Development of a Functional Measure for Persons With Alzheimer’s Disease: The Disability Assessment for Dementia

Isabelle Gélinas, Louise Gauthier, Maria McIntyre, Serge Gauthier

Key Words: activities of daily living evaluation • dementia • reliability of tests

Objectives. This article describes the development of an assessment of functional disability for use with proxy-respondents of community-dwelling persons who have Alzheimer’s disease as well as a study testing its reliability.

Method. Panels composed of health care professionals and caregivers of persons with Alzheimer’s disease were used to develop the Disability Assessment for Dementia (DAD). Fifty-nine caregivers participated in the refinement of the content and the testing of reliability.

Results. The DAD includes 40 items: 17 related to basic self-care and 23 to instrumental activities of daily living. It demonstrated a high degree of internal consistency (Cronbach’s alpha = .96) and excellent interrater (N = 31, ICC = .95) and test–retest (N = 45, ICC = .96) reliability. In addition, it was found not to have gender bias.

Conclusion. This instrument may help clinicians and caregivers of the population with Alzheimer’s disease make decisions regarding the choice of suitable interventions.

Alzheimer’s disease is the most frequent cause of dementia (Lerner & Whitehouse, 1994; Small et al., 1997). Estimates of the prevalence of Alzheimer’s disease range from 6% to 8% in persons 65 years of age and older, depending on the sample studied or the methodology used (Ritchie & Kildea, 1995; Small et al., 1997). The prevalence of Alzheimer’s disease is expected to rise drastically by the year 2040 (Canadian Study of Health and Aging Working Group, 1994; Evans, 1990).

Usually appearing in late life, Alzheimer’s disease is a progressive deteriorating disease of insidious onset that affects higher mental functions. Motor and sensory functions usually are spared until later stages. Deterioration in functional performance is also a predominant feature of Alzheimer’s disease (American Psychiatric Association, 1994; Small et al., 1997; Teunisse, Derix, & Van Crevel, 1991), and it is included as a criterion for the diagnosis of probable dementia of the Alzheimer’s type by the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Associations (NINCDS-ADRDA) work group (McKhann et al., 1984; Tierney et al., 1988) and the 4th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria for a dementia of the Alzheimer’s type (American Psychiatric Association, 1994). Decline in functional abilities has an important impact on the quality of life of the person with Alzheimer’s disease (Albert et al., 1996). Instruments for measuring functional performance in the population with Alzheimer’s disease are
needed in order to monitor disease progression, make decisions regarding care, and plan intervention strategies. Instruments would also be helpful for evaluating interventions, conducting research, and making decisions on legal issues like guardianship.

This article describes the development of a functional measure, the Disability Assessment for Dementia (DAD) scale, designed to meet the specific needs of community-dwelling persons who have Alzheimer’s disease and provides evidence of its content validity, internal consistency, and interrater and test–retest reliability. The model of health put forth by the World Health Organization (WHO, 1980) served as the conceptual basis for the development of the DAD. According to this model, functional disability is defined as any restriction in the ability to perform an activity, a task, or a behavior of everyday life, such as basic self-care or instrumental activities.

**Literature Review**

**Changes in Functional Activities of Daily Living in Alzheimer’s Disease**

Changes in functional activities observed in persons with Alzheimer’s disease usually appear insidiously, and abilities progressively deteriorate over several years. Functional decline appears to occur in a hierarchical pattern, from the loss of complex occupational tasks (e.g., work, hobbies) early in the disease process to more difficulty in performing instrumental activities (e.g., dealing with finances, shopping) to deterioration in single overlearned tasks (e.g., basic self-care activities) that appear later in the course of the disease and continue to deteriorate over time. This observed characteristic pattern of progressive deterioration was empirically tested by Sclan and Reisberg (1992) in a study of 56 persons with Alzheimer’s disease. The progressive loss of functional ability in a hierarchical pattern also has been substantiated by several longitudinal studies (Carswell & Eastwood, 1993; Galasko et al., 1995; Green, Mohs, Schneider, Aryan, & Davis, 1993; Stern, Hesdaorfer, Sano, & Mayeux, 1990).

Several authors have discussed the origin of these functional changes observed in Alzheimer’s disease. These changes appear to be of multiple origin. Behavior alterations and cognitive deficits—particularly in memory, concentration, praxis, gnosis, and executive functions (e.g., spontaneity, planning and organization, completion of the task, judgment, sequencing, volition)—all have an impact on functional changes in activities of daily living (ADL) (Baum & Edwards, 1993; Borell, 1996; Galasko et al., 1995; Laberge & Gauthier, 1994; Skurla, Rogers, & Sunderland, 1988; Weintraub, 1986). Indeed, when Reed, Jagust, and Seab (1989) evaluated 59 persons with dementia with established measures of cognition and function in ADL, they found that the cognitive measure explained only about one third of the variance in ADL.

**Evaluation of Functional Disability in Alzheimer’s Disease**

Several studies support the need to test functional abilities in ADL separately from mental status when evaluating severity of dementia because functional tests and mental status tests evaluate different functions (Baum, Edwards, & Morrow-Howell, 1993; Reed et al., 1989; Teunisse et al., 1991). Baum et al. (1993) cautioned against the tendency to rely on results of cognitive tests for evaluation and management of patients with dementia. According to these authors, performance in ADL reflects the ability to integrate different cognitive functions in the production of common behaviors. It would seem that functional assessments provide a more concrete and meaningful way to show families that this integration is successful. This is fundamental to the development of adequate interventions (Carswell & Eastwood, 1993).

The most common way to measure functional disability is through the evaluation of basic ADL (BADL) and instrumental ADL (IADL; Kempen & Suurmeijer, 1990). These tasks, as opposed to advanced activities such as work or leisure, are of major concern to health professionals working with elderly persons who have dementia because these tasks are key factors for determining a person’s ability to live alone or the level of care that he or she may require. Although the number and the kind of activities evaluated in each of these areas vary from author to author, all agree that in order to adequately evaluate functional disability in elderly persons, a scale should measure both BADL and IADL (Fitzgerald, Smith, Martin, Freedman, & Wolinsky, 1993; Norstrom & Thorslund, 1991).

Scales developed for elderly persons in general that target the area of physical performance are not appropriate for use with the population with Alzheimer’s disease, whose functional disabilities are related to deficits in mental functions. Specific scales for the group with Alzheimer’s disease should include items representative of the activities of elderly persons in general, as well as activities affected by the disease process, and show progressive disability in Alzheimer’s disease if it occurs (Carswell, Carson, Walop, & Zgola, 1992; Gelinas & Auer, 1996; Sclan & Reisberg, 1992). A suitable measure for guiding intervention would be one that measures whether the person is able to perform the activities and describes how a task is performed (Baum & Edwards, 1993; Carswell et al., 1992).

Several measures of functional status for the elderly person with dementia have been designed, but none have met all of the desirable criteria (e.g., appropriateness of the content, practicality, adequate psychometric properties). Several scales focus on either BADL or IADL, whereas the literature suggests that to evaluate a wide range of disability in Alzheimer’s disease, several activities in both BADL and IADL should be included (Galasko et al., 1995; Green et al., 1993). For example, the Direct Assessment of Functional Status (Loewenstein et al., 1989) and the Activities of Daily Living Situational Test...
primarily evaluate IADL and may not adequately measure functional deficits when a person can no longer perform complex activities, whereas the Psychogeriatric Basic ADL Scale (Laberge & Gauthier, 1994) focuses only on BADL. Some other scales target specific areas of functioning such as feeding (Athlin, Norberg, Axelsson, Moller, & Nordstrom, 1989; Rogers & Snow, 1982), dressing (Beck, 1988), meal preparation (Baum & Edwards, 1993), or driving (Lucas-Blaustein, Filipp, Dungan, & Tune, 1988). Only the Kitchen Task Assessment (Baum & Edwards, 1993) and the Psychogeriatric Basic ADL Scale (Laberge & Gauthier, 1994) indicate which impairments affect performance (Blessed, Tomlinson, & Roth, 1968; Loewenstein et al., 1989; Mahurin, DeBettignies, & Pirizzolo, 1991; Reisberg et al., 1984). Finally, some of these instruments have weak psychometric properties (Blessed et al., 1968; Skurla et al., 1988). Therefore, our purpose was to develop an appropriate assessment of functional disability designed for caregivers of community-dwelling persons who have Alzheimer’s disease.

Method

The DAD was developed in six consecutive stages (see Table 1), using methodology proposed by Spitzer et al. (1981). The first four stages aimed at defining the content of the instrument, whereas Stages 5 and 6 involved the testing of content validity and reliability. In addition, because French and English versions of the instrument were needed, the development and content validation were carried out simultaneously in both French and English, using a forward-backward translation procedure according to methods proposed by Streiner and Norman (1989).

Content Development

Panels composed of experts and caregivers were used at different stages of the study. Experts who had extensive experience in evaluating or treating patients with Alzheimer’s dis-

<table>
<thead>
<tr>
<th>Stage Obtained</th>
<th>DAD</th>
<th>Versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Identification of the domains and item generation</td>
<td>Panel 1: Mail questionnaire on a preliminary version (DAD 1)</td>
<td>DAD 1</td>
</tr>
<tr>
<td>2. Pretest</td>
<td>Panel 2: Meeting on functional impairments</td>
<td>DAD 2</td>
</tr>
<tr>
<td>3. Content validation of DAD</td>
<td>Panel 3: Meeting on cognitive impairments</td>
<td>DAD 2</td>
</tr>
<tr>
<td>4. Tests of internal consistency and item reduction</td>
<td>Administration of DAD 3 to 59 caregivers of persons with Alzheimer’s disease</td>
<td>DAD 3</td>
</tr>
<tr>
<td>5. Psychometric tests of the final version of the DAD: Content validity</td>
<td>Final validation of content through the mail with Panel 4</td>
<td>DAD 4</td>
</tr>
<tr>
<td>6. Psychometric tests of the final version of the DAD: Reliability</td>
<td>Reliability Verification of interrater and test–retest reliability with DAD 4</td>
<td>DAD 4</td>
</tr>
</tbody>
</table>

Note. DAD = Disability Assessment for Dementia.

<table>
<thead>
<tr>
<th>Disabilities</th>
<th>Impairments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-care activities</td>
<td>Initiation</td>
</tr>
<tr>
<td>Hygiene</td>
<td>Planning and organization</td>
</tr>
<tr>
<td>Dressing</td>
<td>Effective performance</td>
</tr>
<tr>
<td>Undressing</td>
<td></td>
</tr>
<tr>
<td>Continence</td>
<td></td>
</tr>
<tr>
<td>Eating</td>
<td></td>
</tr>
<tr>
<td>Instrumental activities</td>
<td></td>
</tr>
<tr>
<td>Meal preparation</td>
<td></td>
</tr>
<tr>
<td>Telephoning</td>
<td></td>
</tr>
<tr>
<td>Going on an outing</td>
<td></td>
</tr>
<tr>
<td>Finance and correspondence</td>
<td></td>
</tr>
<tr>
<td>Medications</td>
<td></td>
</tr>
<tr>
<td>Housework</td>
<td></td>
</tr>
<tr>
<td>Leisure</td>
<td></td>
</tr>
</tbody>
</table>

Table 2

Note. DAD = Disability Assessment for Dementia.
impairments was retained. Information from Panel 1 and Panel 3 was used to determine which impairments would be included in the items (see Table 2). Several cognitive impairments were mentioned by members of Panel 3. When considering which to include, it was important to consider those that could be easily observed by caregivers in this proxy-respondent evaluation. Because caregivers do not have the specialized training of the health care professional, they may not be in a position to evaluate or understand the impact of cognitive abilities such as memory or perceptual integration. Executive functions that are related to cognitive performance in functional activities seemed to be most amenable for evaluation by caregivers.

Decisions regarding scaling format were made after considering feedback from all three panels, the investigator’s expertise, and information provided in the literature. The dichotomous scale (Yes and No) was retained, and although these types of scales are not usually as sensitive to change, it was believed that sensitivity would be adequate because the scale evaluated functional performance according to two spheres organized in a hierarchical fashion. One sphere pertained to deterioration in functional disability where instrumental activities are lost first, followed by self-care activities (Galasko et al., 1995; Sclan & Reisberg, 1992; Stern et al., 1990). Moreover, progressive deterioration in skills is observed even within these activities. For example, in self-care activities, dressing and hygiene are abilities that are lost before eating and continence (Sclan & Reisberg, 1992). The second sphere concerned executive functions where deficits have been found to deteriorate from problems in initiation in early stages to planning and organization and finally to automatic activities (Gauthier, 1988). At the end of this stage, 36 retained items were included in a dichotomous scale, DAD 2, which yielded a continuous global score.

Stage 2. In Stage 2, DAD 2 was pretested with four caregivers of persons who had “probable” Alzheimer’s disease in Stages 3, 4, or 5 of the Global Deterioration Scale (GDS) (Reisberg, Ferris, DeLeon, & Crook, 1982) in order to determine the clarity, completeness and the practicality of the rating scale. Overall, the scale was deemed to be clear, practical, and complete and did not require major changes. The scale was reported to be quick and easy to use, taking an average of 14 min. Caregivers provided extra information and comments in addition to answering the questions. The scale was found to be adequate for evaluating functional disabilities by both raters and caregivers.

Stage 3. In Stage 3, the content of DAD 2 was validated using a fourth panel of experts and caregivers. Panel members (n = 11) were questioned by mail using a structured content validation questionnaire on the adequacy of the domains included in the assessment and on the appropriateness, clarity, completeness, and weighting of each item. They were also consulted on the method of administration, the scaling format, and the ability of the measure to discriminate between participants who were healthy and those with Alzheimer’s disease. This was followed by a 3-hr meeting where these points were discussed. Panel members were asked to reach an agreement on all items. Decisions about whether items were kept, modified, or rejected were determined by whether 51% of the panel members supported it. The same criterion was used regarding scaling format.

Most members from the panel agreed that the DAD 2 scale content was valid to evaluate functional disability in a community-residing population with Alzheimer’s disease and that evaluating functional disability according to executive functions was appropriate and useful. Individual items were found to be important, clear, and complete for a scale aiming at detecting a change in disability. Items on grooming, driving, and taking public modes of transportation were suggested by panel members and considered for addition to the scale. Most of the experts accepted the dichotomous scaling format and agreed that the scale could discriminate between participants who are healthy and those who have Alzheimer’s disease on functional disabilities as well as between persons who are at different stages of the disease. The DAD 2 scale was thus modified to DAD 3, and it became a 46-item questionnaire rather than a 36-item questionnaire.

Stage 4. In this stage, DAD 3 was used to collect data on functional performance of participants with Alzheimer’s disease, from a caregiver sample, to obtain information on item performance. This collection provided information about the need for modification. The data used for analyses at this stage also were collected as part of the test–retest and interrater reliability studies.

Fifty-nine community-residing caregivers of persons with Alzheimer’s disease (35 English, 24 French), recruited from the McGill Center for Studies in Aging and the Alzheimer Society of Montreal, were interviewed at home in the language of their choice by a trained French-speaking or English-speaking rater. To be selected, each had to be a primary caregiver providing daily care to someone with “probable” dementia of the Alzheimer’s type according to the DSM-III-R (3rd ed., revised; American Psychiatric Association, 1987) and the NINCDS-ADRDA criteria (McKhan et al., 1984; Tierney et al., 1988). The 59 caregivers, predominantly women, ranged from 34 years to 84 years of age. Most were spouses of the participants (78%) with a small proportion being children (17%) or others, such as friends (5%). Most caregivers were unemployed (69%) and in good health (80%). They spent a mean of 7 hours a day (SD = 8.6) in direct caregiving contact with the Alzheimer sample.

The participants with Alzheimer’s disease completed the Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) to provide information on cognitive abilities. Information on the GDS stage of the disease, the date of first diagnosis, and presence of neurological or psychiatric conditions was also collected from their physi-
The performance of individual items was evaluated by a fifth panel of experts \((n = 5)\) during a meeting. All panel members had to agree for an item to be removed or modified. Decisions regarding item reduction were based on face validity and results from three types of analyses. These were frequencies of endorsement, tests of internal consistency, and reliability estimates of individual items using the unweighted kappa statistic. Frequency of endorsement is related to the proportion of persons who chose each alternative from an item. Items where one alternative showed very low endorsement rate (lower than .20) or very high endorsement rate (higher than .90–.95), or those frequently rated as nonapplicable, were considered for removal because they provided very little information and lengthened the scale (Steiner & Norman, 1989). The scale was evaluated for internal consistency using three statistical measures: Cronbach’s alpha, item-total correlations, and inter-item correlations. Cronbach’s alpha (Cronbach, 1951) provides an average of all possible split-half reliabilities of a scale and indicates which items may contribute to low reliability. Item-total correlations, which represent the correlation of individual items with the total score of the scale omitting that item, were also produced to make sure that the items did not evaluate different concepts but rather different components of the same concept. Inter-item correlations were determined to identify items that might be redundant or not related to the construct being studied. Unweighted kappa (Cohen, 1960), which is the index of choice for the evaluation of observer agreement with nominal data, was computed to determine test–retest and interrater reliability for individual items. This statistic corrects for agreement expected by chance. This process allowed identification of items that were less reliable and should be considered for deletion to improve global reliability of the scale. Individual items were also examined using multiway tables to identify which ones may be gender specific. Face validity was an important criterion considered in the decision of whether an item would be retained or removed. The criteria used for item elimination are presented in Table 4.

This process led to further modification of the scale. Twenty-two items out of 46 were scrutinized for removal as a result of the statistical analysis. Six items were eliminated because they were too highly correlated to other items \((r > .81)\) and thus were redundant, had low frequencies of endorsement \((< .20)\), had high frequencies of nonapplicable \((> .90)\), or were not important for face validity. One item, which related to the participant’s ability to stay safely at home, was modified because it was not clearly understood by caregivers. As a result of this stage, the final version of the DAD scale (DAD 4) was established, and it became a 40-item questionnaire (see Figure 1).

### Content Validity and Reliability Studies

#### Stage 5

In Stage 5, the content validity of DAD 4 was reverified after the removal of items. The same experts and caregivers who participated in the panel in Stage 3 were contacted by mail for participation in this stage of the study. Panel members \((n = 11)\) were asked whether they agreed or disagreed to each modification suggested in the reduced version of DAD 3. To keep the modification, 51% of the panel had to support it; otherwise, the original item was retained. Following this stage, the final version of the scale—DAD 4—was produced.

#### Stage 6

In Stage 6, test–retest and interrater reliability of DAD 4 were verified. The influence of explanatory variables on the DAD scale also was evaluated. The data collected in Stage 4 with DAD 3 were used for the reliability studies after the six items eliminated for the final version of

---

**Table 3**

Characteristics of Participants With Alzheimer’s Disease and Their Caregivers

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Participants With Alzheimer’s Disease</th>
<th>Caregivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years) (M \pm SD)</td>
<td>68.0 ± 8.2</td>
<td>61.3 ± 10.6</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>51%</td>
<td>32%</td>
</tr>
<tr>
<td>Female</td>
<td>49%</td>
<td>68%</td>
</tr>
<tr>
<td>Language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>French</td>
<td>41%</td>
<td>41%</td>
</tr>
<tr>
<td>English</td>
<td>59%</td>
<td>59%</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>17%</td>
<td>7%</td>
</tr>
<tr>
<td>Secondary</td>
<td>41%</td>
<td>46%</td>
</tr>
<tr>
<td>Postsecondary</td>
<td>42%</td>
<td>47%</td>
</tr>
<tr>
<td>Duration DAT (months) (M \pm SD)</td>
<td>23.7 ± 22.5</td>
<td></td>
</tr>
<tr>
<td>GDS stages ((n = 58))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>39%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>44%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>MMSE (max 30) (M \pm SD)</td>
<td>16.2 ± 8.3</td>
<td></td>
</tr>
</tbody>
</table>

*Note. \(n = 59\) for both groups. DAT = dementia of the Alzheimer’s type; GDS = Global Deterioration Scale (Reisberg, Ferris, DeLeon, & Crook, 1982); MMSE = Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975).*

**Table 4**

Criteria for Item Elimination

<table>
<thead>
<tr>
<th>Statistical Tests</th>
<th>Elimination Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of endorsement (f &lt; .20) or (f &gt; .90)</td>
<td></td>
</tr>
<tr>
<td>Cronbach’s alpha (\alpha &lt; .80)</td>
<td></td>
</tr>
<tr>
<td>Item-total correlations (r &lt; .40) or (r &gt; .80)</td>
<td></td>
</tr>
<tr>
<td>Inter-item correlations (r &lt; .20) or (r &gt; .80)</td>
<td></td>
</tr>
<tr>
<td>Unweighted kappa (\kappa &lt; .40)</td>
<td></td>
</tr>
</tbody>
</table>

*Note. Face validity and gender specificity of items were also considered for item selection.*
the scale (DAD 4) were removed for analysis.

Forty-five caregivers participated in the test–retest reliability study, and 31 caregivers were evaluated for the interrater reliability study. All caregivers who agreed to a second or third interview or both were reevaluated with the DAD by the same rater for the test–retest study and by a second rater for the interrater study. Both interviews were performed within 1 week after the initial interview. This time interval was believed to be appropriate because performance is evaluated over a 2-week period in the DAD. In addition, using this length of time prevented decreased reliability caused by progression of the disease.

Test–retest and interrater reliability were determined with the intraclass correlation coefficient (ICC) (Streiner & Norman, 1989). The ICC for interrater reliability was estimated from a two-way random effects model. The test–retest reliability was estimated from a one-way random effects model. A 95% confidence interval (CI) also was calculated. An ICC of .80 or higher was accepted as showing good reliability. Additional analyses were conducted to determine the presence of gender difference in the scale and the influence of explanatory variables on results obtained on the final version of DAD with multiple regression analysis.

Results

Descriptive and Comparative Analyses on the DAD Scale

The mean score obtained on the DAD for the total group of participants was 56 (SD = 28). Individual scores ranged from 5% to 100%. The distribution was quite symmetrical, with a median of 54 and a skewness of only –.09. Distributions of scores when the scale was divided into self-care (ADL) and instrumental (IADL) subsections are presented in Figure 2. As expected with this population, the distributions were not normal but approximated exponential configurations with marked negative skewness (–.85) for ADL and positive skewness (.42) for IADL. The median was 88 for the ADL and 35 for the IADL scores, showing that the participants with Alzheimer’s disease scored high in self-care activities compared to instrumental activities where the ratings were low. Thus, the participants were not very impaired on basic tasks but were markedly more impaired on instrumental activities. The French and English subgroups did not differ significantly in their scores for each subsection.

Figure 3 shows the distributions of DAD scores across the different stages of the disease for all participants. This figure demonstrates that, as the severity of the disease...
increases, the global score of the DAD decreases, reflecting more disability.

Content Validity

Content validity of the final version of the DAD scale was established by the panel of experts. Panel members unanimously accepted the proposed removal of six items and the modification of one.

Reliability

The 40-item DAD scale demonstrated a high internal consistency. The overall alpha remained very high for raw variables ($r = .957$) and standardized variables ($r = .956$). Item-total correlations remained in the moderate level with coefficients between .20 and .80. Only six items, considered important for face validity, showed correlations below .40. The scale demonstrated very high stability over time (ICC = .96, with a 95% CI of .90, .97) as well as reproducibility between raters (ICC = .95, with a 95% CI of .90, .97).

Factors That May Influence Results on the DAD Scale

The relationship between explanatory variables and the final version of the DAD was examined. The DAD score was found to be significantly correlated only with GDS stage (Pearson’s $r = -.70$) and MMSE (Pearson’s $r = .54$) score ($p < .01$). The DAD scale was not significantly correlated with age, marital status, education, or duration of the disease. When multiple regression with the forward procedure was performed to determine which variables would best explain scores obtained on the DAD, GDS stage was the only one that significantly contributed to the prediction of DAD score (partial $F [3,40] = 6.43, p < .001$). In addition, men and women did not score differently on the DAD scale ($F = 3.49, p > .05$).

Discussion

The intent of this study was to develop a disability assessment for community-residing persons with dementia that would be valid with regard to content and that would demonstrate good reliability over time and consistency among raters. It was also important to create an instrument that would have no gender bias and would be practical to use in clinical and research settings. These objectives were reached.

Several steps were taken for the elaboration of the scale. Different sources, such as information from the literature and the judgment of experts, were used to ensure that the domains of functional disability were represented and that suitable items would be developed. The experts identified five self-care items for the ADL domain, six items for the IADL domain, and one item on leisure activities. Support for including these domains and items is found in the literature. Most of the existing functional disability scales evaluate either ADL or IADL or both. In fact, several studies support the use of these activities for the evaluation of functional disability with the population with dementia. Galasko et al. (1995) and Green et al. (1993) emphasized the importance of including both self-care and instrumental activities to have a scale that will evaluate the severity of disability with populations that are impaired. On the other hand, leisure activities usually are not included in these types of scales. They were found only in the Record of Independent Living from Weintraub (1986) as a general item. This omission may be because of the difficulty in evaluating these types of activities. This item was included in our scale because caregivers and health care professionals from different panels considered that it represented a crucial aspect of living.

Disability in the DAD scale is evaluated according to impairments in order to have a scale that not only would point out problematic activities, but also would identify which aspects of performance are impaired. Impairments in executive functions were retained in the DAD scale because they are directly related to performance in ADL (Winograd, 1993). They are also easier for the caregiver to observe than other cognitive abilities. Problems in judgment and executive functions were recognized by Baum and Edwards (1993) and Skurla et al. (1988) as having an impact on the ability to complete functional activities.

Information from the literature and consultations with experts were also used for determining the method of administration as well as scaling format. Conventionally, functional scales designed for the population with dementia use either direct observation or questionnaire for a proxy respondent; however, the latter seems to be more frequent.
(Blessed et al., 1968; Laberge & Gauthier, 1994; Reisberg et al., 1984; Weintraub, 1986). Myers, Holliday, Harvey, and Hutchinson (1993) evaluated whether performance-based measures of functional status were superior to self-assessment by questionnaire in elderly persons. They found that although one method was not superior to the other, each provided a different perspective and each had its own value. The proxy approach was favored for the DAD scale; it was considered to be the most reliable method of administration in a community setting with the Alzheimer population because lack of insight is a common manifestation of the disease (McGlynn & Kaszniak, 1991). Additionally, the use of proxy-respondent allows continuity over time, particularly in later stages when patients can no longer respond. It is also less time consuming and often more practical. The use of interview with a proxy allows the evaluation of a variety of activities that may not be easily observable because of time constraints and lack of appropriate resources.

The DAD scale fulfills many of the criteria previously described for good evaluation of functional disability in Alzheimer's disease. The scale is based on a recognized conceptual definition of disability from the WHO. The content is valid for evaluating functional disability with patients who have Alzheimer's disease. Its concurrent and construct validity has been determined in a parallel cross-sectional study (McIntyre, 1994). The scale has demonstrated high internal consistency because Cronbach's alpha exceeded the .80 criterion, which is considered satisfactory (Feinstein, 1987). It can therefore be summarized that the scale measures various aspects of functional disability rather than different constructs. The DAD scale also exceeded the established criteria of .80, as indicative of good reliability, for both test–retest and interrater reliability using a reproducibility test statistic.

The scale was not found to have gender bias when the global score or individual items (for 38 out of 40 items) were considered. Issues related to whether a scale has gender bias have not been reported for most of the functional scales for dementia. This is a matter of importance, particularly with the present generation of elderly persons for whom gender-specific activities are prevalent (Lawton & Brody, 1969). These activities are closely linked to learned gender roles and thus pertain mainly to IADL. For example, it is not unusual to find an elderly man who cannot cook or an older woman who has never driven a car or taken care of finances. Lawton and Brody (1969) addressed this issue in their Physical Self-Care and Instrumental Activity of Daily Living Scales, which were designed for the general elderly population. Because the DAD scale was to be free of gender bias and performance based, care was taken to exclude gender-specific questions and to formulate items so that gender bias would be avoided. In addition, the inclusion of a nonapplicable response choice and the conversion of the total score on a percentage prevented penalizing a person who did not perform an activity.

The influence of explanatory variables on scores obtained from the final version of the DAD scale was examined. The scale was found to correlate moderately with MMSE score. Many studies have established a relationship between mental status scores (particularly scores on the MMSE) and functional performance in dementia (Carswell & Eastwood, 1993; Galasko et al., 1995; Reed et al., 1989; Teunisse et al., 1991; Warren et al., 1989). The reported strength of the relationship between these variables varies from low to high. Several studies, however, including this one, indicated a moderate association between these variables (Carswell & Eastwood, 1993; Reed et al., 1989; Warren et al., 1989). The fact that only moderate correlations were obtained between the MMSE and the DAD score in this study is congruent with increasing evidence in the literature showing that although they are related, the course of deterioration between cognition and function is not parallel (Gauthier & Gauthier, 1990; Reisberg, 1986). The DAD scale appears able to capture these differences in deterioration. These results also indicate that the DAD scale evaluates a different concept than cognition. Indeed, adequate performance in ADL involves the ability to integrate many skills that are not evaluated by cognitive tests, such as perception, executive functions, and motivation.

DAD scores were not associated with other variables such as age or education, meaning that the scores were not influenced by age or education. The DAD was highly correlated with GDS stages. Moreover, a linear regression identified GDS stage as being the only factor influencing scores of the DAD scale. The fact that the other variables did not appear in the regression model was not surprising because they were not significantly correlated to the DAD. The progressive loss of functional abilities in Alzheimer's disease according to a hierarchical pattern has been well demonstrated in longitudinal studies (Carswell & Eastwood, 1993; Gauthier & Gauthier, 1990; Green et al., 1993; Stern et al., 1990) and cross-sectional studies (Baum et al., 1993). Furthermore, it has been established that the progressive deterioration in functional abilities followed the changes delineated in the GDS (Sclan & Reisberg, 1992). Therefore, the relationship between GDS stages and scores on the DAD scale indicates that scores on DAD scale change with stages of the disease, allowing determination of patterns of regression in functional disability. Moreover, the scale could be useful in discriminating between persons who are at different levels of disability.

The DAD scale should prove to be useful for clinical practice because it measures the ability to perform basic self-care ADL and IADL. It also evaluates the ability of the person to initiate, plan, and execute each of these activities (see Appendix). The DAD scale is very easy and quick to administer, taking less than 15 min. In addition, it does not need to be used in a particular setting nor does it require...
any special equipment. (It is, however, recommended that the scale be administered to the caregiver alone because the presence of the person with Alzheimer's disease may influence the accuracy of answers.) Another advantage of this scale is that it does not require any particular expertise or extensive training for administration. A user guide is available to ensure proper administration and scoring of the instrument.

It may be argued that a limitation of this study pertains to the sampling procedure used for data collection. The sampling technique was not random but based on volunteer participation, and inclusion was limited to persons living in the Montreal area who had no physical disabilities that could interfere with the performance of ADL. Because data from the study sample were used for making decisions regarding whether an item would be retained or discarded, there could be concerns about whether the items selected (content) would be representative of the domains of functional disability for the larger community-dwelling population with Alzheimer's disease. However, because decisions regarding item reduction were not based solely on the statistical results but also on face validity and information from the experts, we are confident that the items retained are representative of the larger population with Alzheimer's disease. Moreover, the final version of the scale was reevaluated by a panel of experts and found to be content valid.

Another issue that should be addressed is the number of interviews conducted on the same participant. Some caregivers were used for both the Interrater and the test–retest reliability studies and, as a result, were interviewed as many as three times with the disability scale. A possible bias that could have occurred as a result of this procedure is that the caregivers may have familiarized themselves with the scale and remembered answers from one interview to the next. This would have the consequence of falsely increasing the correlation coefficient. However, the fact that the scale contained many items (46) made it difficult for participants to remember questions. Underestimation of the correlation coefficient also could have occurred as a result of this situation because, in some instances, the third interview was done more than 1 week after the initial visit. Changes in the construct under study may have occurred because of the passage of time. To avoid these biases, separate groups of participants could have been used for the two reliability studies, but this was not feasible during the period over which the study was conducted. Another solution might have been to use two raters during the second interview for the Interrater reliability study. However, bias may occur when this method is used because answers from one of the raters may be influenced by questions or clarifications asked by the other rater during the interview. Efforts were made to minimize possible sources of bias when scheduling interviews and during administration.

Conclusion
We developed a content-valid, French and English instrument to evaluate functional disability in community-dwelling persons with Alzheimer's disease through the use of a proxy-respondent. The DAD scale has demonstrated a high degree of reliability over time and across different raters, is practical and easy to administer, and avoids gender bias. The scale, which is now ready for future testing of construct validity and responsiveness, looks promising because it shows changes across different stages of the disease. The uniqueness of the instrument rests in the fact that this will be one of the few functional instruments for Alzheimer's disease available in French and English, thus facilitating multicentered and international studies. It is also one of the few instruments with the Alzheimer's disease population that evaluates not only which activities are problematic, but also which aspects of performance are impaired.

The DAD scale can have a positive impact on geriatric rehabilitation, and on clinical and research activities with the Alzheimer's disease population by helping clinicians and caregivers make decisions regarding the choice of suitable interventions. It will allow the planning of treatment strategies that will take into account the patient's particular cognitive disabilities in relation to his or her daily functioning. It may also guide decision making with regard to the need for home care or institutionalization. As a research tool, it will be useful in describing the functional characteristics of populations with Alzheimer's disease and the course of the disease and as an outcome variable in intervention studies. Future research activities should include longitudinal studies to better understand change in functional ability over time with the population with Alzheimer's disease and to test responsiveness of the instrument. Studies showing the usefulness of the scale for all types of dementia also should be instigated.

Appendix
Components of the Disability Assessment for Dementia (DAD) Scale

Functional disability is measured with the DAD scale through the evaluation of basic, instrumental, and leisure activities. The DAD scale includes:

- Basic activities of daily living (ADL; activities that are important for self-care), which include dressing, hygiene, continence, and eating.
- Instrumental activities of daily living (IADL; activities that are important for maintenance in a specific environment), which include meal preparation, telephoning, housework, taking care of finance and correspondence, going on an outing, taking medications, and the ability to stay safely at home.
- Leisure activities (activities that are beyond self-maintenance and are for the purpose of recreation), which are evaluated in terms of the interest that is shown toward these activities.

Because the objective of the DAD is also to understand the cognitive dimensions of disabilities in ADL, the activities have been subdivided (into initiation, planning and organization, and effective performance) and are evaluated according to executive functions that have shown regression patterns in dementias.
• Initiation consists of the ability to decide or to start an action. This requires spontaneity on the part of the person and must be accomplished at an appropriate moment and place.
• Planning and organization consist of the ability to identify the different components of a task, to be able to structure them in an appropriate sequence, to elaborate a strategy for action, and to be able to prepare the required material before the action. Planning and organization also include the ability to monitor actions during the activity and thus involve problem-solving and decision-making abilities to make appropriate corrections when needed.
• Effective performance consists of the ability to complete an action. The quality of the performance with regard to whether the task is done in a safe and acceptable manner is also an important component (e.g., a person can be able to plan and complete an action but cannot initiate it).

Acknowledgments
We thank all of the caregivers, participants with Alzheimer's disease, and professionals who participated in this study, which was funded by the Alzheimer Society of Canada. This article was based on a thesis by the first author for the PhD degree in Rehabilitation Science, School of Physical and Occupational Therapy, McGill University, Montreal, Quebec.

References


