CHAPTER 6

Validation of the Amsterdam IADL questionnaire®, a new tool to measure instrumental activities of daily living in dementia

Sietske A.M. Sikkes [1, 2], Dirk L. Knol [2], Yolande A.L. Pijnenburg [1, 3],
Elly S.M. de Lange-de Klerk [2], Bernard M.J. Uitdehaag [2, 3], Philip Scheltens [1, 3]

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

Submitted

ABSTRACT

Introduction: Interference in everyday functioning by cognitive disturbances is one of the key diagnostic criteria of dementia. Questionnaires measuring instrumental activities of daily living (IADL) are used to measure this interference, but their quality is questionable. A new informant-based tool aimed at detecting incipient dementia was found to have a high internal consistency and test-retest reliability. The aim of the current study was to further investigate the validity of this new IADL questionnaire. **Methods:** Informants of patients who visited the Alzheimer Center of the VU University Medical Center completed the Amsterdam IADL Questionnaire®, [a]. Item response theory was used to estimate individuals' trait levels. Construct validity was tested by correlating estimated trait levels with age, education, measures of global cognition, daily functioning and depression using Pearson's and Kendall's tau correlation coefficients. In addition, differences in estimated trait levels between patients with and without dementia and between patients with early- and late-onset dementia were tested using an independent t-test. **Results:** A total of 206 informants of patients completed the questionnaire. Estimated trait levels correlated medium with the daily functioning scores (tau =-.47 and tau =.44). Patients diagnosed with dementia (n=93) had higher Amsterdam IADL scores than patients without dementia (n=96), Cohen's effect size, d=1.07, t(187)=7.1, p<.001. We found no differences in trait level between early- and late-onset dementia patients. **Conclusion:** Results suggest that the Amsterdam IADL questionnaire® is a reliable and valid instrument in the evaluation of dementia, for both early and late onset patients.

[a] Amsterdam IADL questionnaire is a registered trademark of Alzheimer Center VU University Medical Center, Amsterdam, the Netherlands

INTRODUCTION

Dementia is one of the most common syndromes among elderly. It leads to a progressive loss of cognitive functions, causing interference with social and occupational functioning.[1] This interference in everyday functioning is generally measured using informant-based questionnaires aimed at instrumental activities in daily living (IADL).[2] However, the psychometric quality of these questionnaires has been questioned in several reviews and improvements were recommended.[3-5]

Based on the need for a reliable and valid IADL informant-based questionnaire for detecting incipient dementia, a new questionnaire was recently developed.[6] In a previous study, we showed that the Amsterdam IADL Questionnaire® had good content validity, consisted of a single factor, had a high internal consistency and good test-retest reliability.[6]

Because this is a newly developed questionnaire, further studies into the validity of this questionnaire are necessary. As a 'gold standard' for IADL does not exist, construct validation is the best approach. Construct validity refers to the degree to which the scores of an instrument are consistent with hypotheses, based on the assumption that the instrument validly measures the construct to be measured.[7]

The aim of the current study is to investigate the construct validity of this new IADL questionnaire. We defined the construct being measured as 'complex activities with little automated skills for which multiple cognitive processes are necessary'.[6] In this study, we intended to provide a comprehensive test of validity.

METHODS

All consecutive informants of patients who visited the Alzheimer Center of the VU University Medical Center for the first time between October 2009 and May 2010 completed the Amsterdam IADL Questionnaire®.

All patients underwent a dementia assessment including clinical history, medical and neurological examination, screening laboratory tests, neuropsychological test battery, MRI and EEG. Diagnoses were made in a multidisciplinary consensus meeting. Participants of this meeting were unaware of the Amsterdam IADL results. The diagnosis of dementia was made using the DSM-IV clinical diagnostic criteria.[1] Alzheimer's Disease was diagnosed according to NINCDS-ADRDA criteria [8], vascular dementia according to NINCDS-AIREN criteria [9], Lewy Body dementia according to McKeith criteria [10] and fronto-temporal lobe dementia according to the Neary criteria.[11] Patients diagnosed with dementia were classified as early-onset dementia (<65 years of age) and late-onset dementia (≥65 years of age).

The study was approved by the Ethics Committee of the VU University Medical Center. All patients gave written informed consent and all informants gave oral informed consent.

The Amsterdam IADL Questionnaire®

The Amsterdam IADL Questionnaire® is a disease-specific IADL questionnaire, aimed at measuring IADL problems in incipient dementia. It is a self-administered questionnaire, completed by an informant of the patient. The questionnaire is assessed using a web-survey based internet tool.[12] The questionnaire consists of 70 items, with each item having a five point scale response option. Using the web-survey based internet tool, items were tailored to the individual responses. For example, no detailed questions on computer use were asked if the patient did not use a computer. This assessment method optimizes the individual differences in complex daily activities.[6]

This tailored approach lead to difficulties in using a total score for indicating difficulties in IADL due to missing items. We decided to use an item response theory (IRT) method of scoring. IRT is a powerful statistical framework in which it is assumed that ordered-categorical item responses represent an underlying construct or 'latent trait'.[13] This construct for the new questionnaire is IADL functioning, ranging from ability to disability. IRT is increasingly used in clinical measurement, as it has several advantages over traditional psychometric methods.[14]

Among the advantages of IRT is the ability to deal with missing items. The assumption for this ability is that items are 'missing at random', that is when the missingness is related to other observed data, but not to unobserved data. Another advantage is that trait estimates are relatively more spread out at the distribution tails.[14] This of particular interest in an IADL questionnaire, as skewed distributions with floor and ceiling effects are often found.[3]

In this study, we used the estimated latent trait as indication of IADL impairment.

Construct validity

Standard criteria were used to evaluate construct validity: we used the Consensus-based standards for the selection of health measurement instruments (COSMIN) checklist for construct validity.[15] These criteria were constructed through an international Delphi study.

Construct validity was tested by correlating the Amsterdam IADL estimated trait levels with age, education, measures of global cognition, daily functioning and depression. We hypothesized low correlations (0.0 – 0.2) between Amsterdam IADL scores and age, education and depression. Medium correlations (0.2 – 0.4) were

expected with measures of cognition and medium high correlations (0.4 – 0.7) with measures of daily functioning.

In addition, we investigated differences between patient groups. We hypothesized patients with dementia to have higher estimated trait levels (representing more IADL disability) than patients without dementia. As the Amsterdam IADL Questionnaire® was designed to be useful for all age groups, we hypothesized no differences between early- and late-onset dementia patients. The relationship with caregiver burden was explored without prior hypotheses.

Measure of global cognitive functioning: the Mini Mental State Examination (MMSE) and the Cambridge Cognitive Examination (CAMCOG)

The MMSE was originally developed to differentiate organic from functional disorders.[16] It is nowadays widely used as a screening test for dementia and assesses several areas of cognition; orientation to time and place, registration of three words, attention and calculation, recall of three words, language and visual construction. MMSE scoring ranges from 0 to 30, with higher scores indicating better performance. Internal consistency and test-retest reliability are generally good, ranging from 0.80 to 0.95.[17] Interrater reliability tends to vary though, suggesting differences in scoring or administration between different testers.[17]

We also administered the CAMCOG, a global measure of cognitive functioning. The CAMCOG is part of the Cambridge Examination of Mental Disorders in the Elderly, an interview schedule for the diagnosis and measurement of dementia in the elderly.[18] The CAMCOG consists of brief neuropsychological tests, which assess cognitive functions such as orientation, memory, language, attention, perception, praxis, calculation and executive functioning.[18] The CAMCOG was found to have excellent internal reliability and test-retest reliability.[19] Also, high sensitivity and specificity values for differentiating between demented and non-demented subjects were found.[19]

Measures of daily functioning: the Disability Assessment for Dementia (DAD) and the Clinical Dementia Rating Scale (CDR)

The DAD is a disease-specific interview-based questionnaire aimed at evaluating functional disability in community-dwelling persons with Alzheimer's disease.[20,21] This informant-based interview consists of 40 items, 17 related to BADL and 23 related to IADL. Every item contains a non-applicable answer option. A total weighted score is calculated by adding the items scores and divide these by the number of items answered (excluding the non-applicable answers), with lower scores indicating more ADL and IADL dysfunction. The DAD was reviewed in a systematic review as having reasonable psychometric properties.[3] The DAD interview was conducted by a specialist nurse.

The CDR is an interview-based global rating device for dementia.[22] Six domains of cognitive and functional performance are rated: memory, orientation, judgment & problem solving, community affairs, home & hobbies and personal care.[23] Every domain is rated using ordered categories, representing increased severity of dementia. An overall CDR score is composed, based on the category ratings with memory as the primary category. This overall score also consists of ordered categories, indicating the level of impairment: healthy, questionable dementia, mild dementia, moderate dementia and severe dementia.[22] Interrater reliability was moderate to high.[22,24] A specialist nurse interviewed the informant to compose a CDR score.

Measures of depression: the Geriatric Depression Scale (GDS) and the Center for Epidemiologic Studies Depression scale (CES-D)

The GDS is a self-rated questionnaire, developed to assess depression in old age.[25] Depressive symptoms during the past week are scored using yes/no questions. The original GDS consisted of 30 items and a high internal consistency, good concurrent validity with other depression scales and adequate psychometric properties for the detection of depression were found.[25,26] We used the Dutch abbreviated GDS. Total scores range from 0 to 15, with higher scores indicating more depressive symptoms. A recent study showed a limited reliability of this version compared to other measures of depression in a non-clinical elderly population.[27]

The CES-D is a self-administered questionnaire designed to measure depressive symptomatology in the general population. The questionnaire consists of 20 statements associated with depression. Each statement is scored on the occurrence of the statement in the past week, ranging from 0 (rarely or never) to 3 (mostly or nearly always). The total score ranges from 0 to 60, with higher scores indicating more depressive symptoms. Exploration of the psychometric properties showed a high internal consistency, moderate test-retest correlations and substantial evidence of construct validity.[28] The psychometric properties of the Dutch translation showed favourable results.[27,29,30]

Measure of caregiver burden: Zarit Burden Inventory (ZBI)

Caregiver burden was measured using the ZBI, one of the most commonly used questionnaires for assessing burden experiences by caregivers of dementia patients.[31,32] We used the Dutch short version of the ZBI (obtained through the Mapi Research Trust). The ZBI consists of 22 items and each item is scored on a 5 point scale. The total score ranges from 0 to 88, with higher scores indicating greater caregiver distress.[32] Psychometric properties were investigated and found to be adequate.[33]

Statistical analysis

As indicated previously, we used IRT to estimate trait levels. Several assumptions must be met for this method and in a previous study we showed that the Amsterdam IADL Questionnaire® met two assumptions: unidimensionality and local independence of items.[6]

For questionnaires with polytomous item responses, the graded response model (GRM) is the most commonly used IRT model.[34] In the GRM, each item (i) is described by one item discrimination or slope parameter (α_i) and difficulty or threshold parameters (β_{ij}).[13] In this model the cumulative probability (P^*) of responding in category j or higher on item i for a person with disability θ (the underlying latent variable) is given by

$$P_{ij}^*(\theta) = \frac{\exp[\alpha_i(\theta - \beta_{ij})]}{1 + \exp[\alpha_i(\theta - \beta_{ij})]} .$$

The difficulty parameter β_{ij} represents the trait level necessary to respond in category j or higher with a probability of .50. A higher value of discrimination parameter α_i indicates that the response categories differentiate well among trait levels.[13] From the P^*, the probability of responding in category j is obtained by

$$P_{ij}(\theta) = P_{ij}^*(\theta) - P_{i,j+1}^*(\theta) .$$

It is assumed that the distribution of the person parameter is standard normal (mean=0 and SD=1). We estimated the item discrimination and item difficulty parameters using Mplus. The estimation method was maximum likelihood, a method suitable for handling missing data.

The items have to meet another assumption which will be investigated in the current study: monotonicity. This means that the probability of responding to a given item should increase monotonically with higher scores on the scale.[7] This was investigated using item tests developed by Stone [35] using SAS. Items were considered misfitting if $p<.01$.

Expected a posteriori estimates of the person trait values were subsequently obtained. We used the estimated trait level score in the current study to indicate IADL impairment. Reliability of the estimated trait level was investigated using a reliability coefficient developed by Samejima (1994).[36]

Correlations between these estimated trait levels, age, education, MMSE, CAMCOG, DAD, CDR, GDS and ZBI were investigated using Pearson's or Kendall's tau correlation coefficients as appropriate.

Differences on estimated trait level between patients with and without dementia and patients with early- and late-onset dementia were investigated using independent t-tests.

The significance level was set at $p<.05$, unless indicated otherwise.

RESULTS

A total of 206 informants of patients completed the questionnaire. Patient and informant characteristics can be found in Table 1.

Table 1. Patient and informant characteristics and comparison of informant characteristics between demented and non-demented patients.

	Patients (n=206)	Informants (n=206)			
			Dementia (n = 93)	No dementia (n = 96)	p-value
Age	64 (10)	59 (12)	58 (11)	59 (13)	.67
Female gender	95 (46.1%)	132 (64.1%)	55 (59.1%)	64 (66.7%)	.28
Level of Education [a]	5 (4-6) [b]	5 (5-6) [c]	5 (5-6) [d]	6 (5-6)	.65
Relationship is spouse		159 (77.2%)	77 (82.8%)	67 (69.8%)	.45
Duration of relationship > 10 years		191 (92.7%)	87 (93.5%)	87 (90.6%)	.19
Living together with patient		153 (74.3%)	66 (71.0%)	71 (74.0%)	.73

Data are presented as mean (SD), median (interquartile range) or n (%). [a]Education according to Verhage's classification, ranging from 1 (low) to 7(high), [b]n=203, [c]n=205, [d]n=92. Differences between groups were tested using the independent t-test, Pearson's Chi-square or Mann-Whitney test.

Item discrimination and difficulty parameters can be found in Table 2. The goodness of fit item tests of Stone are also shown in Table 2. All items fitted the model. The reliability of the estimated trait scores was 0.96.

Correlations

The correlations with clinical measures are shown in Table 3. Estimated IADL trait levels correlated highly with DAD and CDR scores. An intermediate correlation was found with the MMSE and CAMCOG. Low correlations were found with patient's age, education and depression of informant. A low, but significant correlation was found with depressive symptoms of the patient (tau =.14, p=.02). All correlations fell within the expected range.

We explored the relationship with ZBI and found an intermediate correlation (tau = .37, p<.001).

Differences between groups

A total of 93 patients were diagnosed with dementia. For 17 patients, diagnosis was postponed and these patients were excluded from this analysis. No differences in informant characteristics between patients with and without dementia were found (Table 1). Informants of patient with early-onset dementia were more often spouse

Table 2. Item parameter estimates and item fit statistic.

Item	Content	α	β_1	β_2	β_3	β_4	Goodness of fit (p-value) Stone
1	Household duties	2.187	-0.430	0.401	1.467	2.441	.63
2	Shop 1	2.131	-0.629	0.408	1.378	2.027	.88
3	Shop 2	1.718	-0.068	0.704	0.862	0.956	.68
4	Shop 3	1.736	-0.338	0.687	1.293	1.332	.47
5	Shop 4	2.375	0.382	1.009	1.767	2.480	.30
6	Cook 1	2.589	-0.338	0.502	0.999	1.271	.63
7	Cook 2	2.106	0.022	0.655	1.192	1.273	.36
8	Cook 3	2.887	-0.178	0.563	1.189	1.310	.45
9	Cook 4	2.088	0.880	1.822	2.508	2.631	.55
10	Repairs	2.674	-0.811	0.218	0.568	0.798	.17
11	Domestic appliances 1	2.388	-0.018	0.979	1.680	2.272	.52
12	Domestic appliances 2	2.405	0.116	0.903	1.343	1.489	.35
13	Domestic appliances 3	3.309	-0.273	0.473	0.633	*	.16
14	Domestic appliances 4	3.409	0.107	0.741	0.822	0.895	.33
15	Domestic appliances 5	2.275	0.767	1.236	1.821	*	.39
16	Domestic appliances 6	2.244	0.738	1.566	1.946	2.454	.36
17	Domestic appliances 7	2.001	0.619	1.270	1.664	1.720	.58
18	Paying 1	2.307	-1.071	-0.003	0.405	1.003	.55
19	Paying 2	2.607	-0.330	0.400	0.823	0.975	.68
20	Paying 3	1.585	0.236	1.141	1.404	1.608	.98
21	Telephone 1	2.277	-0.313	0.728	1.257	1.455	.86
22	Telephone 2	1.797	-0.419	0.545	0.908	1.437	.50
23	Finances 1	2.761	-0.683	0.105	0.406	0.600	.83
24	Finances 2	2.429	-0.338	0.603	0.857	1.151	.25
25	Finances 3	3.022	-0.308	0.382	0.652	0.815	.45
26	Finances 4	3.436	0.281	0.771	1.014	1.299	.08
27	Finances 5	3.854	-0.097	0.514	0.827	1.092	.12
28	Finances 6	2.029	0.295	1.010	1.364	1.744	.86
29	Finances 7	2.722	0.543	1.185	1.566	1.697	.52
30	Finances 8	2.275	0.489	1.340	2.025	2.281	.56
31	Appointments	1.756	-0.641	0.215	0.920	1.392	.32

Table 2. Continued.

Item	Content	α	β_1	β_2	β_3	β_4	Goodness of fit (p-value) Stone
32	Forms	3.463	-0.644	0.104	0.478	0.655	.78
33	Work 1	1.700	-0.962	-0.225	0.045	0.185	.16
34	Work 2	1.395	-0.890	0.300	0.875	1.393	.13
35	Computer 1	2.567	-0.591	0.357	0.812	1.287	.75
36	Computer 2	2.846	-0.337	0.520	0.782	0.997	.83
37	Computer 3	3.146	-0.055	0.553	0.822	0.951	.51
38	Computer 4	4.838	-0.137	0.342	0.655	0.757	.09
39	Computer 5	2.010	0.451	1.295	1.485	1.693	.79
40	Computer 6	3.753	0.034	0.501	0.862	1.059	.20
41	Computer 7	2.512	-0.283	0.495	0.658	*	.96
42	Computer 8	4.498	-0.454	0.195	0.573	0.896	.47
43	Computer 9	5.685	-0.550	-0.040	0.106	*	.15
44	Computer 10	3.322	-0.769	-0.198	-0.011	0.101	.24
45	Computer 11	2.415	-0.195	0.224	0.398	*	.96
46	Operate devices 1	2.377	-0.073	0.972	1.605	2.429	.56
47	Operate devices 2	2.036	0.181	1.424	2.009	2.940	.47
48	Operate devices 3	4.011	-0.347	0.392	0.611	0.651	.20
49	Operate devices 4	4.033	-0.415	0.478	1.050	1.211	.45
50	Operate devices 5	2.403	-0.211	0.569	0.683	0.796	.80
51	Operate devices 6	3.520	-0.472	0.610	0.898	1.814	.65
52	Operate devices 7	2.224	-0.039	0.498	0.731	0.912	.95
53	Instruction manual 1	3.647	-0.780	0.160	0.411	0.612	.16
54	Instruction manual 2	3.579	-0.887	0.260	0.446	0.956	.66
55	Smartphone	4.500	-0.325	0.088	*	*	.40
56	New devices	2.728	-1.071	-0.074	0.193	0.253	.04
57	Play games	1.571	-0.218	0.848	1.285	1.933	.58
58	Booking	3.943	-0.510	0.043	0.300	*	.18
59	Driving 1	2.047	-0.223	0.321	0.646	1.051	.57
60	Driving 2	1.870	-0.211	0.551	1.125	1.243	.22
61	Driving 3	1.776	-0.102	0.834	1.283	1.354	.90
62	Driving 4	3.933	-0.249	0.071	0.327	0.445	.69
63	Driving 5	3.460	-0.493	0.416	0.756	0.971	.66
64	Driving 6	4.773	-0.580	0.352	0.511	0.983	.35
65	Driving 7	2.367	-0.178	0.664	0.953	1.031	.90
66	Public transport	2.840	-0.100	0.689	1.083	1.434	.54
67	Look for things 1	1.606	-1.946	-0.109	0.847	2.172	.87
68	Look for things 2	1.325	-1.515	0.231	1.378	3.546	.42
69	Deal with the unexpected	1.486	-1.496	-0.011	1.063	3.406	.93
70	Medication	2.008	0.057	0.832	1.409	2.003	.96

Person distribution standard normal. * Difficulty parameter could not be estimated due to zero respondents in category.

than informants of patients with late-onset dementia (80.4% versus 63.8%, Chi-square=8.08, df=3, p=.04). Informants of patient groups did not differ on other characteristics.

Table 4 shows the differences in estimated trait level between patients with and without dementia. Patients diagnosed with dementia had higher estimated trait levels than patients without dementia, Cohen's effect size, d=1.07, t(187)=7.1, p<.001. We found no differences in estimated trait level between early- and late-onset dementia patients.

Table 3. Pearson's or Kendall's tau correlation coefficients of Amsterdam IADL estimated trait level and other clinical measures.

		Hypothesized correlations			Found correlations
Measure	Instrument	Direction	Range	N	(95%CI)
Demographic	Age	+	.00 - .20	206	.04 (-.09 - .18)
	Education	-	.00 - .20	203	-.11 (-.22 - -.01)
Everyday functioning	CDR	+	.40 - .70	57	.44 (.27 - .61)
	DAD	-	.40 - .70	135	-.47 (-.56 - -.37)
Cognitive functioning	MMSE	-	.20 - .40	171	-.32 (-.41 - -.23)
	CAMCOG	-	.20 - .40	170	-.33 (-.42 - -.25)
Depression	GDS	+	.00 - .20	153	.13 (.02 - .24)
	CES-D (informant)	+	.00 - .20	76	.03 (-.12 - .18)

CDR Clinical Dementia Rating Scale; DAD Disability Assessment for Dementia; MMSE Mini Mental State Examination; CAMCOG Cambridge Examination of Cognitive disorder in the Elderly; GDS Geriatric Depression Scale; CES-D Center for Epidemiologic Studies Depression scale.

Table 4. Group differences Amsterdam IADL estimated trait levels.

Hypothesized group differences	Found group differences			
	Diagnosis	N	Estimated trait level	p-value
Dementia > no dementia				
	Dementia	93	0.49 (0.81)	
	No dementia	96	-0.44 (0.98)	<.001
Early onset dementia = late onset dementia				
	Early onset	46	0.53 (0.86)	
	Late onset	47	0.46 (0.78)	.693

Data are presented as mean (SD). Differences between groups are tested using an independent t-test.

CONCLUSION

This is the first validation study of the Amsterdam IADL questionnaire®. We demonstrated good construct validity by the correlations with disability interviews and other clinical variables, indicating that the construct IADL was indeed being measured. Differences between groups were as expected.

The need for a new IADL measure is high, as indicated by several recent studies.[3,37] IADL is important for diagnosing patients with dementia and might also be important for identifying patients at risk for dementia. Patients with IADL impairment and mild cognitive impairment (MCI) were found to be at a higher risk of developing dementia.[38,39]

We found low correlations with age and education. Previous studies with IADL measures found a relationship with age, indicating the need for age-adjusted norms for IADL.[40,41] The low correlation with age implies that the Amsterdam IADL Questionnaire® is not strongly influenced by age. This is in concordance with our finding in early- and late-onset dementia patients. The lack of age effects can be seen as an advantage of this questionnaire.

In this study, we used two other measures of everyday functioning; the DAD and CDR. We found medium correlations with these measures. One might expect higher correlations, but it must be noted that these measures are only partly aimed at measuring IADL. The DAD consists of items aimed at BADL and the CDR is a general indication of dementia severity. Our findings contribute to the idea that the new questionnaire is measuring IADL instead of other aspects of dementia. In addition, the difference in administration (interview versus self-administered) also influences correlations, as scores on a scale are not only determined by the attribute being measured, but also by aspects of the measuring process itself.[7]

In the exploratory analysis, we found an intermediate correlation with caregiver burden. This finding correspond to a previous study in which caregiver burden was found to be associated with informant ratings.[40] This relation can be two-sided: problems in IADL may lead to caregiver burden, but it has also been suggested that burden induces an underestimation of patient's actual performance.[42] Studies on the influence of caregiver burden on IADL informant ratings are scarce. Hence further studies are necessary.

To avoid problems with missing item scores, we used an estimated latent trait score. Other IADL questionnaires calculated a weighted score, disregarding differences between items. Consider for example the difficulty of telephone versus computer use and it the drawback of a weighted score becomes clear. The advantage of the estimated latent trait score, is that reliable estimations can be obtained form a limited

number of items.[13] Another advantage is that we correlated this questionnaire with a broad range of clinical measures, both in patients and caregivers.

In conclusion, in this first validation study of the Amsterdam IADL Questionnaire® we demonstrated good construct validity. Studies into the relevance of IADL for diagnosing different types of dementia and MCI will be among the next steps. In addition, studies measuring change over time are envisioned. This is of particular interest for intervention studies, since these studies focus on patients in early stages of the disease and are in need for clinical relevant outcomes.

The Amsterdam IADL Questionnaire® is free for use in all Public Health and not-for-profit agencies, and can be obtained from the authors following a simple registration.

ACKNOWLEDGEMENTS

The authors would like to thank Freek Gillissen and Rolinka Romkes for their help with the development of the Amsterdam IADL Questionnaire® and Sofie Boom and Paul Dekker for their help with the data collection. The development of the Amsterdam IADL Questionnaire® is supported by grants from Stichting VUmc Fonds and Innovatiefonds Zorgverzekeraars.

REFERENCES

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed. Washington DC; Author: American Psychiatric Publishing Press; 1994.
2. 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.
3. Sikkes SAM, de Lange-de Klerk ESM, Pijnenburg YAL, Scheltens P, Uitdehaag BMJ. A systematic review of Instrumental Activities of Daily Living scales in dementia: room for improvement. Journal of Neurology, Neurosurgery & Psychiatry 2009;80:7-12.
4. Avlund K. Methodological challenges in measurements of functional ability in gerontological research. A review. Aging-Clinical and Experimental Research 1997;9:164-174.
5. 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;Suppl 6, 1998, pp. 27-32:27-32.
6. Sikkes SAM, de Lange-de Klerk E, Pijnenburg YAL, Gilissen F, Romkes R et al. A new informant-based questionnaire for instrumental activities of daily living in dementia. Submitted 2011.
7. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 4th ed. Oxford: Oxford University Press; 2008.
8. Mckhann G, Drachman D, Folstein M, Katzman R, Price D et al. Clinical-Diagnosis of Alzheimers-Disease - Report of the Nincds-Adrda Work Group Under the Auspices of Department-Of-Health-And-Human-Services Task-Force on Alzheimers-Disease. Neurology 1984;34:939-944.

9. Roman GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC et al. Vascular Dementia - Diagnostic-Criteria for Research Studies - Report of the Ninds-Airen International Workshop. Neurology 1993;43:250-260.
10. McKeith IG, Galasko D, Kosaka K, Perry EK, Dickson DW et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): Report of the consortium on DLB international workshop. Neurology 1996;47:1113-1124.
11. Neary D, Snowden JS, Gustafson L, Passant U, Stuss D et al. Frontotemporal lobar degeneration - A consensus on clinical diagnostic criteria. Neurology 1998;51:1546-1554.
12. Roelofsma P, Bottema C, Smeets J. Examine: an websurvey tool for research design and analyis. Amsterdam: SLA Press; 2005.
13. Embretson SE, Reise SP. Item Response Theory for Psychologists. Mahwah, NJ: Erlbaum; 2000.
14. Reise SP, Waller NG. Item Response Theory and Clinical Measurement. Annual Review of Clinical Psychology 2009;5:27-48.
15. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Quality of Life Research 2010;19:539-549.
16. Folstein MF, Folstein SE, Mchugh PR. Mini-Mental State - Practical Method for Grading Cognitive State of Patients for Clinician. Journal of Psychiatric Research 1975;12:189-198.
17. Tombaugh TN, Mcintyre NJ. The Mini-Mental-State-Examination - A Comprehensive Review. Journal of the American Geriatrics Society 1992;40:922-935.
18. Roth M, Tym E, Mountjoy CQ, Huppert FA, Hendrie H et al. Camdex - A Standardized Instrument for the Diagnosis of Mental Disorder in the Elderly with Special Reference to the Early Detection of Dementia. British Journal of Psychiatry 1986;149:698-709.
19. Huppert FA, Jorm AF, Brayne C, Girling DM, Barkley C et al. Psychometric properties of the CAMCOG and its efficacy in the diagnosis of dementia. Aging Neuropsychology and Cognition 1996;3:201-214.
20. 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.
21. 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.
22. 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.
23. Morris JC. The Clinical Dementia Rating (Cdr) - Current Version and Scoring Rules. Neurology 1993;43:2412-2414.
24. Rockwood K, Strang D, MacKnight C, Downer R, Morris JC. Interrater reliability of the clinical dementia rating in a multicenter trial. Journal of the American Geriatrics Society 2000;48:558-559.
25. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V et al. Development and Validation of A Geriatric Depression Screening Scale - A Preliminary-Report. Journal of Psychiatric Research 1983;17:37-49.
26. Montorio I, Izal M. The Geriatric Depression Scale: a review of its development and utility. International Psychogeriatrics 1996;8:103-112.
27. van de Rest O, van der Zwaluw N, Beekman ATF, de Groot LCPG, Geleijnse JM. The reliability of three depression rating scales in a general population of Dutch older persons. International Journal of Geriatric Psychiatry 2010;25:998-1005.
28. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Applied Psychological Measurement 1977;1:385-401.
29. Beekman ATF, Deeg DJH, VanLimbeek J, Braam AW, DeVries MZ et al. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): Results from a community-based sample of older subjects in the Netherlands. Psychological Medicine 1997;27:231-235.

30. Beekman AT, van LJ, Deeg DJ, Wouters L, van TW. [A screening tool for depression in the elderly in the general population: the usefulness of Center for Epidemiological Studies Depression Scale (CES-D)]. Tijdschrift voor Gerontologie en Geriatrie 1994;25:95-103.
31. Zarit SH RKB-PJ. Relatives of the impaired elderly; correlates of feelings of burden. Gerontologist 1980;20:649-655.
32. Zarit SH, Zarit JM. The memory and behavior problems checklist and the burden interview. University Park, PA: Penn State Gerontology Center; 1990.
33. Hebert R, Bravo G, Preville M. Reliability, validity and reference values of the Zarit Burden Interview for assessing informal caregivers of community-dwelling older persons with dementia. Canadian Journal on Aging 2000;19:494-507.
34. Samejima F. Estimation of Latent Ability Using A Response Pattern of Graded Scores. Psychometrika Monograph Supplement 1969;34:100.
35. Stone C. IRTFIT-RESAMPLE: A computer program for assessing goodness of fit of items response theory models based on posterior expectations. Applied Psychological Measurement 2004;28:143-144.
36. Samejima F. Estimation of Reliability Coefficients Using the Test Information Function and Its Modifications. Applied Psychological Measurement 1994;18:229-244.
37. Chang YL, Bondi MW, McEvoy LK, Fennema-Notestine C, Salmon DP et al. Global clinical dementia rating of 0.5 in MCI masks variability related to level of functioning. Neurology 2011;76:652-659.
38. Luck T, Luppa M, Angermeyer MC, Villringer A, Konig HH et al. Impact of Impairment in Instrumental Activities of Daily Living and Mild Cognitive Impairment on Time to Incident Dementia Results of the Leipzig Longitudinal Study of the Aged (LEILA75+). Zeitschrift fur Gerontologie und Geriatrie 2010;43:63.
39. Peres K, Chrysostome V, Fabrigoule C, Orgogozo JM, Dartigues JF et al. Restriction in complex activities of daily living in MCI: impact on outcome. Neurology 2006;67:461-466.
40. Jorm AF. The informant questionnaire on cognitive decline in the elderly (IQCODE): a review. International Psychogeriatrics 2004;16:275-293.
41. Sikkes SAM, Visser PJ, Knol DL, de Lange-de Klerk E, Tsolaki M et al. Do instrumental activities of daily living predict dementia at one and two year follow-up? Findings from the DESCRIPA study. Submitted 2011.
42. Mangone CA, Sanguinetti RM, Baumann PD, Gonzalez RC, Pereyra S et al. Influence of feelings of burden on the caregiver's perception of the patient's functional status. Dementia 1993;4:287-293.