STATISTICS IN MEDICINE
Statist. Med. 2000; **19**:1393–1400

Measurement of progression in Alzheimer's disease: a clinician's perspective

Douglas J. Gelb*,†

Department of Neurology, University of Michigan Medical School, Ann Arbor, Michigan 48109-0316, U.S.A.

SUMMARY

Patients with Alzheimer's disease (AD) and their families must confront two fundamental truths. First, AD is a uniformly progressive disease that ultimately results in debilitating cognitive impairment. Second, although there is now evidence that some medications may produce transient improvement or possibly even slowing of disease progression, there is currently no way to halt the progression of AD. Consequently, patients and their families consistently ask the following questions: 1. What new management issues can be anticipated, and when? 2. What clinical developments are atypical and merit evaluation for a superimposed problem? 3. Is the current treatment working? These questions can only be answered by referring to the natural course of AD, and specifically, information regarding measures of functional impairment and how they change over time. The information that is currently available on this topic is limited and often embodies implicit assumptions that have not been adequately tested. This information will be reviewed, and directions for future research will be outlined. Copyright © 2000 John Wiley & Sons, Ltd.

1. INTRODUCTION

Progressive cognitive impairment is a fundamental characteristic of Alzheimer's disease (AD). Accurate measurement of progression is essential for the evaluation of therapeutic interventions, and it is also important in many other avenues of research, ranging from studies of clinicopathological relationships to investigations of public health issues. This has prompted the development of a wide variety of assessment instruments, each designed to suit particular research needs. In principle, these instruments could also be extremely valuable in guiding individual patient management, but this application has received relatively little attention. In this review, I will explore the extent to which currently available assessment instruments pertain to the practical management issues that confront patients and their caregivers. I will not evaluate the relative strengths and weaknesses of the instruments as research tools; these have been summarized in several recent reviews [1–5].

* Correspondence to: Douglas J. Gelb, Department of Neurology, University of Michigan Medical School, Ann Arbor, Michigan 48109-0316, U.S.A.
† E-mail: dgelb@umich.edu

Contract/grant sponsor: NIH-NIA; contract/grant numbers: 1-P50 AG08671.

Copyright © 2000 John Wiley & Sons, Ltd.

2. COGNITIVE TESTING

The most direct way to measure severity of dementia is to test cognitive function, either with a comprehensive neuropsychological battery or with one of the abbreviated assessment instruments that have been designed for this purpose [4, 6–11]. Results of these tests are often difficult to apply to practical management issues, however. A patient's ability to copy a drawing of intersecting pentagons does not have obvious implications for caregivers, for example, and the ability to count backwards by seven does not readily translate into a skill that is required in daily activities. Although the practical relevance is more apparent for some of the more fundamental tasks included in the batteries, such as orientation to person, place and time, or the ability to name simple objects, caregivers usually do not require the assistance of a professional to know when patients have deficits in these tasks. At times, fundamental deficits can remain unrecognized because of a patient's ability to compensate for them, or because of denial or lack of insight on the part of caretakers, but these situations are not common. Conversely, formal cognitive testing can sometimes identify intellectual abilities that are relatively preserved and that can be incorporated by patients and caregivers into compensatory strategies, but it may be just as reasonable to try those strategies empirically and abandon the ones that are not helpful.

Even though most of the items on standard cognitive assessment batteries are not directly relevant to significant management issues, the items might still have clinical value if it could be shown that they function as reliable surrogates for important practical abilities. For example, all patients with AD will eventually reach a point where they do not have adequate visuospatial skills, motor sequencing abilities, reaction times, judgement, or reasoning skills to drive safely, but it can be very difficult to determine precisely when they have reached that point [12–21]. Dementia in and of itself does not preclude the ability to drive. Premature restrictions on driving can undermine a patient's mobility and sense of autonomy unnecessarily, yet delayed recognition of serious impairment has obvious safety implications both for the patient and for society. Patients often lack the insight necessary to detect their impaired driving performance, and, indeed, they may retain enough 'instinctive' ability to perform adequately in routine circumstances. Too often, the first convincing evidence that a patient is too impaired to drive safely comes when the patient is involved in one or more accidents. A reliable indicator of a patient's capacity to drive safely would be enormously valuable. For the indicator to be clinically useful, a high positive predictive value would be required; a simple statistical correlation would not be sufficient. In fact, even the mini-mental status examination (MMSE) score has a moderate correlation with driving ability, but the relationship is not strong enough to predict driving performance for any individual [14–21]. Measures of reaction time or visual attention may prove to be reliable indicators of driving safety [14, 19, 20], but this has not yet been established conclusively.

Although driving safety has received the most attention, other important aspects of practical management could conceivably be assessed by using cognitive test instruments [1, 22–24]. In principle, there could even be a surrogate measure of legal competence. While that seems unlikely, surrogate measures of ability to cook, handle finances, or other tasks required to live independently would be extremely helpful to the clinician and caregivers, if cut-off values could be established that distinguished competent from incompetent patients with a high sensitivity and specificity. This would require the identification of a measure that had a much stronger correlation with the clinical variable of interest than has been established for any measure to date [25], or else the identification through multivariate analysis of a combination of measures that had high predictive power.

Copyright © 2000 John Wiley & Sons, Ltd.

Even if static cognitive measures did not correlate with individual performance of concrete tasks, it is possible that the rate of change of those measures might. Another potential application of the rate of change could be the identification of individual patients with a rate of disease progression that was faster or slower than expected [1, 26–29]. This could be helpful in evaluating response to a treatment, or in recognizing superimposed disease processes. To be used in this way, a cognitive measure would need to change in a reliable way as AD progresses, with very little variation between patients in the rate of progression [26, 30].

3. GLOBAL ASSESSMENT INSTRUMENTS

A single measure that roughly summarizes a patient's overall level of impairment can be used to screen candidates for clinical research, to group patients when analysing data, or to assess the outcome of an intervention [2, 4, 31–35]. The value of a global measure in routine clinical practice is less apparent, however. What is to be gained by knowing whether a patient has 'mild', 'moderate', or 'severe' disease? Although caregivers often seem to appreciate having a frame of reference, their response to the information may be as much a function of their own outlook as anything else. Some caregivers will be reassured and comforted by the knowledge that a patient is 'only stage 1', for example, whereas others will be depressed by the realization that things can only get worse. Conversely, some caregivers will be disheartened when told that a patient is severely impaired, while others will see it as a validation of the burden they have been experiencing.

There are at least two ways in which global measures could conceivably be useful in clinical practice. First, if there were a treatment strategy that produced a robust improvement in global function but only in a subgroup of patients, then global assessment instruments would provide an efficient way to decide which patients should continue to receive the treatment. Second, if the severity of AD, as assessed by a global assessment instrument, progressed in an extremely reproducible way, then the rating could be used both for prognostic purposes and to detect departures from the expected rate of progression due to superimposed disease processes or a therapeutic intervention.

Before a global scale can be used to direct patient management, however, the assumptions underlying the scale must be considered. By definition, global scales combine information about patients' performance across a wide range of cognitive domains. The way in which this information is combined necessarily reflects implicit or explicit assumptions about the disease process. As an illustration, consider the Global Deterioration Scale (GDS), a widely used seven-point scale in which each stage is defined by describing representative deficits that a patient at that stage would typically exhibit [31]. For example, a patient at stage 5 (moderately severe cognitive decline) 'may have difficulty counting backward from 40 by 4s or from 20 by 2s'. People at this stage 'require no assistance with toileting or eating but may have some difficulty choosing the proper clothing to wear...'. They 'invariably know their own names and generally know their spouses' and children's names'. This presupposes that the deficits of AD progress in an orderly fashion in parallel across cognitive domains. If there are any patients who are incontinent or unable to remember their own names but who still retain the ability to calculate, it is not clear how those patients should be classified.

One way to address this problem is to assign a separate rating to a patient's performance in each of a variety of domains, and then to combine these ratings into a global score. This approach still presupposes that the deficits within a particular domain progress in a predictable

Copyright © 2000 John Wiley & Sons, Ltd.

order in patients with AD, but this is more a reflection of the complexity of specific tasks than an assumption about the disease itself. An example of this type of instrument is the Clinical Dementia Rating (CDR) [32], in which patients receive a score of 0 (no impairment), 0.5 (questionable impairment), 1 (mild impairment), 2 (moderate impairment), or 3 (severe impairment) in six independently rated categories: memory; orientation; judgement and problem solving; community affairs; home and hobbies; and personal care. As with the GDS, each score is defined by describing representative deficits, but this time each description is confined to deficits within a single domain. The global CDR is computed from the six component scores using an algorithm in which memory is considered the primary score and all others are secondary.

Unlike the GDS, in which assumptions about the progression of deficits across domains are inherent in the structure of the rating system, the structure of the CDR does not intrinsically require any assumptions about disease progression. The *algorithm* for computing the global rating *does* embody such assumptions, however. Indeed, when the CDR was originally published [32], the authors assumed that the secondary scores would always cluster tightly around the memory score. As a result, the algorithm does not permit calculation of an unambiguous global score for many combinations in which the component scores are somewhat dispersed [36].

When the CDR was adopted by the Consortium to Establish a Registry for Alzheimer's Disease (CERAD), the algorithm was presented in a way that eliminated these indeterminate cases. None the less, the new formulation of the algorithm still incorporated the assumption that the component scores would always cluster tightly around the memory score. If this assumption is invalid, then some counterintuitive and undesirable situations can occur [36]. For example, a decrease in a component score can produce an increase in the global score. Indeed, this type of inconsistency occurs in 5 per cent of all near-match pairs (two sets of component scores that are identical except for a single component score) – the set with the *higher* component score is assigned a *lower* global CDR. In an additional 1 per cent of near match pairs, the algorithm produces an anomaly in which the global score increases by two points or more when a component score increases by one point or less. This is not simply a theoretical concern; these undesirable combinations represented 12 per cent of the CDR determinations in one registry of patients with AD, and were also present in the CERAD database [36, 37]. A modified version of the CDR algorithm was subsequently published [38], but it failed to eliminate this type of inconsistency [37].

Alternative algorithms exist. The 'sum of boxes' score [39], which is conceptually equivalent to calculating the mean of the six category scores, avoids logical inconsistencies but abandons the original emphasis on the memory score. Another published algorithm retains the emphasis on the memory score and eliminates logical inconsistencies by defining the global CDR as the median of the six category scores; when the median falls between two category scores, the one closer to the memory score is chosen as the global CDR [36].

Any attempt to develop a single numerical measure to summarize something as complex and multifaceted as the severity of AD inevitably requires assumptions about disease progression. The assumptions underlying a global assessment instrument may not be immediately apparent, but they should be carefully analysed before the instrument is adopted.

4. FUNCTIONAL ASSESSMENT INSTRUMENTS

Functional assessment instruments are attempts to quantify patients' abilities to perform routine practical tasks, including self-care activities (ADLs, or activities of daily living) and tasks that

Copyright © 2000 John Wiley & Sons, Ltd.

require the use of tools (IADLs, or instrumental activities of daily living) [4, 40–47]. Unlike cognitive testing, for which practical clinical relevance is a significant concern, functional assessment measures translate readily into concrete issues of patient management. The main thing that limits the clinical utility of functional assessment scales is the fact that the information is derived almost exclusively from caregivers' reports. Although some ADLs and IADLs can be evaluated directly in a controlled setting [44, 45], this is time-consuming and not necessarily representative of patients' performance in their usual environment. Thus, functional assessment instruments typically require questioning a caregiver. Since a clinician caring for a patient would do this routinely, it is not clear that there is anything to gain by coding the information in terms of the discrete, standardized categories required for a formal assessment instrument. One potential advantage of using the assessment instrument could be that it would ensure that the clinician did not forget to ask about significant practical issues, but this objective could also be met by using a checklist of topics, without constraining the clinician to conform to the response categories of a specific instrument.

A functional assessment instrument could conceivably be useful to a clinician trying to determine whether a patient had improved in response to a medication or some other therapeutic intervention, but it is probably better to use cognitive tests or global scales for this purpose because placebo effects could cast doubt on a measure that is based purely on caregiver reporting. In principle, functional assessment instruments that had predictive power could be useful to the clinician. For example, if there were evidence that a particular pattern on an assessment instrument reliably identified those individuals who would have difficulty with toileting within a year, that would be very useful in counselling caregivers about the need to plan for support services or placement. As with cognitive tests and global scales, a general statistical correlation would not be adequate for this purpose; to be applicable to patient care, the pattern would have to predict with a high degree of confidence when an individual patient would reach a specific endpoint, preferably an endpoint limited to a single functional domain. Some functional assessment scales resemble global rating scales in that they group performance across domains [41, 48]; as with the global scales, the grouping procedure necessarily embodies assumptions about disease progression that could result in misleading predictions.

5. SEVERE IMPAIRMENT BATTERIES

For research purposes, it is important to have measures of disease progression even late in the course, when performance on many clinical, neuropsychological and functional tests tend to show floor effects [49, 50]. By the time these stages are reached, however, it is usually not particularly important for the clinician to have an accurate measure of disease progression. The relevant information can generally be obtained by asking a few simple questions about basic daily functions. Thus, the severe impairment batteries that have been developed have very little clinical utility. The most likely exception would be if one of these batteries were shown to be strongly predictive (on an individual level, not just a group level) for mortality.

6. BEHAVIOURAL RATING SCALES

Behavioural manifestations of AD, such as depression, agitation, hallucinations and delusions, can have profound significance for patient management, but they have typically been ignored in

Copyright © 2000 John Wiley & Sons, Ltd.

the kinds of rating scales discussed above. There have been several recent attempts to develop rating instruments for these aspects of the disease [3, 51–54]. As with the functional assessment instruments, these rely on caregiver reporting, so in principle they could be redundant for clinicians who are already obtaining complete histories from caregivers. In practice, however, clinicians probably tend to ignore many of these behavioural problems unless caregivers think to report them. Moreover, clinicians may not find these issues as intuitive as they find ADLs and IADLs, so the specific response categories provided by a rating instrument are probably useful. In principle, behavioural rating scales could also be clinically useful if there were evidence that a certain pattern of results reliably identified when individual patients were likely to experience a specific outcome.

7. FUTURE DIRECTIONS

It would be very helpful for clinicians to be able to predict when patients with AD are likely to lose the ability to perform specific functions. The instruments that are already available for measuring disease progression might provide one way to accomplish this, but only if specific measures could be shown to have predictive power at an individual level, not just at a group level. An essential first step would be to determine which aspects of AD are most distressing to patients and caregivers, in terms of both emotional and financial burden [54–56]. Functional assessment instruments that optimally address those functions could then be selected as outcome measures, making each behaviour an independent outcome (rather than grouping functions together and incorporating unnecessary and unwarranted assumptions).

It would also be helpful for clinicians to know when their patients with AD deviated significantly from the expected pattern of disease progression. For this purpose, it is essential to use global instruments that do not impose unwarranted assumptions about disease progression. Here, too, general statements about group averages are of limited clinical utility; the ideal must be to identify individual patients who are progressing in ways that are outside the expected range of variation for the population.

ACKNOWLEDGEMENTS

This work was supported in part by the Michigan Alzheimer's Disease Research Center, NIH-NIA grant 1-P50 AG08671.

REFERENCES

1. Panisset M, Stern Y. Prognostic factors. In *Clinical Diagnosis and Management of Alzheimer's Disease*, Gauthier S (ed.). Butterworth-Heinemann: Boston, 1996; 129–139.
2. Rockwood K, Morris JC. Global staging measures in dementia. In *Clinical Diagnosis and Management of Alzheimer's disease*, Gauthier S (ed.). Butterworth-Heinemann: Boston, 1996; 141–153.
3. Weiner MF, Koss E, Wild KV *et al*. Measures of psychiatric symptoms in Alzheimer patients: a review. *Alzheimer Disease and Associated Disorders* 1996; **10**:20–30.
4. Camicioli R, Wild K. Assessment of the elderly with dementia. In *Handbook of Neurologic Rating Scales*, Herndon RH (ed.). Demos Vermande: New York, 1997; 125–160.
5. Ferris SH, Mackell JA, Mohs R *et al*. A multicenter evaluation of new treatment efficacy instruments for Alzheimer's disease clinical trials: overview and general results. *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 2: S1–S12.
6. Folstein MF, Folstein SE, McHugh PR. Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 1975; **12**:189–198.

Copyright © 2000 John Wiley & Sons, Ltd.

7. Kaufman DM, Weinberger M, Strain JJ, Jacobs JW. Detection of cognitive deficit by a brief mental status examination. The Cognitive Capacity Screening Examination, a reappraisal and a review. *General Hospital Psychiatry* 1979; **1**:247–254.
8. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *American Journal of Psychiatry* 1984; **141**:1356–1364.
9. Schwamm LH, Van Dyke C, Kiernan RJ, Merrin EL, Mueller J. The Neurobehavioral Cognitive Status Examination: comparison with the Cognitive Capacity Screening Examination and the Mini-Mental State Examination in a neurosurgical population. *Annals of Internal Medicine* 1987; **107**:486–491.
10. Kokmen E, Naessens JM, Offord KP. A short test of mental status: description and preliminary results. *Mayo Clinic Proceedings* 1987; **62**:281–288.
11. Mohs RC, Knopman D, Petersen RC *et al*. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 2:S13–S21.
12. Dubinsky RM, Williamson A, Gray CS, Glatt SL. Driving in Alzheimer's disease. *Journal of the American Geriatrics Society* 1992; **40**:1112–1116.
13. Drachman DA, Swearer JM. the Collaborative Study Group. Driving and Alzheimer's disease: the risk of crashes. *Neurology* 1993; **43**:2448–2456.
14. Fitten LJ, Perryman KM, Wilkinson CJ *et al*. Alzheimer and vascular dementias and driving. A prospective road and laboratory study. *Journal of the American Medical Association* 1995; **273**:1360–1365.
15. Trobe JD, Waller PF, Cook-Flannagan CA, Teshima SM, Bieliauskas LA. Crashes and violations among drivers with Alzheimer disease. *Archives of Neurology* 1996; **53**:411–416.
16. Fox GK, Bowden SC, Bashford GM, Smith DS. Alzheimer's disease and driving: prediction and assessment of driving performance. *Journal of the American Geriatrics Society* 1997; **45**:949–953.
17. Lundberg C, Johansson K, Ball K *et al*. Dementia and driving: an attempt at consensus. *Alzheimer Disease and Associated Disorders* 1997; **11**:28–37.
18. Dobbs AR. Evaluating the driving competence of dementia patients. *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 1:8–12.
19. Ball K. Attentional problems and older drivers. *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 1:42–47.
20. Duchek JM, Hunt L, Ball K, Buckles V, Morris JC. The role of selective attention in driving and dementia of the Alzheimer type. *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 1:48–56.
21. Johansson K, Lundberg C. The 1994 international consensus conference on dementia and driving: a brief report. *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 1:62–69.
22. Drachman DA, O'Donnell BF, Lew RA, Swearer JM. The prognosis in Alzheimer's disease: 'how far' rather than 'how fast' best predicts the course. *Archives of Neurology* 1990; **47**:851–856.
23. Kraemer HC, Tinklenberg J, Yesavage JA. 'How far' vs 'how fast' in Alzheimer's disease. *Archives of Neurology* 1994; **51**:275–279.
24. Bracco L, Gallato R, Grigoletto F *et al*. Factors affecting course and survival in Alzheimer's disease. A 9-year longitudinal study. *Archives of Neurology* 1994; **51**:1213–1219.
25. Reed BR, Jagust WJ, Seab JP. Mental status as a predictor of daily function in progressive dementia. *Gerontologist* 1989; **29**:804–807.
26. Galasko D, Corey-Bloom J, Thal LJ. Monitoring progression in Alzheimer's disease. *Journal of the American Geriatrics Society* 1991; **39**:932–941.
27. Haxby JV, Raffaele K, Gillette J, Schapiro MB, Rapoport SI. Individual trajectories of cognitive decline in patients with dementia of the Alzheimer type. *Journal of Clinical and Experimental Neuropsychology* 1992; **14**:575–592.
28. Morris JC, Edland S, Clark C *et al*. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part IV. Rates of cognitive change in the longitudinal assessment of probable Alzheimer's disease. *Neurology* 1993; **43**:2457–2465.
29. Locascio JJ, Growdon JH, Corkin S. Cognitive test performance in detecting, staging, and tracking Alzheimer's disease. *Archives of Neurology* 1995; **52**:1087–1099.
30. Salmon DP, Thal LJ, Butters N, Heindel WC. Longitudinal evaluation of dementia of the Alzheimer type: a comparison of 3 standardized mental status examinations. *Neurology* 1990; **40**:1225–1230.
31. Reisberg B, Ferris SH, DeLeon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. *American Journal of Psychiatry* 1982; **139**:1136–1139.
32. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *British Journal of Psychiatry* 1982; **140**:566–572.
33. Knopman DS, Knapp MJ, Gracon SI, Davis CS. The Clinician Interview-Based Impression (CIBI): a clinician's global change rating scale in Alzheimer's disease. *Neurology* 1994; **44**:2315–2321.
34. Clark CM, Ewbank DC. Performance of the Dementia Severity Rating Scale: a caregiver questionnaire for rating severity in Alzheimer disease. *Alzheimer Disease and Associated Disorders* 1996; **10**:31–39.
35. Schneider LS, Olin JT, Doody RS *et al*. Validity and reliability of the Alzheimer's Disease Cooperative Study – Clinical Global Impression of Change. *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 2:S22–S32.

Copyright © 2000 John Wiley & Sons, Ltd.

36. Gelb DJ, St. Laurent RT. Alternative calculation of the global Clinical Dementia Rating. *Alzheimer Disease and Associated Disorders* 1993; **7**:202–211.
37. Gelb DJ, St. Laurent RT. Clinical dementia rating. *Neurology* 1994; **44**:1983.
38. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 1993; **43**:2412–2414.
39. Berg L. Mild senile dementia of the Alzheimer type: diagnostic criteria and natural history. *Mt Sinai Journal of Medicine* 1988; **55**:87–96.
40. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *British Journal of Psychiatry* 1968; **114**:797–811.
41. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *Journal of Gerontology* 1982; **37**:323–329.
42. Fillenbaum GG. Screening the elderly. A brief instrumental activities of daily living measure. *Journal of the American Geriatrics Society* 1985; **33**:698–706.
43. Spector WD, Katz S, Murphy JB, Fulton JP. The hierarchical relationship between activities of daily living and instrumental activities of daily living. *Journal of Chronic Diseases* 1987; **40**:481–489.
44. Loewenstein DA, Amigo E, Duara R *et al*. A new scale for the assessment of functional status in Alzheimer's disease and related disorders. *Journal of Gerontology* 1989; **44**:P114–P121.
45. Mahurin RK, DeBettignies BH, Pirozzolo FJ. Structured assessment of independent living skills: preliminary report of a performance measure of functional abilities in dementia. *Journal of Gerontology* 1991; **46**:P58–P66.
46. Bucks RS, Ashworth DL, Wilcock GK, Siegfried K. Assessment of activities of daily living in dementia: development of the Bristol Activities of Daily Living Scale. *Age and Ageing* 1996; **25**:113–120.
47. Galasko D, Bennett D, Sano M *et al*. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 2:S33–S39.
48. Reisberg B. Functional assessment staging (FAST). *Psychopharmacology Bulletin* 1988; **24**:653–659.
49. Panisset M, Roudier M, Saxton J, Boller F. Severe impairment battery. A neurospsychological test for severely demented patients. *Archives of Neurology* 1994; **51**:41–45.
50. Schmitt FA, Ashford W, Ernesto C *et al*. The Severe Impairment Battery: concurrent validity and the assessment of longitudinal change in Alzheimer's disease. *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 2:S51–S56.
51. Devanand DP, Miller L, Richards M *et al*. The Columbia University Scale for Psychopathology in Alzheimer's disease. *Archives of Neurology* 1992; **49**:371–376.
52. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994; **44**:2308–2314.
53. Tariot PN, Mack JL, Patterson MB *et al*. The Behavior Rating Scale for Dementia of the Consortium to Establish a Registry for Alzheimer's Disease. *American Journal of Psychiatry* 1995; **152**:1349–1357.
54. Kaufer DI, Cummings JL, Christine D *et al*. Assessing the impact of neuropsychiatric symptoms in Alzheimer's disease: the Neuropsychiatric Inventory Caregiver Distress Scale. *Journal of the American Geriatrics Society* 1998; **46**:210–215.
55. Kane RL. Which outcomes matter in Alzheimer disease and who should define them? *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 6:12–17.
56. Gwyther LP. The perspective of the person with Alzheimer disease: which outcomes matter in early to middle stages of dementia? *Alzheimer Disease and Associated Disorders* 1997; **11**, Suppl 6:18–24.

Copyright © 2000 John Wiley & Sons, Ltd.