats noted, the current study provides an important window to recorded delirium in elderly patients within a large health care system. Taking into account prior prospective prevalence estimates, the overall rate of delirium/related confusional diagnoses recorded in the VA health system is very likely an underestimate of true prevalence and possibly reflects poor recognition and recording of delirium syndromes in older veterans. We suspect that dementia-related delirium, perhaps in part due to greater attention to mental status in a preexisting neurologic disorder, and organic psychoses and alcohol-related delirium, perhaps in part due to hyperactive or disruptive behavioral symptoms, are more commonly diagnosed and recorded. Other types of delirium (other etiologies or motoric subtypes) are likely underrecognized and undertreated. Racial differences in delirium or other confusional diagnoses received may reflect both patient (differential symptom presentation) and provider (conscious or unconscious bias) factors. Future research should prospectively examine the relationship between motoric and etiologic delirium subtypes and recognition, as well as racial differences in the diagnosis of delirium.

References

Long-Term Effects of Donepezil on P300 Auditory Event-Related Potentials in Patients With Alzheimer’s Disease

Eiichi Katada, MD, PhD, Koichi Sato, MD, PhD, Akira Sawaki, MD, Yasuaki Dohi, MD, PhD, Ryuzo Ueda, MD, PhD, and Kosei Ojika, MD, PhD

ABSTRACT

The P300, one of the cognitive event-related potentials (ERPs) of the cerebral cortex, reflects the functioning of the neurochemical system involved in cognitive processes. We investigated clinical significance of the components of auditory P300 ERPs, in comparison with neuropsychologic tests including the Mini-Mental State Examination and the Japanese version of the Alzheimer’s Disease Assessment Scale–cognitive subscale (ADAS–J cog), for evaluating the effect of donepezil (DPZ) (5 mg daily for 6 months), an acetylcholinesterase inhibitor, in patients with Alzheimer’s disease (AD). Reduction of P300 latency associated with a parallel improvement of ADAS–J cog scores was observed after administration of 5 mg/day of DPZ in patients with AD. P300 latency gives very useful information on the progression of AD, especially in the longitudinal follow-up of patients with AD during treatment with DPZ acting on cholinergic pathways. (J Geriatr Psychiatry Neurol 2003;16:39–43)

Keywords: Alzheimer’s disease; event-related potentials; auditory P300; donepezil

Alzheimer’s disease (AD), one of the most common causes of mental deterioration in the elderly, is a progressive neurodegenerative disorder characterized by cognitive and behavioral dysfunction. The etiology of AD is not yet elucidated, but the cholinergic hypothesis of AD is stated on the ground of the presynaptic deficits found in the brains of patients with AD and studies of the role of acetylcholine in animal and human behavior. To assess the progression of AD, neuropsychological (Alzheimer’s Disease Assessment Scale–cognitive subscale [ADAS–cog]), Mini-Mental State Examination [MMSE]), neuroimaging (computed tomography, magnetic resonance imaging, single-photon emission computed tomography, positron emission tomography), and neurophysiological (electroencephalography, auditory P300 event-related potentials [ERPs]), investigations have been widely used. Among these methods listed above, P300, one of the cognitive ERPs of the cerebral cortex, reflects the functioning of the neurochemical systems involved in cognitive processes. The anticholinergic scopolamine delays P300 latency and decreases P300 amplitude, and physostigmine (cholinesterase inhibitor) reduces P300 latency in the short term, indicating that cholinergic neurons are important in the neuronal networks generating the P300 potential. P300 latency is very useful in determining the progression of AD while the ADAS–cog and the MMSE show the degree of cognition. Furthermore, P300 latency is known to change even at the early stage of AD with little cognition deterioration that cannot be detected by the MMSE. Therefore, changes in P300 latency may reflect the effects of treatment in patients with AD. In this study, we investigated clinical significance of P300 latency in comparison with MMSE and the Japanese version of the ADAS-cog (ADAS–J cog), for evaluating the effect of donepezil (DPZ), an acetylcholinesterase inhibitor, in patients with AD.

PATIENTS AND METHODS

Subjects

Thirteen subjects, aged 70 to 88 (5 men, 8 women; average age [mean ± SD] = 78.0 ± 6.1) were enrolled in this study. The subjects were outpatients diagnosed as having probable AD by the diagnostic criteria of the National
Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association and had undergone a neurological examination, neuroimaging, and laboratory workup to rule out other treatable causes of dementia. None had other significant psychiatric or medical illness or had taken medications that might affect cognitive ERPs. Patients in whom no reliable ERPs to “odd” stimuli could be recorded at baseline (since undetectable P300 is a common minority finding) were excluded because we considered recording of detectable P300 to be the basic prerequisite in this study.

The patients orally received DPZ (3 mg/day) for the first week to avoid gastrointestinal symptoms such as nausea or vomiting that might appear during the initial administration period and 5 mg/day thereafter over 6 months. Neuropsychologic findings and auditory P300 ERPs were obtained at baseline (ie, before initiation of DPZ treatment) and after 1, 3, and 6 months of the treatment with DPZ. The neuropsychologic tests were performed on the days the ERPs were recorded.

Neuropsychologic Assessment
Neuropsychologic tests included the MMSE and the ADAS–J cog in this study. The MMSE was used for grading the cognitive state of patients with AD. To assess the efficacy of DPZ, ADAS–J cog was used. The ADAS–J cog, the Japanese version of ADAS-cog that is a sensitive and reliable psychometric scale, is used as the selection criteria of a clinical trial. ADAS–J cog consists of 11 items that evaluate selected aspects of memory, orientation, attention, language, reasoning, and praxis, and the range of its total score is from 0 to 70. At baseline in this study, the selection criteria were as follows: MMSE scores ranging from 12 to 26 (average [mean ± SD] = 18.5 ± 4.3) and ADAS–J cog scores ranging from 9.4 to 58.0 (average [mean ± SD] = 23.6 ± 12.6).

Event-Related Potential Recordings
ERPs were recorded with Ag/AgCl disk scalp electrodes placed on the Fz, Cz, and Pz positions of the international 10-20 system. Reference electrodes corresponded to the linked earlobes. Ground was placed at Fpz. Interelectrode impedance was reduced to below 5 Kohm. The binaurally presented auditory stimuli were produced by a system applying averaging and a P300 stimulator: Neuropack Sigma (Nihon Kohden, Tokyo, Japan). We used a simple auditory oddball paradigm. Tones of 2 different frequencies, 1 kHz (frequent nontarget tone) and 2 kHz (rare target tone), were presented at a rate of 80% and 20%, respectively, in a random order at an interstimulus interval of 1.4 seconds. The rise-fall and plateau times were 10 and 100 msec, respectively. The stimuli were delivered binaurally through light headphones at an intensity of 70 dB nHL. The subjects were required to respond to the stimulus of the rare target tone by pressing a push button with their right index finger. We assessed the push button response capability and accuracy by observing the Neuropack Sigma display. The responses were amplified with filter band passes of 0.1 and 50 Hz and were averaged until the individual artifact-free trial was completed after 30 target tones. The P300 wave was identified according to the operating definition based on its probability sensitivity and the sequence of preceding components. The latency of P300 was measured at the peak of the potential at Pz, and the amplitude of the potential at Pz was measured in comparison with prestimulus baseline.

Statistical Evaluation
All 13 patients had a detectable P300 component during all recording sessions. Amplitude and latency measurements of P300, which were recorded during this study, were fed into a computer data management and analysis system program for statistical evaluation of data recorded at different times of the study. Continuous variables were reported as mean ± SD. Systematic differences of amplitude and latency of P300 between baseline and each trial period were evaluated using repeated-measure one-way analysis of variance. If the difference was significant, Fisher’s protected least significant difference was applied. Systematic differences of the scores of the MMSE and ADAS–J cog between baseline and each trial period were evaluated using the Friedman test, and if the difference was significant, the Wilcoxon signed-rank test with Bonferroni correction was applied. The Spearman rank correlation was applied to verify the correlation between MMSE scores and P300 and ADAS–J cog scores and P300.

RESULTS
All the patients completed the whole study without any side effects. The premedication and postmedication effects of treatment at DPZ were assessed in all 13 subjects.

MMSE Baseline Versus 1, 3, and 6 Months
Values of the MMSE demonstrated an improvement in 10 patients (76.9%) treated with DPZ at 1-month evaluation in comparison with those at premedication baseline (Figure 1A). An increase in mean MMSE scores was significant (vs premedication baseline) at 1 month and 3 months (P < .05, P < .01, respectively) (Table 1; Figure 2A). MMSE score decrements were significant (vs 3 months) at 6 months (P < .05) (Table 1; Figure 2A).

ADAS–J cog Baseline Versus 1, 3, and 6 Months
Decrements of the ADAS–J cog score were observed in 11 patients (84.6%) treated with DPZ at 1 month relative to those at premedication baseline (Figure 1B). ADAS–J cog score decrements were significant (vs premedication baseline) at 1, 3, and 6 months (P < .01) (Table 1; Figure 2B). The ADAS–J cog increase was significant (vs 3 months) at 6 months (P < .05) (Table 1; Figure 2B).
Long-Term Effects of Donepezil on P300 Auditory Event-Related Potentials in Patients With Alzheimer’s Disease / Katada et al.

P300 Baseline Versus 1, 3, and 6 Months
Typical recordings of P300 obtained from a subject in this study from baseline to 6 months are shown in Figure 3. Decrements of P300 latency were observed in 10 patients (76.9%) treated with DPZ at 1 month relative to those at premedication baseline (Figure 1C). P300 latency reduction was significant (vs premedication baseline) at the 1-month follow-up after the first administration ($P < .05$) (Table 1; Figure 2C). P300 latency was delayed in 6 patients (46.2%) treated with DPZ at the 6-month evaluation relative to those at the 1-month evaluation. P300 latency delay was significant (vs 1 month) at the 6-month follow-up ($P < .05$) (Table 1; Figure 2C).

Correlation Between Neuropsychological Tests and P300
Figure 4 shows the scatter plot of P300 latencies and ADAS–J cog scores recorded at baseline (A) and on completion of the study (6 months) (B) in patients with AD who were treated with DPZ. Because of the reduction of P300 latency and the modest improvement of ADAS–J cog scores, the correlation values changed on completion of the study. The ADAS–J cog score was significantly correlated with P300 latency but not with P300 amplitudes (data not shown) at baseline ($r = 0.676, P < .01$) (Figure 4A) and on completion of the study (6 months) ($r = 0.701, P < .01$) (Figure 4B). However, the MMSE score was not correlated

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Individual values of Mini-Mental State Exam (MMSE) scores (A), Alzheimer’s Disease Assessment Scale-cognitive subscale (Japanese version) (ADAS–J cog) scores (B), and P300 latencies (C) showing the difference between baseline and 1 month after the first administration of donepezil in 13 patients with Alzheimer’s disease.

![Figure 2](https://example.com/figure2.png)

**Figure 2.** Effect of donepezil (5 mg/day) on Mini-Mental State Exam (MMSE) scores (A), Alzheimer’s Disease Assessment Scale-cognitive subscale (Japanese version) (ADAS–J cog) scores (B), and P300 latencies at Pz (C) in 13 patients with Alzheimer’s disease. The mean ± SD values are expressed as changes from baseline to 6 months. *$P < .05$. **$P < .01$.

![Table 1](https://example.com/table1.png)

<table>
<thead>
<tr>
<th>Baseline</th>
<th>1 Month</th>
<th>3 Months</th>
<th>6 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>18.5 ± 4.3</td>
<td>21.8 ± 4.8</td>
<td>22.1 ± 4.6</td>
</tr>
<tr>
<td>ADAS–J cog</td>
<td>23.6 ± 12.6</td>
<td>18.4 ± 9.6</td>
<td>18.3 ± 10.7</td>
</tr>
<tr>
<td>P300 latency (ms)</td>
<td>404.3 ± 49.9</td>
<td>381.5 ± 42.2</td>
<td>394.8 ± 39.2</td>
</tr>
<tr>
<td>P300 amplitude (µV)</td>
<td>12.7 ± 8.1</td>
<td>13.0 ± 9.4</td>
<td>10.1 ± 7.2</td>
</tr>
</tbody>
</table>

MMSE = Mini-Mental State Examination; ADAS–J cog = Alzheimer’s Disease Assessment Scale-cognitive subscale (Japanese version). Values are expressed as mean ± standard deviation.
with P300 latency in patients with AD throughout the study.

DISCUSSION

Our study showed that P300 latency was reduced by the 6 months’ treatment with 5 mg/day of DPZ with a parallel improvement of ADAS–J cog scores in patients with AD. Two previous reports were published concerning the effect of DPZ on auditory P300 ERPs.13,15 Reeves et al15 reported that auditory P300 latency was not normalized in 11 patients with AD after 1 month of treatment with a dosage of 5 mg/day of DPZ, but the treatment was associated with a significant reduction in P300 latency (mean pretreatment latency = 401.5 msec; mean posttreatment latency = 392.7 msec). They also indicated that DPZ had no discernible effect on the auditory P300 peak-to-peak amplitude or baseline-to-P300 peak amplitude. Thomas et al13 reported that in 20 patients with AD, treatment with the DPZ (5 mg/day for 15 days and 10 mg/day thereafter) reduced the P300 latency (mean pretreatment latency = 382.7 msec; mean posttreatment latency = 368.8 msec at 1 month, 366.9 msec at 3 months, and 367.4 msec at 6 months). Our data concerning the changes of P300 latency were similar to the results of these previous reports. However, there were a few differences. Our study indicated that the P300 latency reduction was significant (versus baseline) at 1 month but the P300 latency delay was significant (versus 1 month) at 6 months, although Thomas et al reported that the P300 latency reduction persisted for 6 months after the initial treatment with DPZ. The characteristics of patients with AD were quite different between our study and that of Thomas et al; the mean age was 66.5 ± 9.1 years, the mean MMSE score was 16 ± 0.5, and the mean ADAS-cog score was 33.34 ± 2.7 in the study by Thomas et al. Thus, the difference in the effect of DPZ on P300 latency between the studies is probably based on the selection of responders and the dosage of DPZ.

The effect of DPZ on P300 latency was rapid and consistent, indicating that P300 ERP recordings might represent a rapid and convenient measurement that can be performed repeatedly in the evaluation of patients with AD during treatment with DPZ acting on cholinergic pathways. However, there are several problems related to P300 ERP recordings in the evaluation of patients with AD. For example, methodologically, no reliable P300 components to odd stimuli at baseline could be recorded in some subjects: undetectable P300 is a common finding in a minority (5%-10%) of normal subjects.7 Moreover, neurophysiological change does not necessarily imply clinical or behavioral change. Results of neuropsychological tests sometimes remained in the abnormal range, and clinical improvement was not significant despite of the improvement of P300 latencies. The discrepancy between P300 latencies and clinical features may be explained by the fact that P300 latency reflects cognitive dysfunction more sensitively than clinical or behavioral features in patients with AD. Indeed, P300 latency changes in patients with AD at an
Long-Term Effects of Donepezil on P300 Auditory Event-Related Potentials in Patients With Alzheimer’s Disease / Katada et al. 43

early stage with very little cognitive deterioration. Thus, P300 is very useful in monitoring the efficacy of treatment with DPZ in AD.

We also found that the ADAS–J cog score was significantly correlated with P300 latency at baseline and on completion of the study (6 months). Correlation between latency of P300 ERPs and cognitive dysfunction appears to be well established: P300 latency is correlated negatively with mental function in normal subjects, and severe delays are observed in patients with AD in comparison with age-matched control subjects. These results suggest that P300 latency may serve as a marker for measuring the subclinical and clinical course of AD.

We should note that there were no elderly normal-control subjects treated with DPZ, and the specificity of these observations for AD is unknown. Furthermore, this was not a double-blind, placebo-controlled study. However, we can conclude that P300 latency gives very useful information on the progression of AD, especially in the longitudinal follow-up of patients with AD during treatment with DPZ acting on cholinergic pathways.

References
Affective prosody is the “melody of speech” that provides emotional and attitudinal information during discourse. The acoustic features of affective prosody include aspects of frequency, duration, and intensity. Varying fundamental frequency (F0) over time is perceived by listeners as pitch or intonation variation and has been shown to be one of the most salient cues to determining the affective intent of an utterance. The absence of normal pitch variation in speech causes individuals to sound “flat,” as if they have no feeling even though they may experience normal emotion. Although pitch variation is important, loudness (measured as intensity), timing (measured as duration), quality (measured as spectral information), and variations in syllable structure also provide acoustic cues to affective prosodic intent.

Although affective prosody appears to be a dominant function of the right hemisphere, production and comprehension deficits have been reported following both right-brain damage (RBD) and left-brain damage (LBD). Such deficits have been explained in various ways. For example, brain injury may cause individuals to rely preferentially on one set of acoustic cues opposed to another. Van Lancker and Sidtis found that RBD patients relied mostly on durational cues whereas LBD patients relied mostly on F0 variability, but these findings were not replicated by Pell and Baum. Also, in LBD patients with dense aphasias, the severity of aphasic comprehension deficits has been correlated with the severity of affective prosodic comprehension deficits, suggesting that the verbal and articulatory impairments associated with aphasia might explain impaired affective prosodic performance.

In an attempt to avoid the confounding contribution of aphasic deficits, an Aprosodia Battery was developed using stimuli that are reduced progressively in their verbal articulatory content. It was assumed that reducing the verbal articulatory content causes the tasks to shift from bihemispheric to primarily right hemispheric processing. This assumption is consistent with findings from psychoacoustical studies that suggest a specialized role for the right auditory cortical areas in processing pitch direc-

**ABSTRACT**

This study evaluated the ability to produce and comprehend affective prosody across age groups and compared patterns of impaired performance to deficits observed after focal brain damage. Sixty-nine healthy subjects, ages 22 to 83 years, were given the Aprosodia Battery, a test that distinguishes between affective prosodic processing deficits following right- versus left-brain damage through the use of stimuli with progressively reduced verbal articulatory content. Production of affective prosody, measured by variation in fundamental frequency, was unimpaired in older subjects, whereas comprehension of affective prosody was impaired, particularly for tasks with reduced verbal articulatory content. The pattern of performance across affective comprehension tasks in the older subjects resembled the pattern found after right-brain damage. The results demonstrate age-related loss in comprehension of affective prosody that is most likely due to a processing deficit involving the right hemisphere. (J Geriatr Psychiatry Neurol 2003; 16:44–52)

**Keywords:** affective prosody; aging; auditory comprehension; stroke
tion and a specialized role for the left hemisphere in processing temporal information.

In the Aprosodia Battery, acoustic cues to affective prosody are reduced along with the reduction of verbal articulatory content. The word stimuli have a full complement of prosodic cues, including variations in pitch, rhythm, loudness, and voice quality along with normal syllabic structure, thus recruiting both left and right hemispheres for processing. The monosyllabic stimuli also include a full complement of acoustic cues to affective prosody, although the syllabic information is no longer linguistically meaningful. Finally, the asyllabic stimuli consist mainly of suprasegmental variations in pitch, loudness, and voice quality that are carried over the entire utterance with limited rhythmic and no segmental cues. The syllabic condition therefore taxes predominantly right hemisphere processing.

Using the Aprosodia Battery, 2 distinct patterns of affective prosodic performance were observed in brain-damaged subjects. LBD performance, for both comprehension and production, improved relative to control subjects as the verbal-articulatory content of the test stimuli was reduced. In contrast to previous findings in densely aphasic patients, affective prosodic deficits in the LBD group did not correlate with aphasic deficits. This difference may have been due to subject selection because some of the LBD patients had mild or no aphasic deficits. In this less impaired group of LBD patients, the presence of deep white matter lesions adjacent to the corpus callosum best predicted affective prosodic deficits. Thus, affective-prosodic deficits after LBD were attributed to a disruption of callosal integration between affective prosody centers of the right hemisphere and propositional language centers of the left hemisphere. Conversely, RBD performance for both production and comprehension remained impaired or worsened as the verbal-articulatory content of test stimuli was reduced with the cortical rather than the deep distribution of the lesions best predicting affective-prosodic deficits. Thus, the affective prosodic deficits after RBD were attributed to a disruption of the right hemisphere’s dominant processing of affective prosody, a conclusion also reached by other investigators using different test methods.

While studying processing of affective prosody in various clinical populations, we observed that our older control subjects seemed to have difficulty comprehending affective prosody compared to younger subjects. A literature review showed that other investigators have reported age-related impairments in the recognition of auditory affect. Brosgole and Weisman compared 6 groups with ages spanning from 3 to 83 years on tasks of auditory and facial affect recognition and discrimination, as well as recognition of nonemotional voice inflections. They found that the ability to identify facial affect, nonemotional voice inflections, and emotional voice intonations began to decline in their 45- to 64-year-old group with further worsening in their 65- to 83-year-old group. The authors postulated, based on their findings and others, that plaque deposits may be the underlying cause of age-related changes in mood recognition and that the deficits might represent an early sign of impending dementia.

The goals of the present study were (1) to explore the relationship between Aprosodia Battery scores and age and (2) to compare impaired patterns of performance found in apparently healthy subjects to those observed after RBD and LBD.

METHODS

Subjects

The primary study group included 69 control subjects (29 men, 40 women) from our previous clinical studies of affective prosody. Demographic information is provided in Table 1. All subjects were free from medical conditions known to cause cognitive deficits, such as history of major psychiatric (eg, schizophrenia, bipolar illness), medical (eg, serious cardiovascular disease), or neurologic (eg, head trauma, vascular dementia) conditions. Subjects were also free from obvious hearing loss. All control subjects scored 27 or better on the Mini-Mental State Exam. Individuals were excluded if taking medications that could interfere with testing, such as neuroleptics, antidepressants, high-dose β-blockers, or benzodiazepines. All subjects were native speakers of English and right handed based on self-report and the Edinburgh Handedness Scale.

Because the second goal of this study was to compare patterns of performance in apparently healthy subjects to those found after brain damage, data from 37 brain-damaged patients, 19 of whom have been reported else-

Table 1. Demographic Information for Experimental Subjects

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Normal Subjects</th>
<th>Brain-Injured Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entire Group</td>
<td>Young (n = 38)</td>
</tr>
<tr>
<td>Age in years (Mean ± SD)</td>
<td>58 ± 20</td>
<td>30 ± 6</td>
</tr>
<tr>
<td>Years of education*(Mean ± SD)</td>
<td>14 ± 3</td>
<td>15 ± 2</td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>29/40</td>
<td>9/11</td>
</tr>
</tbody>
</table>

a. Young subjects had significantly more years of education than the left-brain-damaged and right-brain-damaged subjects did, F(4, 101) = 4.01, P = .005.
where were included. The brain-damaged subjects met the above criteria with the exception that each had experienced a unilateral ischemic infarction documented by magnetic resonance imaging involving either the right (n = 20) or left (n = 17) hemisphere (see Table 1 for demographic information). Testing of the brain-damaged patients was completed within 4 to 8 weeks postictus to exclude acute deficits due to diaschisis and improvement from spontaneous recovery due to neural reorganization.

Aprosodia Battery

Production

Production of affective prosody was assessed by tape recording subjects on 3 tasks of affective repetition and on a spontaneous production task. The repetition stimuli consisted of 3 sets of randomized utterances reduced progressively in verbal articulatory content. Each set consisted of 12 utterances, 2 renditions each of 6 emotions (happy, sad, disinterested, neutral, surprised, and angry), with 1 rendition having emphatic stress early in the utterance and the other rendition having emphatic stress late in the utterance. The carriers were “I am going to the other movies” for word repetition, “ba ba ba ba ba” for monosyllabic repetition, and “aaaaahhhhh” for asyllabic repetition. Subjects were instructed to listen to each stimulus and then repeat what they heard using the same tone of voice. They were allowed 4 practice utterances before each set of stimuli. All stimuli were played in a free sound field at a comfortable loudness level using a prerecorded audio compact disc and a Sony compact disc player (Sony Electronics Inc., Park Ridge, NJ).

Spontaneous production of affective prosody was assessed by recording each subject during a 5- to 10-minute interview, during which they were encouraged to discuss various emotionally provocative personal life events that made them feel happy, frightened, angry, or sad.

Subject responses were recorded on analogue tape using a Marantz PMD 340 tape recorder and a Shure SM12A microphone mounted on an adjustable boom attached to a headset. The microphone was position.22ed just to the side of the subject's air stream to avoid recording artifacts from head movement and to ensure a high signal-to-noise ratio.

Utterances were evaluated using a PM Pitch Analyzer (Voice Identification Inc., Somerville, NJ), which extracted fundamental frequency (F₀) in Hertz (Hz). Next, the mean, standard deviation, and coefficient of variation (CV) of F₀ were calculated for each utterance. Finally, a mean F₀ CV percentage (F₀ CV%) was calculated for all the utterances within each repetition task. The same analyses were conducted on a total of 10 seconds of each subject's spontaneous production, using phrases 2.5 to 4 seconds long, taken from sections of discourse involving recall of emotional situations. Fundamental frequency variation was chosen as an objective measure of the affect in speech because it has been found that in speakers of English, a nontone language, it correlates with clinical loss of affective prosody after brain injury.15 If individuals lose general affective expression, we would expect to find lower mean F₀ CV% compared to the younger subjects. There is, however, no evidence in the literature to suggest a change in affective production with age.

Comprehension

Comprehension of affective prosody was assessed using 3 identification tasks—word, monosyllabic, and asyllabic—and an affective discrimination task. The stimuli for the 3 identification tasks were the same as those described above for repetition. The 12 stimuli were randomized and presented twice for 24 trials per task. Subjects were asked to identify the emotional intonation of each stimulus by choosing the appropriate affect from a vertical array of 6 line drawings of faces expressing different affects. Next to each face was the corresponding written label of “neutral,” “happy,” “sad,” “disinterested,” “surprised,” and “angry.” Before testing, each subject demonstrated the ability to identify the facial expressions and to read the written labels.

The affect discrimination stimuli were the same as those used for word repetition after band-pass filtering between 70 and 300 Hz (using a Krohn-Hite Variable filter [Model 3550]), a process that distorted the phonetic information while leaving overall prosodic information intact.35 Twenty-four pairs of stimuli were recorded; 12 pairs had the same affective intonation with different stress patterns, and 12 had different intonation with the same stress pattern. Subjects were asked to indicate whether the represented emotions were the same or different for each pair. Subjects were expected to perform poorly when basing their answers on stress rather than affective prosodic information.

The scores for each of the comprehension tasks were the total number of correct responses out of 24.

Data Analysis

To address the first goal of this study, regression analyses were used to explore the relationship between age and scores on the Aprosodia Battery tasks. To address the second goal of this study, the primary study group was divided into 3 age groups—young (ages 22 to 44 years), middle (ages 45 to 64 years), and elderly (ages 65 to 83)—corresponding to the adult groups used by Brosgole and Weisman.28 Then, patterns of impaired affective prosodic processing were explored across relevant tasks using repeated-measure ANOVAs. Finally, any impaired performance was compared to patterns found after LBD and RBD using paired repeated-measure ANOVAs. Alpha was set at .05 for all analyses. Statistical analyses were carried out using SPSS 8.0 for Windows.

In the original study using the Aprosodia Battery, subject raw scores were converted to z scores by subtracting the control group mean score from the individual subject score and then dividing by the control group stan-
standard deviation. This removed the variability in performance across the repetition and comprehension tasks attributable to unknown factors in the control group while leaving the variability attributable to brain damage intact. Because the focus of the present study is on healthy subjects’ performance, all results are presented as raw rather than z scores.

RESULTS

Preliminary Analysis

To screen for the presence of any unexpected predictive value for gender or education, a preliminary regression analysis was conducted that included these factors along with age in the model. Education did not predict performance on any of the Aprosodia Battery tasks; thus, it was dropped from further analyses. Gender was predictive of asyllabic repetition ($R^2 = 0.116; df = 1, 67; P = .004$); therefore, regression models predicting asyllabic repetition scores with age were run for each gender group separately. Within each gender, the predictive value of age for asyllabic repetition was virtually identical for women ($R^2 = 0.111, r = +0.333$) and men ($R^2 = 0.117, r = +0.342$). Although the relationship between gender and asyllabic repetition is an interesting one that may deserve further attention, because the focus in the present study was on age and the relationship between age and asyllabic repetition was similar for both genders, gender was not considered in further analyses.

Aprosodia Battery Scores and Age

Age was found to be a significant predictor of Aprosodia Battery scores on 6 of the 8 tasks (Table 2). For production, age was related to word and asyllabic repetition (Figure 1). Because these relationships were positive and statistically small ($R^2 < 0.10$), there was no indication that age impairs the ability to produce affective prosody in speech.

For comprehension, age was negatively related to all of the comprehension tasks (Table 2). No sharp changes occurred with age (Figure 2). Instead, there was a gradual fall of test scores with greater variability in test performance as age increased. This finding was most apparent for asyllabic and monosyllabic identification and least apparent for word identification and affective discrimination. The best-fitting trend lines were linear for all tasks except asyllabic identification, which was curvilinear. Trend line equations and $R^2$ values are noted on each graph. Ceiling effects, observed for all of the tasks, may have influenced the slope of the trend lines and may account for the curvilinear best fit for asyllabic identification. However, it should be noted that these ceiling effects make it more rather than less difficult to observe differences related to age. Based on regression analyses, there appears to be impairment of affective prosodic comprehension associated with age.

Patterns of Performance

To further explore patterns of impairment on affective comprehension, the performance across identification tasks for the 3 age groups (Figure 3) were compared using a repeated-measure ANOVA. The ANOVA yielded a significant main effect for task, $F(2, 132) = 29.30, P < .001$, and group, $F(2, 66) = 20.29, P < .001$, and a significant group by task interaction, $F(4, 132) = 7.50, P < .001$. Post hoc tests comparing groups within each identification task showed no significant differences among the age groups on word
identification, but they did show a significant difference between the young and elderly groups on monosyllabic identification and among all 3 groups on asyllabic identification. Based on the presence or absence of significant interactions on paired repeated-measure ANOVAs (Table 3), the two older groups were found to have similar patterns of performance across identification tasks (no significant interaction). However, there was a main effect for group indicating that the elderly group was more impaired than the middle group. The young group’s performance across tasks was different from both the middle and elderly groups based on the presence of significant interactions.

Finally, the patterns of performance from the age groups were compared to those found following brain damage (Figure 4). Based on presence or absence of significant interactions on the paired repeated-measure ANOVAs (Table 4), similar patterns of performance were observed for the elderly, middle, and RBD groups, and similar patterns of performance were observed for the young and LBD groups.

**DISCUSSION**

The primary goal of the present study was to explore the relationship between age and the ability to produce and comprehend affective prosody. The results suggest that production of affective prosody remains intact with age, while comprehension of affective prosody is impaired in the older relative to younger subjects. The data in the present study showed a small positive correlation with the amount of fundamental frequency variation produced by subjects on word and asyllabic repetition, while no relationship between age was found for monosyllabic repetition or spontaneous production. Although the positive correlations with repetition and age are interesting, the findings do not indicate any impairment in the ability to produce affect as a function of age. It may be that speech is such an overlearned and well-developed skill that it remains relatively impervious to the aging process.

The ability to comprehend affective prosody, however, showed a consistent negative relationship with age. Performance on all 4 of the comprehension tasks was negatively correlated to age, with the effect being most impressive for asyllabic identification. The asyllabic task
uses stimuli with the least amount of verbal cues to affect and no cues associated with segmentation. Subjects must make affective distinctions based primarily on changes in pitch, a task modulated predominately by the right hemisphere. Although intonation or pitch variation over time has been shown to be one of the most salient acoustic cues for signaling affective prosody in spoken English, a nontone language, older individuals may be relying more heavily than younger individuals on cues such as rhythm, loudness, and voice quality or may need greater redundancy of cues to make affective determinations. Either way, the current results suggest that when intonation is the primary cue to affective prosody, the elderly exhibit more difficulty than do the young. This may reflect a right hemisphere processing disorder based on the idea that the Aprosodia Battery changes from a bihemispheric task to primarily a right hemispheric task as the verbal articulatory content of test stimuli is reduced.

The second goal of this study was to compare patterns of performance across tasks that are impaired with age, specifically affective identification. To do this, the healthy study subjects were split into 3 groups after Brosgole and Weisman, and their performance across tasks was compared to brain-damaged subjects. First, we noted that the middle (ages 45-64) and elderly (ages 65–83) groups were impaired relative to the young group on tasks of affective prosodic identification, with the middle group being intermediate and the elderly group performing the worst, a finding consistent with those of Brosgole and Weisman. Second, we noted that the pattern of performance exhibited by the 2 older groups was similar to that of the RBD group, whereas the young group was similar to the LBD group. The 2 older groups’ and the RBD groups’ performance declined across tasks as the verbal articulatory content of test stimuli was reduced, whereas the performance of the young and LBD groups remained consistent across tasks. Thus, the impaired pattern of performance across identification tasks found in the older groups appears consistent with a right rather than a left hemisphere processing disorder.

Although our subjects had no overt hearing difficulty, hearing loss due to changes in both the peripheral and central auditory system are prevalent in the elderly. Thus, it is possible that the older subjects were not able to hear the pitch variations well enough to make affective identifications. In past research, it has been assumed that hearing deficits are not a contributing factor to age-related deficits in affect recognition because subjects had no obvious difficulty with verbal communication in test situations and they passed screening tests involving verbal instructions. Likewise, in the present study, intact performance on the affective repetition tasks suggests that test stimuli were presented at audible levels sufficient for nascent perceptual analysis.

If hearing is not an issue, than what could be the cause of the relatively poor performance on comprehen-
sion tasks exhibited by the older individuals? One possibility is that the tasks simply become more difficult as the stimuli become more unnatural due to reduction in verbal articulatory content. As mentioned above, the reduced verbal articulatory content probably forces subjects to rely solely on intonational cues in the absence of supporting segmental and rhythmic indices of affective prosody. The relatively poor performance on affective prosodic tasks by older subjects compared to younger subjects may imply a difference in cognitive strategy employed by the older subjects.

Differences in cognitive strategy have been inferred from neuroimaging studies showing activation of different brain areas when young and elderly perform identical tasks. More specifically, imaging studies suggest greater use of frontal cognitive systems with age. This recruitment of different and, by extension, possibly less efficient brain areas may result in either equivalent or increased lateralization of language functions to the left hemisphere with age. In contrast, studies involving visual perception of emotional words and the production and perception of right and left hemifacial emotions have suggested a continued and competent role for the right hemisphere in modulating affective behavior across the life span. However, recent studies probing the ability to process facial blends of emotion have reported age-related deficits in the ability to attend and perceive emotion conveyed on the upper half of the face, a finding that was consistent with right hemisphere aging. Additional hypotheses posit that both hemispheres age equally or that the frontal lobes are the focus of age effects.

Lastly, an important question to ask is, How do age-related findings in the comprehension of affective prosody affect elderly people in everyday communication situations? The aging effect observed on regression analysis for word identification was relatively small compared to those found for monosyllabic and asyllabic identification. Using tests similar to word identification, others have found that only a subgroup of healthy elderly was impaired on affect recognition. Thus, many elderly people may be able to interpret basic emotions such as happiness, sadness, and anger in typical communication situations that involve a synthesis of affective prosodic cues with context and facial expression. This argument, however, should be made cautiously because age-related impairments have also been reported for comprehension of facial emotions. In addition, the ability of elderly subjects to interpret the subtler cues underlying attitudinal prosody that convey important psychosocial information in everyday situations, such as sarcasm, tacit approval, and disapproval, needs exploration.

Regardless of our speculation about underlying mechanisms, the results have empirical validity. Age-associated loss of affective prosodic comprehension, coupled with a possible reduction in the ability to detect facial and lexical emotion, carries certain implications. An elderly person’s perception of social and emotional support may be
attenuated if the affective aspects of communication involving social interaction are missed or misinterpreted. Consequently, they may be less adept at interacting with family, friends, and professionals and may be at a disadvantage during deceitful situations.

References


Gait disturbances and falls are common in Parkinson’s disease (PD), often leading to institutionalization and a loss of independence in the advanced stages of the disease.\(^1\)-\(^3\) Gait disturbances typically include slowed walking and increased gait instability, often manifest by increased stride-to-stride variability in walking.\(^4\)-\(^6\) Despite the incapacitating effects of falls,\(^7\) the specific factors that contribute to falls and fall prediction in PD remain elusive.\(^8\) Impaired attention and a reduced ability to carry out dual tasks may exacerbate gait disturbances and contribute to fall risk in PD,\(^1,9\) as is the case in other populations.\(^10,11\) Recent work, however, suggests that the relationship between fall risk and dual-task performance may not be so simple in PD.\(^12\) For example, Bloem et al\(^12\) found that “stops walking while talking” apparently does not predict falls in patients with PD. Thus, while it is clear that enhancing attention to gait via “cueing” improves walking in PD,\(^6\) the effects of reduced attention on gait and fall risk in PD are not yet fully understood.

The present study is motivated by 4 factors: (1) the strong relationship between increased gait variability and fall risk observed in other populations,\(^13,14\) (2) the observation that gait variability is increased in PD,\(^4\) (3) the evidence highlighting the important influences of attention on balance and gait,\(^1,15\) and (4) the poor understanding of the factors that contribute to falls in PD.\(^7\) Here, we sought to gain insight into the effects of dual-task performance and attention on fall risk and gait instability in PD. More specifically, in this pilot study, we tested the hypothesis that walking while performing a cognitively challenging task (CCT) would impair the ability of patients with PD to maintain a stable walk.

### METHODS

#### Subjects

This pilot study is based on the analysis of data from 10 subjects with idiopathic PD. The subjects ranged in age...
from 52 to 82 years old. Hoehn and Yahr stage during testing ranged between 1 and 4 (mean = 3.1). Patients were receiving routine care for advanced PD in a movement disorders unit. Disease duration ranged between 6 and 26 years (mean = 13.4 years). Medication usage was not altered. All subjects were tested while in an “on” state (ie, on medications). All subjects provided informed consent.

Cognitive status and subject characteristics were assessed using the Mini-Mental State Exam (MMSE) and the Unified Parkinson’s Disease Rating Scale (UPDRS). Average total score on the UPDRS during “on” was 33.9 (range = 7-66). Subtotals for items grouped by mental function (UPDRS part I), activities of daily living (UPDRS part II), and motor function (UPDRS part III) were 2.5, 11.8, and 14.0, respectively. Mean score on the MMSE was 27.1 (range = 23-30). No patient met the DSM–IV criteria for dementia. Fall risk (retrospective) and postural instability were assessed using questions 13 and 30 of the UPDRS, respectively.

Assessment of Gait Dynamics
The protocol and methods used to assess the dynamics of walking, for example, gait variability, were similar to those used previously. Briefly, subjects were instructed to walk at their normal pace on level ground from one end of a hallway to the other end (approximately 20 meters away) and then back again. (During normal walking and dual-task walking, a clinician walked near the subject to ensure safety.) To measure the gait rhythm and the timing of the gait cycle, force-sensitive insoles were placed in the subject’s shoe. These inserts produce a measure of the force applied to the ground during ambulation. A small, lightweight (5.5 × 2 × 9 cm; 0.1 Kg) recorder was worn on the ankle and held in place using an ankle strap. An on-board A/D converter (12 bit) sampled the output of the footswitches at 300 Hz and stored the data. Subsequently, the digitized data were transferred to a Linux workstation for analysis using software that extracts the initial contact time of each stride (for 1 foot). With this information, the stride time or duration of the gait cycle (time from initial contact of one foot to subsequent contact of the same foot) was determined for each stride during the walk.

To study the intrinsic dynamics of the gait rhythm, the time series of the stride time was divided into 3 time segments for both normal walking and walking while cognitively challenged. We analyzed the first walk down the hall, before any turns (FirstLap), the complete walk (CompleteWalk), and the initial 5 strides (FirstStrides). Three subjects experienced freezing of gait, but all freezing episodes were excluded from the analysis. For analysis of the complete walk, turns and any starting and stopping episodes were automatically excluded by filtering out those stride times that were greater than 2.5 seconds or less than 0.5 seconds in duration. In addition, to focus on the dynamics of continuous walking, a median filter was applied to remove data points that were 3 standard deviations greater than or less than the median value. Subsequently, the average stride time was determined for each segment. Stride time variability, the magnitude of the stride-to-stride fluctuations in the gait cycle duration, was calculated by determining the standard deviation and the coefficient of variation (CV) of each subject’s stride time. The CV assesses the magnitude of the deviations of the stride time with respect to each subject’s mean value. These 2 measures of the magnitude of the stride-to-stride variability, the standard deviation and the CV, were highly correlated with one another (r > 0.95, P < .001), and all results were similar with both measures. In addition to the above measures, we also determined if a subject stopped walking during each trial (ie, the occurrence of “stop walking while talking” or SWWT). Finally, a neurologist with expertise in PD and freezing of gait documented the occurrence of any freezing episodes (ie, the subject tries to walk but is stuck in place). To a large degree, SWWT and freezing can be distinguished visually. In contrast to freezing, during SWWT, there is no effort made to continue walking.

Dual Task
After completing the normal walk, subjects walked along the same pathway and under the same conditions while being asked to serially subtract 7s from a 3-digit number (eg, 200, 193, 186). Specifically, subjects were instructed to walk at their usual pace, without stopping, while performing serial 7 subtractions out loud. Serial 7 performance was not evaluated. Serial 7 subtraction has been widely used as means of providing a distraction and a cognitive challenge.

Statistical Analysis
Results are reported as mean ± standard deviation. The Wilcoxon signed rank test, a nonparametric analog of the paired t test, was used to compare the results from normal walking with those from walking while performing the cognitive challenge. Spearman’s correlation coefficient was used to evaluate the association between measures. A P value less than .05 (2-tailed) was considered statistically significant. Statistical analysis was performed using SAS version 7.0 software.

RESULTS
An example of the effect of dual tasking on gait dynamics in a subject with PD is shown in Figure 1. During normal walking, relatively small stride-to-stride fluctuations are seen in the stride time of this subject. In contrast, when walking under the dual-task condition, the subject’s gait becomes hesitant and the stride-to-stride fluctuations become very large. In addition, the average stride time is longer compared to walking without dual tasking.

In general, a similar pattern was observed for all of the subjects with PD, no matter which time segment
was analyzed (see Table 1). For example, during First-Lap, stride time variability increased in each of the 10 PD subjects, and the average stride time increased in 9 of the 10 PD subjects. Note that for all 3 analysis segments, the percentage increase in stride time variability tended to be much larger than the percentage increase in average stride time (in 26 of 30 pairwise comparisons), and the 2 measures of change were not correlated ($P > .05$).

The observed, marked effect of the cognitive challenge on the dynamics of gait in PD is likely abnormal. Consistent with this, we note anecdotally a very different response to dual tasking in one subject with “primary freezing of gait.” In this older adult, gait dynamics was similar during normal walking and during dual-task walking (see Figure 2). For example, in the FirstStrides segment, stride time variability was 53 msec during normal walking and 45 msec during dual-task walking. In contrast to the subjects with PD, in this 77-year-old elderly woman who did not have PD, the cognitive challenge essentially had no effect on gait dynamics.

This initial result suggests that age alone may not explain the effects of dual tasking on gait. A closer look at the PD subjects supports this idea. In all 3 analysis periods (ie, FirstLap, FirstStrides, and CompleteWalk), age was not associated with the mean stride time ($P > .35$) or either measure of stride time variability ($P > .56$) during dual-task walking. Moreover, when studying only the 5 youngest subjects (the oldest of these subjects was 61 years of age), we find results that are consistent with those of the entire group. For example, stride time variability increased in all 5 of these subjects during FirstLap ($\Delta$ variability = 153 msec, $P = .043$), during CompleteWalk ($\Delta$ variability = 79 msec, $P = .043$), and during FirstStrides ($\Delta$ variability = 223 msec, $P = .043$).

Dual tasking consistently affected gait variability in the subjects with PD; however, SWWT was observed in only 2 subjects. These 2 had fairly advanced PD (Hoehn and Yahr’s scores of 3 and 4). Nonetheless, other subjects with advanced PD did not exhibit SWWT. For these 2 subjects, relatively high and extreme values of gait variability were obtained when walking while cognitively challenged (eg, during FirstLap, they had the highest gait variability from among all subjects).

In the 10 subjects with PD, gait variability during normal walking was associated with disease severity, disease duration, fall risk, and motor and cognitive function (Table 2). There was a wide range in age, but as noted above, variability was not associated with age. Although the response to dual tasking was more severe in some subjects with more advanced disease, the associations between clinical characteristics (eg, motor function) and gait variability became weaker during dual-task walking. The average stride time was not significantly associated with age, disease duration, fall risk, or any other clinical measure, under both walking conditions. The percentage

<table>
<thead>
<tr>
<th>Time Segment</th>
<th>Measure of Stride Time</th>
<th>Normal Walking</th>
<th>Cognitively Challenged</th>
<th>Percentage Increase</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FirstLap</td>
<td>Average (msec)</td>
<td>940 ± 67</td>
<td>1137 ± 228</td>
<td>21 ± 21</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>Variability (msec)</td>
<td>47 ± 29</td>
<td>223 ± 281</td>
<td>384 ± 595</td>
<td>.002</td>
</tr>
<tr>
<td>CompleteWalk</td>
<td>Average (msec)</td>
<td>946 ± 74</td>
<td>1046 ± 126</td>
<td>10 ± 9</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>Variability (msec)</td>
<td>64 ± 37</td>
<td>142 ± 109</td>
<td>154 ± 164</td>
<td>.014</td>
</tr>
<tr>
<td>FirstStrides</td>
<td>Average (msec)</td>
<td>950 ± 45</td>
<td>1169 ± 283</td>
<td>23 ± 28</td>
<td>.019</td>
</tr>
<tr>
<td></td>
<td>Variability (msec)</td>
<td>33 ± 28</td>
<td>243 ± 361</td>
<td>915 ± 1912</td>
<td>.008</td>
</tr>
</tbody>
</table>

Values are mean ± SD.
Table 2. Association Between Gait Variability and Parkinson’s Disease Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Normal Walking</th>
<th>Walking While Cognitively Challenged</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease duration</td>
<td>0.81 (.005)</td>
<td>0.76 (.010)</td>
</tr>
<tr>
<td>Hoehn and Yahr stage</td>
<td>0.95 (&lt; .0001)</td>
<td>0.57 (.088)</td>
</tr>
<tr>
<td>Disease severity (UPDRS total)</td>
<td>0.84 (.002)</td>
<td>0.63 (.083)</td>
</tr>
<tr>
<td>Mental function (UPDRS part I)</td>
<td>0.72 (.019)</td>
<td>0.17 (.64)</td>
</tr>
<tr>
<td>Activities of daily living (UPDRS part II)</td>
<td>0.56 (.089)</td>
<td>0.59 (.074)</td>
</tr>
<tr>
<td>Motor function (UPDRS part III)</td>
<td>0.74 (.015)</td>
<td>0.63 (.052)</td>
</tr>
<tr>
<td>Fall risk</td>
<td>0.68 (.042)</td>
<td>0.55 (.098)</td>
</tr>
<tr>
<td>Postural instability</td>
<td>0.76 (.023)</td>
<td>0.52 (.122)</td>
</tr>
</tbody>
</table>

Entries are the Spearman correlation coefficient (P value). From a potential list of variables that includes age, the 4 subcategories of the Unified Parkinson’s Disease Rating Scale (UPDRS), and the listed headings, only those associations in which P < .10 during at least 1 walking condition are included. Values are calculated from the first lap using the coefficient of variation of gait; similar results were obtained in the other time segments and using the Pearson correlation coefficient instead. Fall risk and postural instability were determined by questions 13 and 30 of the UPDRS, respectively.

DISCUSSION

Previous investigations demonstrated that gait and balance are altered during dual-task performance in PD. These studies of walking largely focused on the effects on gait speed or the number of steps taken during the walk. However, gait variability, on one hand, and gait speed (and the closely related measure average stride time), on the other, are often affected by different mechanisms. For example, among older adults, gait variability is prospectively associated with fall risk, while gait speed is not. Indeed, in the present study, gait variability was retrospectively associated with fall risk (Table 2). A larger scale, prospective study should help address this issue.

Further study is required to more clearly elucidate the pathophysiological mechanisms that impinge on dual-task performance, in general, and gait variability, more specifically, in PD. In this regard, the association between gait variability and mental function (Table 2) is worth noting. Adequate dopaminergic transmission is apparently necessary for concurrent processing of cognitive information and striatal integration of sensorimotor information required to program cognitive acts. However, stride-to-stride variability of gait in PD may be dopa independent. Some have suggested that depleted central processing resources may impair dual-task performance or the possibility that this represents a deficit in switching processing resources between 2 tasks as the combined demands outweigh available resources. Alternatively, there is a close connection between the basal ganglia and frontal lobe function, and the frontal lobe–like deficits that have been reported in PD may affect dual-task performance.

Further investigations are required to examine if the observed effect of dual tasking on gait variability was age related, if there is a disproportional effect in subjects with PD, over and above a healthy population, and/or how specific the observed response is to PD. It seems likely, however, that aging alone does not explain the sensitivity of gait variability to dual-task performance in PD. On a preliminary basis, we note that dual-task walking did not affect the gait of a 77-year-old woman who did not have PD. Moreover, we note that age was not related to dual performance among the subjects with PD, and even relatively young subjects consistently showed increased stride time and increased variability during dual-task walking.

Other studies suggest that PD pathology contributes to the effect of dual tasking on gait. Camicioli et al observed that dual-task performance did not affect gait speed in healthy older adults. They also found that a verbal fluency task had a much greater effect on gait in subjects with PD compared to aged-matched controls. Similarly, Bond and Morris reported that PD subjects experienced considerable difficulty when they were required to walk while attending to a complex visuomotor task, while aged-matched controls did not. Recent findings by Stolze et al are consistent with the suggestion that the effects of dual tasking and attention on gait may be specific to PD. Thus, previous studies and our preliminary evidence suggest that performance of a second-
ary task has little or no effect on the gait of healthy older adults. Future studies of the effects of dual-task performance on gait variability will be helpful for identifying any role of age. It seems likely, however, that age-related effects are much more subtle than the dramatic response observed and that the more than 5-fold increase in variability is a consequence of PD.

The present study has a number of limitations, perhaps most notably the small number of subjects, the lack of a control group (see above), and the nature of the dual task. Future, larger scale studies should explore the effects of other dual tasks, compare the results to age-matched controls, and study other aspects of gait and motor function (eg, left-right asymmetry, kinematics, dyskinesias; we did not observe any marked increase in dyskinesias during dual-task walking, however, dyskinesias were not quantified). It is possible that there was an ordering effect. All subjects performed dual-task walking after walking under usual conditions. Direct confrontation with the most difficult task could have induced even more gait abnormalities than presently observed. Although we did not quantify serial 7 performance, it might also be helpful in the future to compare performance on the dual task with motor performance. In contrast to subjects with PD, Bloem et al noticed that healthy controls tend to give priority to cognitive tasks over the motor task when attentional resources are divided during dual-task performance. Despite these limitations, the highly consistent results among the PD subjects studied support the idea that in subjects with PD and intact cognitive function, performance of a secondary task increases gait variability.

Further study is also needed to understand why dual-task performance causes a relatively small increase in gait variability in some patients while it almost literally knocks some people off their feet. All subjects were tested while their medications were active (ie, in the “on” state). Subject-to-subject variability in the “on” state may be related to differences in the response to dopamine. Either way, it is important to note that the present results suggest that the clinical features measured in the UPDRS may not explain the degree to which some subjects respond. Perhaps this may also be related to patient-specific differences in the perception of the stimulus provided by serial 7 subtraction. Studies of dual-task effects on balance have shown that both arousal and attention influence balance but differentiate between the specific effects. For some patients, serial 7 subtraction may simply divide attention and provide a distraction. For other patients, an element of arousal/stress may also play a role. Differences in cognitive function not captured by the MMSE, a crude measure of mental function, and variable engagement in the secondary task may also explain subject-to-subject variability. The fact that clinical features evaluated in the UPDRS did not fully explain the degree of the response may suggest that the ability to walk while performing a cognitive task follows an individual course and is influenced by distinct pathways outside the dopaminergic system that affects general motor function. In this regard, it may be helpful to investigate the relationship between sensitive measures of neuropsychological/cognitive function and dual-task walking ability.

Much further work is needed to help reduce falls in PD. Future investigations should confirm the present findings and attempt to more fully identify the factors that modulate the ability of a person to maintain a stable walk while performing a secondary task. The present study provides new insights into PD and gait variability and sets the stage for additional investigations that may help reduce the ubiquity and incapacitating effects of falls in PD. Walking while performing a dual task markedly exacerbates gait variability and impairs the ability of patients with PD to maintain a stable walk. Behavioral or pharmacological interventions designed to minimize the distractions of dual-task walking may be helpful in maintaining gait stability in subjects with PD.

References
Familial Dementia With Lewy Bodies With an Atypical Clinical Presentation

Lauren T. Bonner, MD, Debby W. Tsuang, MD, MSc, Monique M. Cherrier, PhD, Charisma J. Eugenio, BS, Jennifer Q. Du, RN, Ellen J. Steinbart, RN, Pornprot Limprasert, MD, PhD, Albert R. La Spada, MD, PhD, Benjamin Seltzer, MD, Thomas D. Bird, MD, and James B. Leverenz, MD

ABSTRACT

The authors report a case of a 64-year-old male with Alzheimer’s disease (AD) and dementia with Lewy bodies (DLB) pathology at autopsy who did not manifest the core symptoms of DLB until very late in his clinical course. His initial presentation of early executive and language dysfunction suggested a cortical dementia similar to frontotemporal lobar degeneration (FTLD). Core symptoms of DLB including dementia, hallucination, and parkinsonian symptoms were not apparent until late in the course of his illness. Autopsy revealed both brainstem and cortical Lewy bodies and AD pathology. Family history revealed 7 relatives with a history of dementia including 4 with possible or probable DLB. This case is unique because of the FTLD-like presentation, positive family history of dementia, and autopsy confirmation of DLB. (J Geriatr Psychiatry Neurol 2003; 16:59–64)

Keywords: Lewy bodies; familial dementia; Alzheimer’s disease

Dementia with Lewy bodies (DLB) may be the second leading cause of dementia after Alzheimer’s disease (AD).\(^1\) Autopsy studies have demonstrated Lewy bodies (LBs) in the brains of at least 15% to 25% of patients with dementia.\(^2-4\) Approximately 50% to 70% of DLB cases also have sufficient Alzheimer’s pathology to meet criteria for AD.\(^5\)

Clinically, DLB is frequently characterized by dementia, fluctuation in cognition and attention, visual hallucinations, and parkinsonian symptoms.\(^6\) “Cortical” symptoms, such as executive dysfunction and aphasia, occur in more advanced stages of DLB but have not been reported as presenting symptoms.\(^6-9\)

In this report, we describe a case of neuropathologically confirmed DLB, with early executive dysfunction and aphasia with a positive family history of dementia.

CASE PRESENTATION

The proband, a 64-year-old right-handed male aerospace engineer, was referred to a local neurologist for evaluation of a 3-year history of mild depressive symptoms, memory and speech disturbance, and deterioration in handwriting after cardiac bypass surgery. He reported new onset of difficulty recalling names of casual professional and social acquaintances. Although these symptoms were continuous and troublesome to the patient, they were not initially apparent to others. He continued to function at work and at home without substantial difficulties. More recently, his cognitive difficulties had become noticeable to his wife with loss of language and decision-making skills. He had been able to maintain full-time employment as an aerospace engineer. Past medical history was significant for coronary artery
disease and successfully treated pernicious anemia. The neurologist found no signs of parkinsonism on examination. Initial neuropsychological evaluation (see Table 1) revealed constructional dyspraxia and mild difficulties with executive functioning. He was noted to be fluent with superior intellectual functioning and memory that was “high for age” with no recall deficits.

His prior evaluation by his primary care physician had been negative and included serum chemistry, sedimentation rate, VDRL and serum B12. Head computed tomography (CT) at that time demonstrated slight prominence of the sylvian fissure area at the left temporal tip with normal-sized ventricles. The neurologist attributed his cognitive problems to age, whereas the neuropsychologist attributed them to mixed anxiety and depression.

He was subsequently referred to a psychiatrist and underwent a series of treatment trials with antidepressants including selective serotonin reuptake inhibitors and tricyclic antidepressants. However, his cognitive difficulties continued to progress, leading to retirement at age 65.

At age 70, the patient was evaluated by his primary care physician secondary to subjective reports of cognitive decline. The neurological evaluation revealed no parkinsonian signs and symptoms. His primary care physician felt that his cognitive functioning was normal. However, the patient was referred for neuropsychological reassessment in a dementia specialty clinic. During this evaluation, the patient reiterated his complaints of a progressive decline in memory, speech, and cognition with depressed mood. Evaluating neuropsychologists noted he had significant depression with dysphoric mood, early morning awaking, and thoughts of death and euthanasia. On examination (see Table 1), he was aphasic with a Folstein Mini-Mental Status Exam (MMSE) score of 24/30, reflecting difficulties with serial 7s, delayed recall, and inability to correctly copy intersecting pentagons. Examiners noted impairments in executive functioning, language, and verbal memory with significant word-finding difficulties.

Periodic neurological examinations from age 71 to 77 documented continued cognitive decline. Aphasic symptoms included increasingly severe difficulties with word finding, naming, and word comprehension. He also exhibited evidence of apraxia. He developed behavioral disturbances including auditory and visual hallucinations, verbal outbursts, agitation, pacing, and frequent falls. He was treated with multiple psychotropic medications including antipsychotics, antidepressants, anxiolytics, and anticonvulsants, with poor clinical response. The proband maintained some insight into his disorder stating, “My brain is going away and I cannot convert thoughts to words.”

Parkinsonian symptoms including bradykinesia appeared at approximately age 73 and were noted to worsen after treatment with haloperidol with an increase in tremor, rigidity, abnormal gait, and posture.

General neurological examination at age 76, 15 years after the onset of dementia symptoms, noted continued evidence of clinical parkinsonism with cogwheel rigidity; slow, stooped shuffling gait; and positive snout reflex.

### Table 1. Summary of Neuropsychological Evaluations

<table>
<thead>
<tr>
<th>Date of Neuropsychological Evaluation</th>
<th>September 1983</th>
<th>January 1990</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at time of testing (years)</td>
<td>64</td>
<td>70</td>
</tr>
<tr>
<td>Duration of symptoms at time of testing</td>
<td>2 years</td>
<td>8 years</td>
</tr>
<tr>
<td>Cognitive domain/test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intellectual functioning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS (WAIS-R) full-scale IQ</td>
<td>124 (95 percentile)</td>
<td>109 (73 percentile)</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>120 (91 percentile)</td>
<td>104 (81 percentile)</td>
</tr>
<tr>
<td>Performance IQ</td>
<td>124 (95 percentile)</td>
<td>115 (84 percentile)</td>
</tr>
<tr>
<td>Highest subtest</td>
<td>Arithmetic</td>
<td>Picture completion</td>
</tr>
<tr>
<td>Mini-Mental Status Examination</td>
<td>NA</td>
<td>24/30</td>
</tr>
<tr>
<td>Attention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trail-making test, part A</td>
<td>WNL</td>
<td>WNL</td>
</tr>
<tr>
<td>Memory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS/R logical memory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate recall</td>
<td>WNL</td>
<td>Impaired (8 percentile) 2/46</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>WNL</td>
<td>Impaired (8 percentile) 0/46</td>
</tr>
<tr>
<td>Story recall WMS/WMS-R visual reproduction</td>
<td>WNL</td>
<td>WNL (99 percentile) 11/14</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>WNL</td>
<td>WNL (99 percentile) 10/14</td>
</tr>
<tr>
<td>Delayed recall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS/WMS-R visual reproduction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate recall</td>
<td>WNL</td>
<td>WNL (99 percentile) 11/14</td>
</tr>
<tr>
<td>Delayed recall</td>
<td></td>
<td>WNL (99 percentile) 10/14</td>
</tr>
<tr>
<td>Visual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS/WMS-R visual reproduction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate recall</td>
<td>WNL</td>
<td>WNL (99 percentile) 11/14</td>
</tr>
<tr>
<td>Delayed recall</td>
<td></td>
<td>WNL (99 percentile) 10/14</td>
</tr>
<tr>
<td>Construction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aphasia screening test drawings</td>
<td>Mild difficulties</td>
<td>Mild difficulties</td>
</tr>
<tr>
<td>WAIS-R block design subtest</td>
<td></td>
<td>WNL (84 percentile) SS = 13</td>
</tr>
<tr>
<td>Executive functions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trail-making test, part B</td>
<td>Impaired (3 errors) (3 errors)</td>
<td>Impaired (6 to 10 percentile) (3 errors)</td>
</tr>
<tr>
<td>Halstead-Reitan Category Test</td>
<td>WNL (with confusion during subtest 5)</td>
<td>NA</td>
</tr>
</tbody>
</table>

WAIS = Wechsler Adult Intelligence Scale; NA = not administered; WNL = within normal limits; WMS = Wechsler Memory Scale.

a. Details of neuropsychological evaluation were obtained from narrative report. Therefore, individual test scores and percentiles were available only as noted in the report.
b. Test administered but actual score unavailable.
Resting tremor, extraocular movement abnormalities, and other parkinsonian signs were not appreciated. Magnetic resonance imaging and computerized axial tomography revealed mild to moderate cerebral atrophy. Electroencephalogram (EEG) demonstrated diffuse slowing.

His functional status worsened, and his behavioral disturbances including severe agitation and assaultive behavior increased, eventually leading to nursing home placement. Neurological examination at age 78 revealed a severe global aphasia and worsening parkinsonism, although without tremor.

He died at age 79 from acute bronchopneumonia and multiple pulmonary emboli.

**FAMILY HISTORY**

Family history obtained from family members revealed a clinical history of dementia in 7 relatives including his mother and sister (Figure 1).

**First-Degree Relatives**
The proband’s mother (II 4) was approximately 60 years old when she demonstrated difficulties with memory, organization, planning, and visual spatial deficits. She would wander and become lost in her neighborhood on multiple occasions. Approximately 2 to 3 years into the course of her illness, she demonstrated a marked decline in her ability to perform activities of daily living as well as slowing of her movements, difficulty maintaining balance, and urinary incontinence. There was no history of fluctuating mental status. She died at the age of 65.

The proband’s sister (III 7) first manifested symptoms of dementia when she was in her late 50s. Her family reported difficulties with executive functioning and memory with decline in ability to perform activities of daily living. Neurological evaluation at age 61 found severe and continued deterioration in functioning with normal laboratory parameters and head CT demonstrating generalized atrophy. Reexamination at age 62 documented inappropriate affect, with excessive laughing and joking during the examination; conversation was noted to be “rich in puns.” Cognitive evaluation at that time revealed delayed recall of 2 out of 3 items after 5 minutes. On neurological examination, she demonstrated parkinsonian symptoms with rigidity, cogwheeling, and shuffling gait. She was also noted to have a positive glabellar, bilateral grasp, slight Babinski, and poor upward gaze. Head CT demonstrated severe cortical atrophy, normal lumbar puncture, and EEG slowing of background activity.

Her dementia continued to worsen with cognitive and functional decline, auditory and visual hallucinations, and worsening parkinsonian symptoms. She subsequently developed language impairments, and her mental status fluctuated from episodes of coherence to periods of bizarre behavior. Terminally, she was mute and bedridden, dying at age 65.

**Second-Degree Relatives**
The proband’s maternal aunt (II 1) had memory problems that began in her 70s. She died in her late 70s/early 80s. She did not demonstrate parkinsonism, hallucinations, or bizarre behaviors.
The proband’s maternal aunt (II 2) initially manifested symptoms of dementia during her 60s. Her family reported difficulties with memory, visuospatial and executive functioning. She would frequently wander the neighborhood, becoming lost. As her illness progressed, she became paranoid and developed aphasia and apraxia. Eventually, she was institutionalized at a state mental hospital secondary to extreme combativeness and resistiveness to care. At the time of death, she was virtually mute.

The proband’s maternal uncle (II 7) had a history of excessive alcohol intake. His dementia symptoms began in his 60s with his family reporting difficulties with memory, apraxia, and visuospatial and executive functioning. Toward the end of his life, he developed parkinsonian symptoms with slow, stooped gait; bradykinesia; and resting tremor. He also manifested symptoms of agitation, paranoia, and visual and auditory hallucinations. Although he continued to use alcohol until his death, his dementia and behavioral symptoms were generally not associated with intoxication or withdrawal.

Third-Degree Relatives
The proband’s cousin (III 2) is 80 years old and reports a 1-year history of increasing difficulty with short-term memory that interferes with daily activities, visuospatial dysfunction with geographic disorientation, and slight impairment of executive functioning. She had a recent Clinical Dementia Rating (CDR) Scale score of 1.0, indicating mild dementia, and Neuropsychiatric Inventory Score (NPI) of 0, both obtained via a telephone interview with a relative. The CDR and NPI are assessments of cognitive and neuropsychiatric symptomatology, respectively. Both are obtained by a semistructured interview with the caregiver and are easily and reliably performed. Neurological examination revealed an MMSE of 27/30, reflecting a 0/3 score for delayed recall. Administration of the Unified Parkinson’s Disease Rating Scale (UPDRS)–Motor Examination revealed a score of 2, reflecting mild bilateral slowing and/or reduction in leg agility.

NEUROPATHOLOGY
At autopsy, the proband’s brain weighed 1225 grams with gross evidence of moderate to severe fronto-temporal atrophy and mild parietal atrophy. There was mild patchy large vessel atherosclerosis. On gross sectioning, there was severe ventriculomegaly and depigmentation of the substantia nigra. Microscopically, Bielschowsky silver staining demonstrated moderate to severe changes of AD. There was moderate to severe senile plaque formation in the cortex and hippocampal formation. Neurofibrillary tangles were moderate to severe in the medial temporal lobe and hippocampus and mild to moderate in association cortices (Braak Stage V). Hematoxylin and Eosin revealed neuronal loss in the substantia nigra with occasional LB inclusions in the remaining neurons (Figure 2a). Immunohistochemistry for alpha-synuclein (LB 509, Zymed) demonstrated widespread inclusions in the substantia nigra, cingulate gyrus, parahippocampal gyrus, and amygdala (Figure 2b). There were immunopositive neurites in these regions and in the CA-2 subfield of the hippocampus. The case fulfilled neuropathologic diagnostic criteria for DLB (neocortical subtype) and high likelihood AD. There was no evidence grossly or microscopically of any vascular lesions.

MOLECULAR GENETICS
Genomic DNA from the proband and 3 family members was screened for alpha-synuclein mutations. The complete genomic sequences encompassing the alpha-synuclein gene is available through Genbank (http://www.ncbi.nlm.nih.gov/Genbank). Polymerase chain reaction amplification of the 5 coding regions (exons) that encoded for the amino acid sequences was conducted. There were no genetic variants in the gene, specifically, the previously reported A53T mutation. In addition, apolipoprotein E genotyping was conducted using standard methods in the University of Washington Alzheimer’s Disease Research Center Genotyping Core (G. Schellenberg, core principal investigator). The proband’s APOE genotype was ε3/ε4.

DISCUSSION
Our index case fulfilled both clinical and pathological criteria for DLB, although his early clinical presentation made the diagnosis difficult. Ultimately, he exhibited 2 of the 3 core DLB symptoms including hallucinations and motor symptoms of parkinsonism. Fluctuations in mental status were not well documented, although his wife reported nighttime confusion. In addition, his history of
repeated falls and exquisite neuroleptic sensitivity support a diagnosis of DLB. However, he did not exhibit any of these DLB symptoms during the first 10 years of his 18-year course. Pathologically, he exhibited both brainstem and cortical LBs consistent with the neocortical category of DLB. As with many DLB patients, he also exhibited concomitant AD pathology with moderate to severe senile plaque and neurofibrillary tangle changes.

The initial clinical presentation of the proband included clinical and neuropsychological evidence of visuospatial, language, and executive dysfunction. Despite early subjective memory complaints, initial neuropsychological testing found superior memory function and argued against a diagnosis of AD. His early language and executive dysfunction, with subsequent development of a severe global aphasia, suggested a cortical disorder similar to frontotemporal lobar degeneration (FTLD). In fact, his clinical course included several core (insidious onset, gradual progression, nonfluent spontaneous speech) and supportive features (age of onset less than 65 years of age, early preservation of word meaning, late mutism, early preservation of social skills, late behavioral changes, late akinesia, and rigidity) of progressive nonfluent aphasia, a proposed clinical subtype of FTLD. One exclusionary item for FTLD present during the initial presentation was clinical and neuropsychological evidence of visuospatial dysfunction.

DLB patients generally exhibit neuropsychological deficits quite similar to that observed in AD except for relatively more severe visuospatial deficits and less severe memory dysfunction. In several studies, language dysfunction has been found to be equivalent in AD and DLB. In our proband, the only early symptom that suggested a diagnosis of DLB, versus FTLD, was the visuospatial dysfunction. It was only in the second decade of his clinical course that the core symptoms of DLB became evident. To the best of our knowledge, this is the first case report of an FTLD-like clinical presentation in a neuropathologically verified DLB case.

The family history of our case suggests an autosomal dominant inheritance pattern with similar symptoms in affected relatives. Although autopsy was not available for other affected family members, they also demonstrated dementia characterized by symptoms of executive dysfunction, with the sister demonstrating additional language impairment. Seven of the proband’s relatives also demonstrated dementia, some characterized by symptoms suggestive of DLB including clinical parkinsonism and hallucinations.

There have been multiple reports of familial parkinsonism with dementia. In the majority of these families, clinical parkinsonism is the predominant presenting symptom with later development of a dementing disorder. Only a few families include individuals that presented with dementia. However, there are no neuropathological examinations available in these affected individuals. Most recently, Tsuang et al and Galvin et al described 2 families with 2 or more individuals with autopsy-proven DLB. Mutations in α-synuclein have also been found in families with clinical parkinsonism and dementia. Pathologically, most of these familial cases have brainstem, limbic, and cortical LB pathology. In this proband, there was no mutation in the α-synuclein gene. Other disorders associated with a familial frontal dementia with parkinsonism include fronto-temporal dementia with parkinsonism associated with chromosome 17 (FTDP-17) and hereditary dysphasic dementia. The former is characterized by t- associated pathology and LBs are not characteristic, while the latter is associated with inconsistent parkinsonism and LB pathology.

This family is unique in that all 7 potentially affected cases presented with dementia and the autopsy of the index case demonstrated cortical and brainstem LB pathol-
ogy. Our index case demonstrates the degree of clinical heterogeneity in DLB. DLB should be considered in the differential diagnosis of demented patients presenting with FTLD-like symptoms, especially if there is early evidence of visuospatial dysfunction. Other disorders characterized by parkinsonian signs and dementia must also be considered in these patients including Parkinson’s disease with dementia,34 progressive supranuclear palsy,34 FTDP-17,32 and AB.34 Further evaluation of this family and a search for additional families with autopsy-proven DLB will help to further elucidate the genetics and pathophysiology of this important dementing disorder.

References