

CHAPTER

4

A systematic review of IADL scales in dementia:
Room for improvement



Sietske A.M. Sikkes ^{1,2}, Elly S.M. de Lange-de Klerk ¹, Yolande A.L. Pijnenburg ^{2,3}, Philip Scheltens ^{2,3}, Bernard M.J. Uitdehaag ^{1,3}

¹Department of Epidemiology and Biostatistics, ²Alzheimer Center and ³Department of Neurology, VU University Medical Center, Amsterdam, The Netherlands

Journal of Neurology, Neurosurgery and Psychiatry. 2009; 80: 7-12.

ABSTRACT

Background: Instrumental Activities of daily living (IADL) questionnaires can be helpful in diagnosing dementia and are often used for clinical follow-up and treatment evaluation in dementia patients. Despite the large number of questionnaires, their quality has received little attention. **Objective:** To systematically review the measurement properties of all available structured informant-based (I)ADL questionnaires, developed or validated for use in demented patients. **Methods:** A systematic literature search was conducted in MEDLINE, PsycINFO and EMBASE for psychometric articles on (I)ADL questionnaires. In addition, reference lists of all retrieved articles were screened. Standardized criteria were used to assess the quality of the measurement properties. When possible, investigators were contacted to obtain missing information. Two authors independently extracted studies and performed the quality assessment of the questionnaires. **Findings:** Thirty-two articles were selected, covering 12 (I)ADL questionnaires. Information on 52.3% of the quality aspects was not available, 32.4% of the ratings were indeterminate, 8.1% were positive and 7.2% were negative. Out of eight measurement properties, two scales (the DAD and the Bristol ADL) received two positive ratings and were classified as of moderate quality. Five scales (ADL-PI, ADL-IS, B-ADL, CSADL and Lawton IADL) received one positive rating. **Interpretation:** The findings indicate that improvements in and more data on psychometric properties of (I)ADL questionnaires for dementia patients are necessary in order to justify their use.

INTRODUCTION

Functional decline is an essential feature of all dementias and is therefore embedded in the diagnostic criteria for dementia.¹ This decline is commonly assessed using ‘functional ability’ or ‘activities of daily living’ measurement instruments. Activities of daily living can be divided into Basic activities of daily living (BADL) and Instrumental activities of daily living (IADL). BADL are self-maintenance skills such as bathing, dressing and toileting. IADL involve more complex activities, such as preparing a meal, handling finances and shopping.² These instrumental activities generally require a greater complexity of neuropsychological organization and are in consequence more likely to be vulnerable to the early effects of cognitive decline.²⁻⁶ Measuring IADL can therefore be helpful in diagnosing early dementia.⁷⁻¹⁰ Since IADL also gives an indication of patient dependency, it is frequently used for clinical follow-up and to evaluate treatments.^{8,9,11}

Methods to assess IADL comprise self-reported questionnaires, performance-based assessment and informant-based questionnaires. Self-reported questionnaires are difficult to assess in dementia patients since disease insight is frequently impaired.¹²⁻¹⁵ Observation or direct assessment has the advantage of directly obtaining information without relying on self- or informant-report. Nonetheless, a major drawback of this method is the time-consuming aspect of these instruments, with assessment times up to 1.5 hours.^{8,16} Hence, the most common method is the use of informant-based questionnaires.

A large number of these (I)ADL informant-based questionnaires are available and their number is still growing.^{8,17} Despite the widespread use of these questionnaires, little attention has been paid to their quality. For example, in the light of an early diagnosis of dementia, it still remains unclear which of the existing IADL questionnaires might identify people at risk for dementia.^{7,11}

A critical review of IADL questionnaires is therefore timely and needed in view of the expected increase in clinical trials in early Alzheimer’s disease (AD). Here, we provide an overview of all available structured informant-based IADL questionnaires, developed or validated for the use in AD. Additionally, we set out to evaluate the psychometric properties of these questionnaires. Our final aim was to identify questionnaires useful in the identification of early dementia, particularly in young patients.

METHODS

Data sources

Computer-based literature searches were performed in the PubMed (1950-2007), PsycINFO (1887-2007) and Embase (1966-2007) databases, concluding in November 2007. These databases were searched with the search terms Activities of daily living (MeSH), dement*, Alzheimer*, iadl, instrumental adl, instrumental activities of daily living, extended ADL, complex ADL, advanced ADL, functional ability, everyday functioning and activities of daily living. An additional search was conducted with the terms ADL (MeSH) and dementia (MeSH), to ensure no questionnaires were missed. No limits were set in languages. Case reports and clinical trials were excluded. Two authors (EdL-dK and SS) independently screened abstracts and titles to identify those articles relevant to the research question. In addition to the computerized databases, a book with an overview of assessment scales in old age psychiatry was hand-searched.¹⁸ Reference lists of all articles related to the research question were screened and potentially relevant articles were subsequently retrieved and assessed.

Data extraction and data synthesis

Questionnaires aimed at measuring (I)ADL, complex ADL, advanced ADL, functional ability, functional disability or everyday functioning were selected. When the primary measurement aim was otherwise, for example quality of life or general deterioration, the questionnaire was excluded. The questionnaire had to be disease specific, that is developed or validated for use in dementia patients. In addition, the questionnaire had to be structured and informant-based. In consequence, self-report or clinical judgment scales without operationally defined behaviorally anchor points were excluded. Furthermore, the questionnaire had to be developed or validated for use in Western society (Europe or Northern America).

A study was selected when its primary objective was the development or the clinimetric evaluation of an (I)ADL questionnaire. We aimed to identify the original validation article and all subsequent psychometric articles. Studies addressing other objectives, such as treatment evaluation, were excluded to avoid circular reasoning.

All questionnaires meeting the inclusion and exclusion criteria outlined above were incorporated in the review. They were reviewed on relevant psychometric characteristics according to the 'quality criteria for measurement properties of health status questionnaires' developed by Terwee *et al.*¹⁹ Authors were contacted if further information, such as additional unpublished data, was required.

Quality assessment of the questionnaires

The quality of the questionnaires was evaluated on the following eight measurement properties: (1) Content validity, (2) internal consistency, (3) criterion validity, (4) construct validity, (5) reproducibility, (6) responsiveness, (7) floor- and ceiling effects and (8) interpretability. Each aspect was rated as positive, negative or indeterminate, depending on the design, methods and outcomes of the studies. The quality and rating criteria are outlined in appendix III.

In general, a positive rating was given when the study was adequately designed, executed and analyzed, had adequate sample sizes and satisfying results. An indeterminate rating was given when there was an inadequate (description of) design and execution, inadequate methods or analyses, the sample size was too small or there were methodological shortcomings. A negative rating was given when unsatisfactory results were found despite adequate design, execution, methods, analyses and sample size. When information about the relating criteria was lacking, a 'no information available' rating was given.

Two authors (EdL-dK and SS) independently rated the questionnaires and discrepancies were resolved by consensus. Where consensus could not be reached, a third reviewer was consulted.

Content validity

The content validity refers to the way in which items of the questionnaire cover the domain(s) under investigation.²⁰ In order to be able to judge the content validity of a questionnaire, authors should provide information regarding the measurement aim of the questionnaire, the target population, the concepts the questionnaire intends to measure, the process of item selection and the interpretability of the items.¹⁹ The items of the questionnaire should contain relevant items, and to ensure its relevance, the target population should have been involved in the process of item selection to obtain a positive rating on this property.

Internal consistency

Internal consistency concerns the extent to which items in a (sub)scale are correlated, and are as such measuring the same concept.^{19,20} In order to evaluate the internal consistency, authors should have performed a factor analysis to check for subscales in the questionnaire. When subscales are found, the internal consistency should have been tested separately for these subscales. Internal consistency can be determined by calculating Cronbach's alfa.²⁰ A positive rating for this property was obtained when Cronbach's alpha was between 0.70 and 0.95.

Criterion validity

Criterion validity refers to the extent to which scores on a scale relate to another measure of the construct under study, ideally a 'gold standard'.²⁰ A positive rating is given if the correlation with the gold standard is at least 0.70.

Construct validity

When no absolute 'gold standard' is available, the validity must be investigated by means of indirect evidence, such as establishing the construct validity.²¹ A construct is some postulated attribute of people, assumed to be reflected in test performance.²¹ To investigate construct validity, the scores on the questionnaire under study are correlated with scores on other measurement instruments which are known to be related (or not) to the construct under study. Authors should provide hypotheses about the relation between the construct under study and the other constructs in advance and at least 75% of the results should be in correspondence with the hypotheses to receive a positive rating.^{19,20}

Reproducibility

Reproducibility concerns the degree to which repeated measurements in stable patients provide the same results.²² The concept of reproducibility embraces two aspects, namely agreement and reliability. Reliability addresses whether patients can be distinguished from each other, despite measurement error. This can be determined by calculating a reliability parameter, generally an ICC (Intraclass Correlation Coefficient).²³ The ICC for agreement, which reflects both systematic and random differences in test scores, is preferred. A positive rating is given when the ICC is at least 0.70.¹⁹ The second aspect, agreement, reflects the extent to which repeated measurements give the same results. This can be expressed as the standard error of measurement (SEM) or with the limits of agreement of Bland and Altman.^{19,22} In case of high agreement, the measurement error is small. The SEM can be converted into the smallest detectable change (SDC), which is the smallest within-person change, above measurement error. A positive rating is given when the SDC or the limits of agreement are smaller than the minimal important change (MIC).¹⁹ The MIC is the smallest difference in score in the domain of interest which patients perceive as beneficial and would mandate, in the absence of troublesome side effects and excessive costs.¹⁹

Responsiveness

Responsiveness is the ability of a questionnaire to detect a meaningful or clinically important change over time in a clinical state.^{20,24} To detect such a change, the questionnaire should be able to distinguish clinically important change from measurement error. Responsiveness is determined by calculating Guyatt's responsiveness ratio (RR) or the area under the receiver operating characteristics (ROC) curve (AUC). The latter is a measure of the ability to distinguish patients who

have and have not changed, according to an external criterion. A positive rating was given when the RR was at least 1.96 or the AUC was at least 0.70.¹⁹

Floor- and ceiling effects

Floor- or ceiling effects are present if more than 15% of the patients obtain the lowest or highest possible score. In consequence, patients in these upper or lower ends can not be distinguished from each other and change can not be measured. A positive rating was given when these effects were absent.¹⁹

Interpretability

Lastly, the interpretability of the questionnaire is rated. This is the extent to which one can assign qualitative meaning to quantitative scores. One can interpret scores on a questionnaire when information is present concerning what score or change in score is clinically meaningful. Authors received a positive rating when they provided scores of a reference population and relevant subgroups of patients, and when a MIC was defined to enable interpretation of change scores over time.¹⁹

RESULTS

The initial Medline search produced 2104 possible sources referring to (I)ADL in combination with dementia. All abstracts and titles were screened using the inclusion/exclusion criteria. Twenty-three articles relevant to the research question were identified, covering 13 questionnaires. The additional searches in PsycINFO and Embase disclosed one other relevant article. Cross-referencing led to the identification of another six articles. Resources were unavailable to translate two possibly relevant articles (Spanish and French, concerning the Lawton & Brody IADL questionnaire). One questionnaire (the Daily Activities Questionnaire) was excluded because it was developed with item response techniques (instead of classical test theory), and the quality criteria were not suitable for these techniques.^{25,26} For further investigation, 28 articles covering 12 questionnaires were selected. Table 1 provides the full names and abbreviations of the questionnaires.

Description of questionnaires

Table 2 presents an overview of the included questionnaires. The ADCS-ADL, ADCS-ADL-sev, ADL-IS, ADLQ, B-ADL, Blessed DRS, Bristol ADL, CSADL, DAD and IDDD were disease-specific scales for dementia patients.²⁷⁻³⁵ The ADLQ, B-ADL, DAD and IDDD were aimed at community-dwelling dementia patients, in the early stages of the disease. The ADCS-ADL-sev was aimed at patients in later stages of dementia. The ADL-PI was aimed at healthy older people in prevention trials for AD.³⁶ The Lawton IADL was a generic scale, developed for community-resident older people.² This

Table 1. Abbreviations and full names of the identified questionnaires.

	Abbreviation	Full name
1	ADCS-ADL	Alzheimer Disease Cooperative Study Activities of Daily living Inventory
2	ADCS-ADL-sev	Alzheimer Disease Cooperative Study Activities of Daily living Severe
3	ADL-PI	Activities of Daily living Prevention Instrument
4	ADL-IS	The Alzheimer's disease activities of daily living international scale
5	ADLQ	Activities of Daily Living Questionnaire
6	B-ADL	Bayer Activities of Daily Living Scale
7	Blessed DRS	Blessed Dementia Rating scale
8	Bristol ADL	Bristol activities of daily living Scale
9	CSADL	Cleveland Scale for Activities of daily living
10	DAD	Disability assessment for dementia
11	IDDD	Interview for Deterioration in Daily Living Activities in Dementia
12	Lawton IADL	Lawton & Brody Instrumental Activities of Daily Living scale

questionnaire did not meet the inclusion criteria, but was included by exception, because it is currently the most widely applied IADL questionnaire for dementia patients.

Most questionnaires combined Instrumental ADL with Basic ADL, behavior, cognition or social functioning. The Lawton IADL and the ADL-PI were the only scales designed to measure exclusively IADL.²

The measurement aims of the questionnaires were to assess ADL (ADCS-ADL-sev, Blessed DRS, CSADL, DAD, IDDD, Lawton IADL), to detect changes in ADL (ADL-PI, ADL-IS, ADLQ, B-ADL) or both (ADCS-ADL, Bristol ADL). No questionnaires were aimed at the early identification of dementia in younger patients.

Content validity

In the development of most questionnaires, suggestions from experts, investigators or the target population were incorporated. In some validation articles, no information about how the generation and selection of the items took place was found, such as in the Blessed DRS, IDDD and the Lawton IADL articles. For the ADLQ, no information about the target population involvement was provided. Target population was involved in the item selection of the Bristol ADL, ADL-IS, DAD and CSADL. For all these questionnaires except the Bristol ADL, there was also involvement of experts or investigators. No target population was involved in the selection of items for the B-ADL, item selection was performed by experts only.

Internal Consistency

A Cronbach's alpha was reported for the B-ADL, CSADL, DAD, IDDD and Lawton IADL, ranging from 0.78 to 0.98. The only studies in which factor analysis was

Table 2. Descriptions of the questionnaires.

	Name	Function measured	Goal	Assessment	No. of items
1	ADCS-ADL ³⁷	BADL and IADL	To evaluate ADL abilities over time and detect changes in patients with AD	Interview-based questionnaire	23
2	ADCS-ADL-sev ²⁷	BADL and IADL	To assess ADL in moderate to severe AD	Interview-based questionnaire	19
3	ADL-PI ³⁶	IADL and mobility	To rate ADL in prevention of dementia trials	Two versions: one for patient, one for informant	20
4	ADL-IS ²⁸	ADL	An ADL scale for pharmacologic trials in AD	Questionnaire	40
5	ADLQ ^{4,29}	BADL and IADL	To track progression of functional decline over time in dementia.	Questionnaire	28
6	B-ADL ^{30,38-40}	BADL, IADL and cognition	To measure changes in activities of daily living in the early stages of dementia	Questionnaire	25
7	Blessed DRS ^{31,41}	BADL, IADL and behavior	To describe in quantitative terms the degree of intellectual and personality deterioration shown by a dementia patient	Interview-based questionnaire	22
8	Bristol ADL ^{32,42}	BADL and IADL	To provide a baseline assessment of ability of demented subjects and to be sensitive of change	Questionnaire	20
9	CSADL ^{33,43,44}	BADL and IADL	To characterize the functional difficulties of AD patients as they progress from the mild to the more severe stages of dementia	Interview-based questionnaire	47
10	DAD ⁴⁵⁻⁴⁷	BADL, IADL and leisure activities	To evaluate functional disability in community-dwelling persons with AD through the use of a proxy-respondent	Interview-based questionnaire	40
11	IDDD ^{35,48,49}	BADL and IADL	To assess activities of daily living in dementia	Interview-based questionnaire	33
12	Lawton IADL ^{2,50-53}	IADL	To assess the functional ability of older people	Interview-based questionnaire	8

performed prior to the computation of Cronbach's alpha, were studies of the Lawton IADL and the B-ADL scale. For the B-ADL Cronbach's alpha was 0.98, for the Lawton IADL, two subscales were found, with Cronbach's alphas of 0.91 and 0.78.

Criterion validity

Since there is no apparent 'gold standard' available for ADL, most studies did not provide information about criterion validity. For the Blessed DRS, however, the authors suggested the senile plaque count in the brains of (diseased) patients as a gold standard. The correlation with this standard for demented patients was 0.40.

Construct validity

In most studies, other instruments in addition to the (I)ADL questionnaire were administered, most frequently the MMSE. For the ADLQ, hypotheses were presented relating to an instrument measuring both ADL and behavior. It was expected that correlation with ADLQ would be higher for the ADL section than for the behavior section. These hypotheses were confirmed, with a correlation of 0.71 with the ADL section and a correlation of 0.46 for the behavior section. For the Bristol ADL, hypotheses were also formed in advance, which was explained in a personal communication by Dr. Romola S. Bucks (email, 14 May 2008). Hypotheses relating to the magnitude and direction of expected relationships with other instruments were not presented for any other questionnaires.

Reproducibility

Agreement

Information about agreement was only presented for the Blessed DRS. The limits of agreement were presented, but a MIC was not defined. Information regarding agreement was not present for any of the other questionnaires.

Reliability

Reliability was investigated for nine questionnaires. Kappas were computed for the ADLQ and Bristol ADL. Spearman rank-order correlations were presented for the CSADL and B-ADL (personal communication with Dr. Hartmut Lehfeld, email, 8 April 2008). Pearson correlations were presented for the ADL-PI. ICCs were computed for the Blessed DRS, DAD, DAQ, IDDD and Lawton IADL, with values ranging from 0.30 to 0.94. The sample sizes for the Lawton IADL were too small (<50 patients). For the DAD and the Blessed DRS an ICC was obtained with an adequate sample size (ICC=0.95 and ICC=0.30).

Responsiveness

Change over time was addressed in studies of the DAD, ADLQ, ADL-PI, Lawton IADL, Bristol ADL and ADCS-ADL-sev. These studies presented change scores over time only. Several studies also compared change scores with scores on other related measurement instruments. No MIC was defined for any of the questionnaires.

Floor or ceiling effect

Floor or ceiling effects were reported for the CSADL (78% obtained the lowest score), and the Lawton IADL (approximately 20% obtained the highest score). No floor or ceiling effects were present for the B-ADL, as was explained in a personal communication with Dr. Hartmut Lehfeld (email, 8 April 2008). No or unclear information about possible floor or ceiling effects was provided for the remaining questionnaires.

Interpretability

For eight questionnaires, information of different relevant subgroups was available. These subgroups included groups of different dementia severity, healthy elderly, physically impaired, MCI, depression or different age groups. Mean and SD scores were presented in most articles.

Overall quality

Table 3 shows the overall quality assessment of the 12 (I)ADL questionnaires, summarizing each property as good (+), doubtful (?), poor (-) or no information found (NA) according to the criteria of Terwee *et al.*¹⁹ Information on 52.3% of the quality aspects was not available, 8.1% of the ratings were positive, 32.4% of the ratings were indeterminate, and 7.2% were negative. The ADL-PI, ADL-IS, B-ADL, CSADL and Lawton IADL received one positive rating and the DAD and Bristol ADL received two positive ratings.

DISCUSSION

We identified 12 questionnaires for the evaluation of (I)ADL in dementia. None of these questionnaires provided information for all measurement properties. Of the ratings, 8.1% were positive, 7.2% were negative, and 32.4% were indeterminate, often due to inadequate data analyses or sample sizes. Overall, the DAD and the Bristol ADL received the best ratings for their clinimetric properties. Although several questionnaires were aimed at early dementia, no diagnostic questionnaire specific for use in young patients was found.

To our knowledge, no previous study has systematically addressed the psychometric properties of IADL questionnaires for dementia patients using well-defined criteria. Previous reviews described only the content of the scales or a limited number of psychometric properties, without evaluating those using clear criteria.^{8,54-57}

In our study, the DAD and the Bristol ADL received the best ratings, but when considering all eight quality aspects, these questionnaires are only of moderate quality. The rating of the questionnaires was complicated by the lack of psychometric information. When possible, authors were contacted to obtain missing information. Several authors provided additional information, but a large percentage of measurement properties remained unknown or unclear. It was therefore impossible to give a judgment on several important quality criteria such as responsiveness, reproducibility, construct validity and interpretability.

The quality criteria used in this article might be considered rather strict, since no questionnaire received satisfactory ratings on all criteria. For example, an important

Table 3. Summary of the assessment of the measurement properties of all questionnaires measuring IADL in patients with dementia.

Questionnaire	Content validity	Internal consistency	Criterion validity	Construct validity	Reproducibility			Responsiveness	Floor or ceiling effect	Interpretability
					Agreement	Reliability	Reliability			
ADCS-ADL	-	NA	NA	NA	NA	NA	NA	NA	NA	NA
ADCS-ADL-Sev	-	NA	NA	?	NA	NA	?	NA	NA	NA
ADL-PI	+	NA	NA	?	NA	?	?	NA	NA	?
ADL-IS	+	NA	NA	?	NA	NA	NA	NA	NA	?
ADLQ	NA	NA	NA	?	NA	?	?	NA	NA	?
B-ADL	-*	-	NA	?	NA	?*	NA	NA	+	?
Bristol ADL	+	NA	NA	+	NA	?	?*	NA	NA	NA
Blessed DRS	?	NA	?	NA	?	-	NA	NA	NA	?
CSADL	+	?	NA	?	NA	?	NA	NA	-	?
DAD	+	?	NA	?	NA	+	?	?	?	NA
IDDD	NA	?	NA	NA	NA	?	NA	NA	NA	?
Lawton IADL	-(a)/NA(b)	NA(a)/+(b)	NA	?	NA	?	?	-(a)/NA(b)	?	?

Method or result was rated as: + positive; - negative; ? indeterminate; NA no data available, * Data provided by personal communication, (a) Dementia patients, (b) Healthy elderly.

aspect for several psychometric properties was the definition of a minimal important change. Since no minimal important change (MIC) was defined for any of the questionnaires, it was therefore impossible to receive a positive rating on agreement, responsiveness and interpretability. Defining a MIC is one of the most important aspects of an evaluative questionnaire. Since half of the questionnaires are aimed at the evaluation of ADL and measuring change, one would expect more attention for this aspect, especially in view of the use of these scales in clinical trials and the relative weight put on IADL measures as outcome variable by registration authorities. Another example of the strictness of the criteria is the negative rating the Bayer-ADL received for internal consistency. The Cronbach's alpha for this questionnaire was too high according to the quality criteria. A very high Cronbach's alpha is indicative of redundancy of items, but it does not necessarily mean that the internal consistency is not good. Therefore, this negative rating is more indicative of redundant items and the questionnaire could possibly consist of fewer items. We choose these quality criteria because they are the most specific criteria currently available. These criteria clearly pointed out flaws in the existent IADL questionnaires. The methods used are often inadequate and psychometric information is lacking. These findings are in concordance with previous reviews, in which serious limitations of ADL measurement instruments were mentioned.⁵⁵⁻⁵⁷

Making recommendations about the use of IADL questionnaires is extremely difficult because of lack of information. When choosing a questionnaire it is important to bear in mind that a 'good' score for many measurement properties is a start, but in itself not enough. Which questionnaire is best depends on the measurement aims. In general, one may say that responsiveness analyses are of particular importance for evaluative questionnaires. For diagnostic questionnaires, a good reproducibility is important, in order to distinguish individuals from each other. The applicability of the questionnaire, an aspect that has not received attention in this review, can also be of importance in choosing an appropriate questionnaire.

Even though this review indicated several serious limitations in currently available IADL questionnaires, further investigations could result into better psychometric properties. Validation is an ongoing process and can therefore improve over time when new studies are conducted. Questionnaires with a positive rating on content validity (ADL-PI, ADL-IS, Bristol ADL, CSADL and DAD) provided an appropriate justification for the items included in the questionnaire. The DAD also received a positive rating on reliability, and is therefore of particular interest for further study. Our findings indicate that improvements in psychometric properties of these questionnaires are needed in order to justify their use in clinical practice and for research purposes.

ACKNOWLEDGEMENTS

The authors would like to thank Hans Ket, for his suggestions and help with the literature search.

FUNDING

S.A.M. Sikkes is partly funded by Stichting VUmc funds.

COMPETING INTERESTS

PS is associate editor of JNNP but had no role in the review or editorial process.

REFERENCES

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed. Washington, DC: American Psychiatric Publishing Press; 1994.
2. Lawton MP, Brody EM. Assessment of Older People - Self-Maintaining and Instrumental Activities of Daily Living. *Gerontologist* 1969;9:179-186.
3. Peres K, Helmer C, Amieva H, Orgogozo JM, Rouch I et al. Natural history of decline in instrumental activities of daily living performance over the 10 years preceding the clinical diagnosis of dementia: A prospective population-based study. *Journal of the American Geriatrics Society* 2008;56:37-44.
4. Wicklund AH, Johnson N, Rademaker A, Weitner BB, Weintraub S. Profiles of decline in activities of daily living in non-Alzheimer dementia. *Alzheimer Disease and Associated Disorders* 2007;21:8-13.
5. Barberger-Gateau P, Alioum A, Peres K, Regnault A, Fabrigoule C et al. The contribution of dementia to the disablement process and modifying factors. *Dementia and Geriatric Cognitive Disorders* 2004;18:330-337.
6. Njegovan V, Man-Son-Hing M, Mitchell SL, Molnar FJ. The hierarchy of functional loss associated with cognitive decline in older persons. *Journals of Gerontology Series A-Biological Sciences and Medical Sciences* 2001;56:M638-M643.
7. Nygard L. Instrumental activities of daily living: a stepping-stone towards Alzheimer's disease diagnosis in subjects with mild cognitive impairment? *Acta Neurological Scandinavica Suppl* 2003;179:42-46.
8. Desai AK, Grossberg GT, Sheth DN. Activities of daily living in patients with dementia: clinical relevance, methods of assessment and effects of treatment. *CNS Drugs* 2004;18:853-875.
9. Gauthier S, Gelinat I, Gauthier L. Functional disability in Alzheimer's disease. *International Psychogeriatrics* 1997;9 Suppl 1:163-165.
10. Oakley F, Sunderland T. Assessment of motor and process skills as a measure of IADL functioning in pharmacologic studies of people with Alzheimer's disease: a pilot study. *International Psychogeriatrics* 1997;9:197-206.
11. Dubois B, Feldman HH, Jacova C, DeKosky ST, Barberger-Gateau P et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurology* 2007;6:734-746.
12. Graham DP, Kunik ME, Doody R, Snow AL. Self-reported awareness of performance in dementia. *Cognitive Brain Research* 2005;25:144-152.
13. Giovannetti T, Libon DJ, Hart T. Awareness of naturalistic action errors in dementia. *Journal of the International Neuropsychological Society* 2002;8:633-644.
14. Howorth P, Saper J. The dimensions of insight in people with dementia. *Aging & Mental Health* 2003;7:113-122.
15. Onor ML, Trevisiol M, Negro C, Aguglia E. Different perception of cognitive impairment, behavioral disturbances, and functional disabilities between persons with mild cognitive impairment and mild Alzheimer's disease and their caregivers. *American Journal of Alzheimer's Disease and Other Dementias* 2006;21:333-338.
16. Moore DJ, Palmer BW, Patterson TL, Jeste DV. A review of performance-based measures of functional living skills. *Journal of Psychiatric Research* 2007;41:97-118.
17. Feinstein AR, Josephy BR, Wells CK. Scientific and Clinical Problems in Indexes of Functional Disability. *Annals of Internal Medicine* 1986;105:413-420.
18. Burns A, Lawlor B, Craig S. *Assessment Scales in Old Age Psychiatry*. 2nd ed. Martin Dunitz, an imprint of the Taylor & Francis Group plc; 2004.
19. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL et al. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology* 2007;60:34-42.
20. Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use*. 3rd ed. Oxford: Oxford University Press; 2003.

21. Cronbach LJ, Meehl PE. Construct Validity in Psychological Tests. *Psychological Bulletin* 1955;52:281-302.
22. De Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *Journal of Clinical Epidemiology* 2006;59:1033-1039.
23. Mcgraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1996;1:30-46.
24. Guyatt G, Walter S, Norman G. Measuring Change Over Time - Assessing the Usefulness of Evaluative Instruments. *Journal of Chronic Diseases* 1987;40:171-178.
25. Oakley F, Sunderland T, Hill J, Philips S, Makahon R et al. The Daily Activities Questionnaire: A functional Assessment for People with Alzheimer's Disease. *Physical & Occupational Therapy in Geriatrics* 1991;10:67-81.
26. Oakley F, Lai JS, Sunderland T. A validation study of the daily activities questionnaire: an activities of daily living assessment for people with Alzheimer's disease. *Journal of Outcome Measures* 1999;3:297-307.
27. Galasko D, Schmitt F, Thomas R, Jin S, Bennett D. Detailed assessment of activities of daily living in moderate to severe Alzheimer's disease. *Journal of the International Neuropsychological Society* 2005;11:446-453.
28. Reisberg B, Finkel S, Overall J, Schmidt-Gollas N, Kanowski S et al. The Alzheimer's disease activities of daily living international scale (ADL-IS). *International Psychogeriatrics* 2001;13:163-181.
29. Johnson N, Barion A, Rademaker A, Rehkemper G, Weintraub S. The Activities of Daily Living Questionnaire: a validation study in patients with dementia. *Alzheimer Disease and Associated Disorders* 2004;18:223-230.
30. Hindmarch I, Lehfeld H, de Jongh P, Erzigkeit H. The Bayer Activities of Daily Living Scale (B-ADL). *Dementia and Geriatric Cognitive Disorders* 1998;9 Suppl 2:20-26.
31. Blessed G, Tomlinso BE, Roth M. Association Between Quantitative Measures of Dementia and of Senile Change in Cerebral Grey Matter of Elderly Subjects. *British Journal of Psychiatry* 1968;114:797-811.
32. 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.
33. Patterson MB, Mack JL, Neundorfer MM, Martin RJ, Smyth KA et al. Assessment of functional ability in Alzheimer disease: a review and a preliminary report on the Cleveland Scale for Activities of Daily Living. *Alzheimer Disease and Associated Disorders* 1992;6:145-163.
34. Gelinas I, Gauthier L, McIntyre M, Gauthier S. Development of a functional measure for persons with Alzheimer's disease: the disability assessment for dementia. *American Journal of Occupational Therapy* 1999;53:471-481.
35. Teunisse S, Derix MM. [Measurement of activities of daily living in patients with dementia living at home: development of a questionnaire]. *Tijdschrift voor Gerontologie en Geriatrie* 1991;22:53-59.
36. Galasko D, Bennett DA, Sano M, Marson D, Kaye J et al. ADCS Prevention Instrument Project: assessment of instrumental activities of daily living for community-dwelling elderly individuals in dementia prevention clinical trials. *Alzheimer Disease and Associated Disorders* 2006;20:S152-S169.
37. Galasko D, Bennett D, Sano M, Ernesto C, Thomas R et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. The Alzheimer's Disease Cooperative Study. *Alzheimer Disease and Associated Disorders* 1997;11 Suppl 2:S33-S39.
38. Lehfeld H, Reisberg B, Finkel S, Kanowski S, Wied V et al. Informant-rated activities-of-daily-living (ADL) assessments: results of a study of 141 items in the U.S.A., Germany, Russia, and Greece from the International ADL Scale Development Project. *Alzheimer Disease and Associated Disorders* 1997;11 Suppl 4:S39-S44.
39. Erzigkeit H, Lehfeld H, Pena-Casanova J, Bieber F, Yekrangi-Hartmann C et al. The Bayer-Activities of Daily Living Scale (B-ADL): Results from a validation study in three European countries. *Dementia and*

- Geriatric Cognitive Disorders 2001;12:348-358.
40. Folquitto JC, Bustamante SE, Barros SB, Azevedo D, Lopes MA et al. The Bayer: activities of daily living scale (B-ADL) in the differentiation between mild to moderate dementia and normal aging. *Revista Brasileira de Psiquiatria* 2007.
 41. Cole MG. Interrater reliability of the Blessed Dementia Scale. *Canadian Journal of Psychiatry* 1990;35:328-330.
 42. Byrne LM, Wilson PM, Bucks RS, Hughes AO, Wilcock GK. The sensitivity to change over time of the Bristol Activities of Daily Living Scale in Alzheimer's disease. *International Journal of Geriatric Psychiatry* 2000;15:656-661.
 43. Patterson MB, Mack JL. The Cleveland Scale for Activities of Daily Living (CSADL): Its reliability and validity. *Journal of Clinical Geropsychology* 2001;7:15-28.
 44. Mack JL, Patterson MB. An empirical basis for domains in the analysis of dependency in the activities of daily living (ADL): results of a confirmatory factor analysis of the Cleveland Scale for Activities of Daily Living (CSADL). *The Clinical Neuropsychologist* 2006;20:662-677.
 45. Gelinas I, Gauthier L, McIntyre M, Gauthier S. Development of a functional measure for persons with Alzheimer's disease: the disability assessment for dementia. *American Journal of Occupational Therapy* 1999;53:471-481.
 46. Feldman H, Sauter A, Donald A, Gelinas I, Gauthier S et al. The disability assessment for dementia scale: a 12-month study of functional ability in mild to moderate severity Alzheimer disease. *Alzheimer Disease and Associated Disorders* 2001;15:89-95.
 47. Mioshi E, Kipps CM, Dawson K, Mitchell J, Graham A et al. Activities of daily living in frontotemporal dementia and Alzheimer disease. *Neurology* 2007;68:2077-2084.
 48. Bohm P, Pena-Casanova J, Aguilar M, Hernandez G, Sol JM et al. Clinical validity and utility of the interview for deterioration of daily living in dementia for Spanish-speaking communities NORMACODEM Group. *International Psychogeriatrics* 1998;10:261-270.
 49. Teunisse S, Derix MM. The interview for deterioration in daily living activities in dementia: agreement between primary and secondary caregivers. *International Psychogeriatrics* 1997;9 Suppl 1:155-162.
 50. Hokoishi K, Ikeda M, Maki N, Nomura M, Torikawa S et al. Interrater reliability of the Physical Self-Maintenance Scale and the Instrumental Activities of Daily Living Scale in a variety of health professional representatives. *Aging & Mental Health* 2001;5:38-40.
 51. Green CR, Mohs RC, Schmeidler J, Aryan M, Davis KL. Functional decline in Alzheimer's disease: a longitudinal study. *Journal of the American Geriatrics Society* 1993;41:654-661.
 52. Ng TP, Niti M, Chiam PC, Kua EH. Physical and cognitive domains of the Instrumental Activities of Daily Living: Validation in a multiethnic population of Asian older adults. *Journals of Gerontology series A: Biological Sciences and Medical Sciences* 2006;61:726-735.
 53. Tabert MH, Albert SM, Borukhova-Milov L, Camacho Y, Pelton G et al. Functional deficits in patients with mild cognitive impairment: prediction of AD. *Neurology* 2002;58:758-764.
 54. Lindeboom R, Vermeulen M, Holman R, De Haan RJ. Activities of daily living instruments - Optimizing scales for neurologic assessments. *Neurology* 2003;60:738-742.
 55. Demers L, Oremus M, Perrault A, Champoux N, Wolfson C. Review of outcome measurement instruments in Alzheimer's disease drug trials: psychometric properties of functional and quality of life scales. *Journal of Geriatric Psychiatry and Neurology* 2000;13:170-180.
 56. Avlund K. Methodological challenges in measurements of functional ability in gerontological research. A review. *Aging-Clinical and Experimental Research* 1997;9:164-174.
 57. Bavazzano A, Magnolfi SU, Calvani D, Valente C, Boni F et al. Functional evaluation of Alzheimer patients during clinical trials: A review. *Archives of Gerontology and Geriatrics* 1998;26 (6 Suppl), pp. 27-32.