Original Research Article

Dementia and Geriatric Cognitive Disorders

Dement Geriatr Cogn Disord 2010;29:363–370
DOI: 10.1159/000297523

Accepted: March 5, 2010
Published online: April 24, 2010

Korean Version of Frontal Assessment Battery: Psychometric Properties and Normative Data

Tae Hui Kim[a] Yoonseok Huh[a] Jin Yeong Choe[a] Ji Woon Jeong[a]
Joon Hyuk Park[b] Seok Bum Lee[c] Jung Jae Lee[d] Jin Hyeong Jhoo[e]
Dong Young Lee[f, g] Jong Inn Woo[f, g] Ki Woong Kim[a, g]

[a]Department of Neuropsychiatry, Seoul National University Bundang Hospital, Seongnam, [b]Department of Neuropsychiatry, Jeju National University Hospital, Jeju, [c]Department of Psychiatry, Dankook University Hospital, Cheonan, [d]Department of Psychiatry, Kyungbook National University Hospital, Daegu, [e]Department of Neuropsychiatry, Kangwon National University Hospital, Chuncheon, [f]Department of Neuropsychiatry, Seoul National University Hospital, and [g]Department of Psychiatry, Seoul National University College of Medicine, Seoul, Korea

Key Words

Frontal Assessment Battery · Korean · Validity · Reliability · Normative data

Abstract

Background: We developed the Korean version of the Frontal Assessment Battery (FAB-K), evaluated its psychometric properties and constructed normative data for Korean elders. ***Methods:*** FAB-K was administered to 300 Alzheimer's disease (AD) patients and 635 normal controls. Reliability of FAB-K was evaluated by testing its internal consistency, test-retest and inter-rater reliabilities. Validity of FAB-K was evaluated by testing discriminant validity for AD and concurrent validity with other frontal function tests. Age- and education-specific normative data of FAB-K were developed. ***Results:*** Cronbach's α, inter-rater reliability and test-retest reliability of FAB-K were 0.802, 0.980 ($p < 0.001$) and 0.820 ($p < 0.001$), respectively. FAB-K exhibited significant correlations with the scores of MMSE and other frontal function tests ($p < 0.01$). Total and item scores of FAB-K were lower in AD patients than in controls and became worse as clinical dementia rating increased (F = 192.026, d.f. = 4, $p < 0.001$). The optimal cut-off score of FAB-K for AD was determined as 10/11, where sensitivity and specificity for AD were 0.717 and 0.827, respectively. Normative data were stratified by 3 age groups and 4 education groups. ***Conclusion:*** The FAB-K is a valid and reliable instrument for evaluating frontal dysfunction, and may be useful for screening AD.

Introduction

Frontal lobe functions are the abilities that enable an individual to engage successfully in autonomous, purposive, self-serving behavior. These abilities include initiation, planning, self-regulation, set maintenance and shifting, and vigilance. Frontal functions are known to decline with aging and to be impaired in various neurodegenerative conditions, such as frontotemporal dementia [1], vascular dementia [2] and Alzheimer's disease (AD) [3–6]. Patterns of frontal dysfunction were reported to be useful in differentiating types of dementia [7, 8], and

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1420–8008/10/0294–0363$26.00/0

Accessible online at:
www.karger.com/dem

Ki Woong Kim, MD, PhD
Department of Neuropsychiatry, Seoul National University Bundang Hospital
300 Gumi-dong, Bundang-gu
Seongnam-si, Gyeonggi-do 463-707 (Korea)
Tel. +82 31 787 7439, Fax +82 31 787 4058, E-Mail kwkimmd@snu.ac.kr

to be associated with noncognitive symptoms in dementia. For example, deficits in executive function were associated with delusional thoughts [9], apathy/depression [10] and aberrant motor behavior [11] in AD patients. Therefore, identification of dysexecutive syndrome is important for the diagnosis and management of these conditions.

However, frontal lobe functions are difficult to assess clinically. Dubois et al. [12] developed the Frontal Assessment Battery (FAB) as a short bedside test to assess the presence and severity of dysexecutive syndrome affecting both cognitive and motor behavior. The FAB consists of six subtests: conceptualization (by means of a similarities task), mental flexibility (by a phonological fluency task), motor programming (Luria's motor series), sensitivity to interference (a conflicting instructions task), inhibitory control (a go-no-go task), and environmental autonomy (an evaluation of prehension behavior). Global performance on these six subtests gives a composite score summarizing the severity of the dysexecutive syndrome, while individual subscores might suggest a descriptive pattern of executive dysfunction in a given patient [13, 14]. A higher score means a better performance, and the maximum possible sum of the scores is 18 (scores on each item range from 0 to 3). The FAB's reliability and validity have been demonstrated by studies of various neurodegenerative diseases [15–20], and FAB score has been correlated with the blood flow in the frontal lobe [21]. The FAB provides as much information about the presence of executive dysfunction as the EXIT-25 can [22] but takes less time to complete and is less frustrating for patients. Thus, it is more feasible than the EXIT-25 as a screening test for executive dysfunction [23]. Since FAB scores are significantly influenced by age and education [13, 14], interpretation of the FAB's results requires normative data, particularly for populations with diverse demographic characteristics.

In the present study, we developed the Korean version of the FAB (FAB-K) and evaluated its reliability and validity in assessing dysexecutive syndrome in AD among Korean elders. We also constructed normative data for Korean elders, since the variation in educational levels is still very wide among Korean elders.

Methods

Translation

The original English version of FAB [12] was translated into Korean by two psychiatrists and one clinical psychologist who were familiar with both the English and Korean languages. The words used in the similarities task of the English version (tulip, rose, daisy) were changed into 'rose, lily, sunflower' since tulips and daisies are not familiar to Korean elders. The letter used in the fluency task ('S') was translated into the Korean alphabet 'ㄱ' (phonetic substitute in English: k). The other four items were translated directly into Korean.

To standardize the assessment process, we created an instruction manual based on the study by Dubois et al. [12] The psychologists administered all the assessments conformed to this manual. The raters were blinded to the clinical diagnoses of the study participants.

Participants

Normal elderly individuals were drawn from participants in the Korean Longitudinal Study on Health and Aging (KLoSHA) [24]. AD patients were drawn from visitors to the Dementia Clinics of the Seoul National University Bundang Hospital and the Kyunggi Provincial Hospital for the Elderly. All the participants were aged 55 years or older.

A research neuropsychiatrist using the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment packet (CERAD-K) [25] and the Korean version of Mini International Neuropsychiatric Interview (MINI-K) [26] evaluated each participant. After this clinical examination, a panel consisting of four research neuropsychiatrists reviewed all of the available raw data from the clinical evaluations and determined the diagnosis and the Clinical Dementia Rating (CDR). The panel diagnosed dementia and other major psychiatric disorders according to the criteria in the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) [27] and diagnosed AD according to the criteria of the NINCDS-ADRDA (National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association) [28]. The subjects whose CDR was 0 were enrolled in the normal control group. The study excluded participants with major psychiatric disorders and serious medical and neurological disorders that could affect their mental function but did not exclude individuals with minor physical abnormalities (e.g. diabetes with no serious complications, essential hypertension, mild hearing loss). All participants had adequate vision and hearing, although many wore glasses and some required hearing aids.

Data Analysis

In order to evaluate the inter-rater reliability, two raters (one tester and one rater) simultaneously assessed 23 participants (5 normal, 7 MCI and 11 AD). To evaluate test-retest reliability, the same rater administered the FAB-K to these participants again, 2 weeks after the initial assessment. To test inter-rater reliability and test-retest reliability, we calculated Pearson correlation coefficients. Internal consistency was examined using Cronbach's α and item-total correlations.

To evaluate discriminant validity, we compared the mean FAB-K scores among the five severity groups classified by the CDR (designated 0, 0.5, 1, 2 and 3) using ANOVA, after adjusting for age, gender and educational level. To evaluate concurrent validity, we administered the Categorical Verbal Fluency Test [25], Lexical Verbal Fluency Test [29], Design Fluency Test [30], Trail Making Test A and B [25], Wisconsin Card Sorting Test (WCST) [31] and Digit Span Test [32] together and calculated their Pearson correlation coefficients with the FAB-K. In addition, we evaluated

Table 1. Demographic and clinical characteristics of the subjects

	Normal	AD
Number	635	300
Age, years	73.58 ± 8.11	77.06 ± 8.2[c]
Female, %	49.4	74.0
Education, years	9.17 ± 5.47	6.35 ± 5.39[c]
MMSE	25.23 ± 3.35	13.88 ± 5.77[c]
Wisconsin Card Sorting Test, points		
Total correct	27.87 ± 10.44	20.95 ± 9.24[b]
Total errors	36.24 ± 10.50	40.18 ± 11.83
Perseverative responses	19.75 ± 12.88	28.55 ± 16.82[b]
Perseverative errors	16.75 ± 9.41	23.00 ± 12.54[b]
Nonperseverative errors	19.44 ± 11.15	17.18 ± 11.78
Conceptual level responses	17.13 ± 13.11	8.86 ± 9.90[b]
Categories completed	0.86 ± 1.01	0.14 ± 0.35[b]
Trials to complete 1st category	43.19 ± 22.75	63.14 ± 5.71[c]
Failure to maintain set	0.47 ± 0.86	0.36 ± 0.72
Design Fluency Test, points		
Novel output of free condition	10.35 ± 8.99	0.67 ± 1.63[a]
Novel output of fixed condition	7.60 ± 5.90	0.50 ± 0.83[b]
Verbal Fluency Test, points		
Categorical	14.46 ± 4.04	6.68 ± 3.97[c]
Lexical	23.24 ± 10.07	7.19 ± 8.19[c]
Digit Span Test, points		
Forward	6.43 ± 1.50	4.68 ± 1.67[c]
Backward	3.61 ± 1.31	2.12 ± 1.55[c]
Trail Making Test, s		
TMT-A	82.73 ± 66.44	207.03 ± 117.67[c]
TMT-B	205.53 ± 82.90	279.08 ± 53.66[c]

[a] $p < 0.01$, [b] $p < 0.005$, [c] $p < 0.001$ by ANOVA, adjusting age, education and gender, or χ^2 test.

Table 2. Item-total statistics of the FAB-K

	Corrected item-total correlation	Cronbach's α if item deleted
Item 1	0.586	0.761
Item 2	0.658	0.741
Item 3	0.643	0.746
Item 4	0.665	0.739
Item 5	0.572	0.772
Item 6	0.206	0.823

the correlation of the FAB-K with the CDR by means of the Spearman correlation coefficient and determined the optimal cut-off scores satisfying both sensitivity and specificity for AD via receiver operator characteristic (ROC) analyses. To measure the diagnostic accuracy of the FAB-K and MMSE for AD, we calculated the area under the ROC curves (AUC). To examine the difference between the FAB-K's and the MMSE's diagnostic accuracy for AD, we compared the AUC of the FAB-K with that of the MMSE by calculating a critical ratio z, as proposed by Hanley and McNeil in 1983 [33].

A multiple linear regression analysis with stepwise variable selection assessed the relative contributions of age, education, and gender to the FAB-K scores. To obtain normative data of the FAB-K, we also performed a series of 3 × 3 × 2 ANOVAs to determine any main effects and interactions of age (55–69, 70–79 and 80–96), educational level (0–3, 4–6 and ≥7 years), and gender on the performance of the FAB-K. The three age divisions corresponded to the age stratifications used in our previous studies on the normative data of other neuropsychological tests in Korean elders [25, 34, 35]. Educational level was divided into three groups corresponding to the categorical distinctions in the Korean public education system. Post hoc contrasts with Scheffé's method were conducted when any main effect of education was determined to be significant, via ANOVA, at the $p < 0.05$ level. We used the demographic factors that had main effects on the performance of the FAB-K to stratify the FAB-K's norms. To maximize the quantity of information and clinical usefulness of the data, we developed the normative data using overlapping strata, following the procedures described by Pauker [36].

All statistical analyses were performed using SPSS 15.0 for Windows.

Results

Participants

Nine hundred and thirty-five participants (300 AD patients, 635 normal controls) completed the present study. Among the AD patients, 106 had very mild dementia (CDR = 0.5), 114 had mild dementia (CDR = 1), 52 had moderate (CDR = 2), and 28 had severe (CDR ≥3). The control group was significantly younger ($F = 35.033$, d.f. = 1, $p < 0.001$) and better educated ($F = 54.612$, d.f. = 1, $p < 0.001$) than the AD group. The demographic and clinical characteristics of the participants are summarized in table 1.

Reliability

The FAB-K demonstrated good internal consistency. The Cronbach's α coefficient for the FAB-K was 0.802. Item-total statistics are summarized in table 2. The item-total correlations were significant ($p < 0.01$, Pearson correlation tests) for all 6 items, ranging from a minimum of 0.28 (prehension behavior) to a maximum of 0.80 (conflicting instruction).The FAB-K's inter-rater reliability was 0.980 ($p < 0.001$, Pearson correlation test) and the test-retest reliability was 0.820 ($p < 0.001$, Pearson correlation test).

Table 3. Correlation between the FAB-K and other neuropsychological tests for global cognition and frontal lobe function

Tests	Correlation coefficients						
	Total	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
MMSE	0.831[c]	0.606[c]	0.701[c]	0.677[c]	0.739[c]	0.510[c]	0.216[c]
Wisconsin Card Sorting Test							
Total correct	0.249[c]	0.217[c]	0.230[c]	0.098[a]	0.098[a]	0.190[c]	0.014
Total errors	–0.213[c]	–0.199[c]	–0.204[c]	–0.062	–0.048	–0.185[c]	–0.014
Perseverative responses	0.262[c]	–0.185[c]	–0.177[c]	–0.156[b]	–0.163[c]	–0.204[c]	–0.075
Perseverative errors	0.259[c]	–0.182[c]	–0.180[c]	–0.150[b]	–0.157[b]	–0.203[c]	–0.066
Nonperseverative errors	0.023	–0.030	–0.068	–0.069	–0.089	–0.004	–0.043
Conceptual level responses	0.240[c]	0.225[c]	0.224[c]	0.092[a]	0.091[a]	0.181[c]	0.021
Categories Completed	0.249[c]	0.240[c]	0.200[c]	0.109[a]	0.124[b]	0.177[c]	0.047
Trials to complete 1st category	0.236[c]	–0.209[c]	–0.209[c]	–0.063	–0.113[a]	–0.187[c]	–0.053
Failure to maintain set	0.061	0.030	0.091	–0.016	0.030	0.060	–0.020
Design Fluency Test							
Novel output of free condition	0.243[c]	0.306[c]	0.244[c]	0.076	0.066	0.108[a]	–0.035
Novel output of fixed condition	0.396[c]	0.399[c]	0.332[c]	0.167[b]	0.122[a]	0.252[c]	–0.005
Verbal Fluency Test							
Categorical	0.705[c]	0.524[c]	0.632[c]	0.572[c]	0.582[c]	0.427[c]	0.197[c]
Lexical	0.677[c]	0.458[c]	0.823[c]	0.386[c]	0.320[c]	0.362[c]	0.153[b]
Digit Span Test							
Forward	0.597[c]	0.456[c]	0.521[c]	0.480[c]	0.472[c]	0.375[c]	0.215[c]
Backward	0.685[c]	0.498[c]	0.593[c]	0.594[c]	0.549[c]	0.464[c]	0.181[c]
Trail Making Test							
TMT-A	–0.715[c]	–0.524[c]	–0.617[c]	–0.512[c]	–0.540[c]	–0.463[c]	–0.159[c]
TMT-B	–0.605[c]	–0.531[c]	–0.540[c]	–0.360[c]	–0.347[c]	–0.475[c]	–0.100[b]

[a] $p < 0.05$, [b] $p < 0.01$, [c] $p < 0.001$, Pearson correlation test.

Validity

As shown in table 3, the FAB-K's Pearson correlation coefficients with the MMSE and other frontal function tests were statistically significant, indicating that the FAB-K had a high concurrent validity. Table 4 shows the FAB-K's total scores and item scores across the CDR. As the CDR increased, total FAB-K scores and item scores decreased significantly (F = 192.026, d.f. = 4, $p < 0.001$, ANOVA). In the post hoc analysis, the total FAB-K score started to decline significantly at the very mild stage ($p < 0.05$, Bonferroni post hoc test). The subitems' scores began to decline at the very-mild to mild stages, except for that of item 6 (prehension behavior), which began to decrease at the severe stage. The scores of items 2, 3 and 4 were impaired from the very mild stage and became significantly worse as CDR increased. Item 1's score began to decline at the very mild stage and reached a floor at the moderate stage, whereas item 5's score began to decline and reached a floor at the mild stage. In addition, scores of subitem 1 (similarities), 2 (lexical fluency) and 5 (go-no-go) were more strongly correlated with the scores of WCST than those of subitem 3 (Luria's motor series), 5 (conflicting instructions) and 6 (prehension behavior; table 3).

The optimal cut-off score of the FAB-K for AD was determined as 10/11, where sensitivity and specificity for AD were 0.717 and 0.827, respectively. The FAB-K's AUC for AD was 0.854 (95%CI = 0.827–0.881), indicating that the FAB-K's diagnostic accuracy for AD may be good. However, the FAB-K's AUC was smaller than that of the MMSE (AUC = 0.954, 95%CI = 0.940–0.968, $p < 0.001$).

Normative Data

We used the normal control group to construct the FAB-K's normative data for Korean elders. Of these participants, 49% were women, and the males' mean age and educational level were significantly higher than those of the females (t = –2.851, $p < 0.01$ for age; t = –11.268, $p < 0.001$ for education). In the stepwise linear regression analysis, age (B = –0.081, SE = 0.012, standardized B = –0.210, $R^2 = 0.043$) and education (B = 0.315, SE = 0.018, standardized B = 0.551, $R^2 = 0.344$) had significant effects

Table 4. FAB-K scores stratified by CDR

	Normal (CDR = 0)	AD					ANOVA		
		Very mild (CDR = 0.5)	Mild (CDR = 1)	Moderate (CDR = 2)	Severe (CDR ≥3)	All	F [4, 927]	p	Post hoc
Number	635	106	114	52	28	300			
FAB-K									
Item 1	1.60 ± 0.90^a	1.07 ± 0.93^b	0.84 ± 0.85^b	0.40 ± 0.66^c	0.11 ± 0.31^c	0.78 ± 0.87	34.479	<0.001	a>b>c
Item 2	1.79 ± 0.98^a	1.03 ± 1.05^b	0.46 ± 0.80^c	0.17 ± 0.58^d	0.00 ± 0.00^d	0.57 ± 0.91	72.841	<0.001	a>b>c>d
Item 3	2.68 ± 0.70^a	2.11 ± 1.05^b	1.44 ± 1.06^c	0.85 ± 0.95^d	0.36 ± 0.48^d	1.47 ± 1.15	111.565	<0.001	a>b>c>d
Item 4	2.78 ± 0.65^a	1.99 ± 1.21^b	1.21 ± 1.33^c	0.63 ± 1.01^d	0.04 ± 1.8^d	1.28 ± 1.32	160.703	<0.001	a>b>c>d
Item 5	1.80 ± 1.38^a	1.42 ± 1.23^a	0.61 ± 0.98^b	0.29 ± 0.69^b	0.00 ± 0.00^b	0.79 ± 1.11	27.679	<0.001	a>b
Item 6	2.99 ± 0.17^a	2.92 ± 0.40^a	2.94 ± 0.40^a	2.85 ± 0.63^a	2.36 ± 1.25^b	2.86 ± 0.59	20.873	<0.001	a>b
Total	13.62 ± 3.13^a	10.54 ± 3.99^b	7.46 ± 3.59^c	5.19 ± 2.91^d	2.86 ± 1.45^d	7.73 ± 4.27	192.026	<0.001	a>b>c>d

ANOVA adjusted for age, education and gender with Scheffé's post hoc comparison.

on the FAB-K scores ($p < 0.001$), but gender failed to enter the model as a predicting factor.

We stratified the participants into three age groups (55–69, 70–79, and 80–96 years), 4 educational groups (0–3, 4–6, 7–12, and ≥13 years), and two gender groups (men, women). In a three-way ANOVA, the main effects of age ($F [2, 613] = 24.520$, $p < 0.001$) and educational level ($F [3, 613] = 36.914$, $p < 0.001$) were significant, but the gender effect was not ($F [1, 613] = 3.442$, $p > 0.05$). There were no interactions among the demographic variables. The mean FAB-K scores were higher in the younger and more educated participants. Since the main effects of age and education were significant, we performed post hoc contrasts, using Scheffé's method, among the same age and educational groups. These post hoc contrasts revealed that the younger groups showed significantly higher scores than the older groups and that more educated people performed better than the less educated, demonstrating significant differences between each sequential pair of age groups and sequential pair of educational levels. Based on the analysis for the effect of demographic variables, we decided to stratify the norm of the FAB-K by age and education. To maximize the quantity of information and clinical usefulness of the data, we divided the age groups into three overlapping strata (55–75, 65–85, and 75 years and over), with midpoint ages at 10-year intervals (65, 75, 85 years). As a result, the normative data from the age groups 55–75, 65–85, and 75 years and over applied to persons whose ages ranged from 55–69, 70–79, and 80–89 years, respectively. The FAB-K score for each stratum includes a mean, a standard deviation, a median, and a range from the 5th to the 95th percentile (table 5).

Table 5. Normative data of the FAB-K in Korean elders

Age, years	Education, years			
	0–3	4–6	7–12	≥13
55–69[a]				
Number	73	89	135	131
Mean	10.88	13.34	14.65	15.99
SD	2.82	2.58	2.24	1.80
Median	10.00	14.00	15.00	16.00
5th percentile	6.00	8.00	10.00	12.00
95th percentile	16.00	17.00	17.00	18.00
70–79[b]				
Number	88	109	162	140
Mean	10.56	13.15	14.51	15.72
SD	2.87	2.53	2.22	1.91
Median	10.00	13.00	15.00	16.00
5th percentile	6.00	8.00	10.00	12.00
95th percentile	15.55	17.00	17.00	18.00
≥80[c]				
Number	49	53	74	41
Mean	9.51	12.28	13.82	13.98
SD	3.30	2.87	2.56	2.60
Median	10.00	12.00	14.00	15.00
5th percentile	4.00	7.00	9.00	9.00
95th percentile	14.50	17.00	17.00	18.00

[a] Normative data from the people whose ages ranged from 55 to 75 years. [b] Normative data from the people whose ages ranged from 65 to 85 years. [c] Normative data from the people whose ages were 75 years and over.

Discussion

We found that the FAB-K was a valid and reliable instrument for evaluating frontal lobe function in the elderly. Its inter-rater and test-retest reliabilities and internal consistency were good, indicating that it is appropriate to use the FAB-K as a unified scale. The FAB-K's correlations with other frontal lobe function tests were all significant, indicating that the FAB-K has appropriate construct validity for testing frontal lobe function. These results correspond well to earlier observations on the FAB in other populations [12, 15, 16, 18–20, 22, 37]. However, FAB-K score may not reflect frontal lobe function as exclusively as the FAB was believed to do [17] since, in the present study, it was also correlated strongly with MMSE score, which is relatively insensitive to frontal dysfunction. Furthermore, the FAB-K's correlation coefficients with the WCST were much lower than its correlation coefficients with other frontal lobe function tests, such as verbal fluency tests, that may be less specific to frontal function than the WCST is.

We tested the criterion validity of the FAB-K using AD as the gold standard, since AD is the most common cause of dementia [38, 39]. Approximately 60% of AD patients manifest variable executive/frontal dysfunctions [4, 17], and thus AD may be the most common cause of dysexecutive syndrome in clinical settings. In the present study, FAB-K scores not only clearly discriminated the AD group from the control group but also discriminated AD severity as classified by the CDR, which was consistent with earlier observations [12]. These results indicated that the FAB-K might be useful in detecting early cognitive changes in AD. However, these cognitive changes may not be confined to frontal/executive dysfunction since the correlation of FAB-K with MMSE was stronger than that with WCST. The FAB-K's diagnostic accuracy for AD was fairly high, although it was generally lower than that of the MMSE. In addition, specific subsets of the FAB-K seemed to be more sensitive at discriminating cognitive impairments of AD from the normal elderly. Among the six FAB-K items, three (item 2, testing mental flexibility using lexical fluency; item 3, testing motor programming using Luria's motor series, and item 4, testing sensitivity to interference using conflicting instructions) began to decline at the very mild stage and kept decreasing significantly as CDR increased. In contrast, item 6 (testing environmental autonomy using prehension behavior) did not decline until the severe stage of AD. A previous study reported similar results [14]. Although prehension behavior is reportedly useful in discriminating other types of dementia, such as frontotemporal dementia [12, 16, 17, 20] and vascular dementia [18] from normal elders, it may not be suitable for the early detection of dysexecutive syndrome in AD due to this ceiling effect [14]. Therefore, Mok et al. [15] suggested that prehension behavior should be replaced by other brief executive tests. The remaining two items showed different patterns of decline from those of items 2, 3, 4, and 6. Item 1 (testing conceptualization using a similarity task) began to decline at the very mild stage of AD and reached a floor in moderate stage AD, whereas item 5 (testing inhibitory control using a go-no-go task) began to decline and reached a floor at the mild stage of AD. Therefore, differential patterns of decline among item scores may provide additional information for understanding AD patients by severity.

Consistent with previous studies [13, 14], FAB-K's performance in Korean elders was significantly influenced by age and education but not by gender. In particular, education's influence on FAB-K's performance was strong, explaining 34.4% of the total variance in the FAB-K score. Therefore, normative data are essential for using FAB-K in clinical practice, particularly in the least-developed and developing countries, where the range of educational levels in the populations is very wide. In establishing the normative data for FAB-K, we adopted the overlapping age stratification method to resolve the limited sample size of each cell [36]. This procedure allowed us to present more stratified and more accurate normative data tables by providing adequate numbers of participants for most normative cells and a broad normative base for each normative cell. Although we estimated the normative data in each age-overlapping table from a broader age range that overlaps with an adjacent one, these data are applicable to people within a narrow, non-overlapping age range. In addition, this procedure can provide more stable means of stratification, resulting in less abrupt mean shifts between age strata than simple non-overlapping age stratification. The users of these normative data can refer to the title of each normative table, which shows the applicable age range [34].

Our study has two strengths with regard to providing normative information for the FAB-K. First, by means of structured evaluations, we strictly excluded both cognitive disorders, including very mild dementia (CDR = 0.5), and major psychiatric disorders, which are prevalent in elderly populations. Therefore, this normative data is not likely to be confounded by the inclusion of patients misclassified as normal elderly. Second, the participants' range of educational levels was wide enough to evenly cover from the non-educated level to the post-graduate level. Therefore, this normative data can be used both in developing countries, where a substantial proportion of the elderly population is still undereducated, and in developed countries, where the majority of the elderly population is well-educated.

Several limitations warrant consideration in generalizing our observations. First, we did not examine other types of dementia, such as frontotemporal dementia and vascular dementia, that are known to have dysexecutive syndrome more frequently than AD does. This may have contributed to the low correlation between FAB-K scores and WCST scores in the present study, since the types of frontal dysfunction in these types of dementia may differ from those in AD. Second, we did not include in our control group those cognitively impaired but not demented elders whose CDR was 0.5, which may have raised the AUC for both the FAB-K and the MMSE with regard to AD. Third, the sample size of some cells in the normative data was relatively small. Since small sample size increases standard errors and possibly reduces the stability of the estimated results, users should be cautious when interpreting the test scores of individuals within the cells having small sample sizes.

In conclusion, the FAB-K was a reliable and valid bedside instrument for assessing executive dysfunction in Korean elders. The age- and education-specific normative data for the FAB-K may assist clinicians and researchers in their interpretation of FAB-K's performance in elderly populations with wide variations in educational levels.

Acknowledgments

This study was supported by an Independent Research Grant from Pfizer Global Pharmaceuticals (grant No. 06-05-039) and a grant of the Korean Health Technology R&D Project, Ministry for Health, Welfare & Family Affairs, Republic of Korea (grant No. A092077).

References

1 Clinical and neuropathological criteria for frontotemporal dementia: the Lund and Manchester groups. J Neurol Neurosurg Psychiatry 1994;57:416–418.

2 Desmond DW, Erkinjuntti T, Sano M, Cummings JL, Bowler JV, Pasquier F, Moroney JT, Ferris SH, Stern Y, Sachdev PS, Hachinski VC: The cognitive syndrome of vascular dementia: implications for clinical trials. Alzheimer Dis Assoc Disord 1999;13(suppl 3):S21–S29.

3 Greene JD, Hodges JR, Baddeley AD: Autobiographical memory and executive function in early dementia of Alzheimer type. Neuropsychologia 1995;33:1647–1670.

4 Swanberg MM, Tractenberg RE, Mohs R, Thal LJ, Cummings JL: Executive dysfunction in Alzheimer disease. Arch Neurol 2004;61:556–560.

5 Sgaramella TM, Borgo F, Mondini S, Pasini M, Toso V, Semenza C: Executive deficits appearing in the initial stage of Alzheimer's disease. Brain Cogn 2001;46:264–268.

6 Perry RJ, Hodges JR: Attention and executive deficits in Alzheimer's disease: a critical review. Brain 1999;122:383–404.

7 Fukui T, Lee E, Kitamura M, Hosoda H, Bokui C, Ikusu K, Okita K: Visuospatial dysfunction may be a key in the differentiation between Alzheimer's disease and subcortical cognitive impairment in moderate to severe stages. Dement Geriatr Cogn Disord 2009; 28:288–294.

8 Hanyu H, Sato T, Kume K, Takada Y, Onuma T, Iwamoto T: Differentiation of dementia with lewy bodies from Alzheimer disease using the frontal assessment battery test. Int J Geriatr Psychiatry 2009;24:1034–1035.

9 Nagata T, Ishii K, Ito T, Aoki K, Ehara Y, Kada H, Furukawa H, Tsumura M, Shinagawa S, Kasahara H, Nakayama K: Correlation between a reduction in frontal assessment battery scores and delusional thoughts in patients with Alzheimer's disease. Psychiatry Clin Neurosci 2009;63:449–454.

10 Nakaaki S, Murata Y, Sato J, Shinagawa Y, Hongo J, Tatsumi H, Hirono N, Mimura M, Furukawa TA: Association between apathy/depression and executive function in patients with Alzheimer's disease. Int Psychogeriatr 2008;20:964–975.

11 Nagata T, Shinagawa S, Ochiai Y, Kada H, Kasahara H, Nukariya K, Nakayama K: Relationship of frontal lobe dysfunction and aberrant motor behaviors in patients with Alzheimer's disease. Int Psychogeriatr 2009: 1–7.

12 Dubois B, Slachevsky A, Litvan I, Pillon B: The FAB: a frontal assessment battery at bedside. Neurology 2000;55:1621–1626.

13 Iavarone A, Ronga B, Pellegrino L, Lore E, Vitaliano S, Galeone F, Carlomagno S: The frontal assessment battery (FAB): normative data from an Italian sample and performances of patients with Alzheimer's disease and frontotemporal dementia. Funct Neurol 2004;19:191–195.

14 Appollonio I, Leone M, Isella V, Piamarta F, Consoli T, Villa ML, Forapani E, Russo A, Nichelli P: The frontal assessment battery (FAB): normative values in an Italian population sample. Neurol Sci 2005;26:108–116.

15 Mok VC, Wong A, Yim P, Fu M, Lam WW, Hui AC, Yau C, Wong KS: The validity and reliability of Chinese frontal assessment battery in evaluating executive dysfunction among Chinese patients with small subcortical infarct. Alzheimer Dis Assoc Disord 2004;18:68–74.

16 Slachevsky A, Villalpando JM, Sarazin M, Hahn-Barma V, Pillon B, Dubois B: Frontal assessment battery and differential diagnosis of frontotemporal dementia and Alzheimer disease. Arch Neurol 2004;61:1104–1107.

17 Lipton AM, Ohman KA, Womack KB, Hynan LS, Ninman ET, Lacritz LH: Subscores of the FAB differentiate frontotemporal lobar degeneration from AD. Neurology 2005;65:726–731.

18 Oguro H, Yamaguchi S, Abe S, Ishida Y, Bokura H, Kobayashi S: Differentiating Alzheimer's disease from subcortical vascular dementia with the FAB test. J Neurol 2006;253:1490–1494.

19 Kugo A, Terada S, Ata T, Ido Y, Kado Y, Ishihara T, Hikiji M, Fujisawa Y, Sasaki K, Kuroda S: Japanese version of the frontal assessment battery for dementia. Psychiatry Res 2007;153:69–75.

20 Nakaaki S, Murata Y, Sato J, Shinagawa Y, Matsui T, Tatsumi H, Furukawa TA: Reliability and validity of the Japanese version of the frontal assessment battery in patients with the frontal variant of frontotemporal dementia. Psychiatry Clin Neurosci 2007;61:78–83.
21 Yoshida H, Terada S, Sato S, Kishimoto Y, Ata T, Ohshima E, Honda H, Ishihara T, Kuroda S: Frontal assessment battery and brain perfusion imaging in early dementia. Dement Geriatr Cogn Disord 2009;27:133–138.
22 Royall DR, Mahurin RK, Gray KF: Bedside assessment of executive cognitive impairment: the executive interview. J Am Geriatr Soc 1992;40:1221–1226.
23 Moorhouse P, Gorman M, Rockwood K: Comparison of exit-25 and the frontal assessment battery for evaluation of executive dysfunction in patients attending a memory clinic. Dement Geriatr Cogn Disord 2009;27:424–428.
24 Park JH, Lim S, Lim JY, Kim KL, Han MK, Yoon IY, Kim JM, Chang YS, Chang CB, Chin HJ, Choi EA, Lee SB, Park YJ, Paik NJ, Kim KJ, Jang HC, Kim KW: An overview of the Korean Longitudinal Study on Health and Aging. Psychiatry Invest 2007;4:84–95.
25 Lee JH, Lee KU, Lee DY, Kim KW, Jhoo JH, Kim JH, Lee KH, Kim SY, Han SH, Woo JI: Development of the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease assessment packet (CERAD-K): clinical and neuropsychological assessment batteries. J Gerontol B Psychol Sci Soc Sci 2002;57:P47–P53.
26 Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC: The mini-international neuropsychiatric interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59(suppl 20):22–33; quiz 34–57.
27 American Psychiatric Association: Diagnostic and statistical manual of mental disorders, ed 4. Washington, DC, American Psychiatiric Association, 1994.
28 McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM: Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on Alzheimer's disease. Neurology 1984;34:939–944.
29 Benton AL, Hamsher, K de S.: Multilingual Aphasia Examination. Iowa City, University of Iowa, 1976.
30 Jones-Gotman M, Milner B: Design fluency: the invention of nonsense drawings after focal cortical lesions. Neuropsychologia 1977;15:653–674.
31 Robinson AL, Heaton RK, Lehman RA, Stilson DW: The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. J Consult Clin Psychol 1980;48:605–614.
32 Newton RL: A comparison of two methods of administering the Digit Span Test. J Clin Psychol 1950;6:409–412.
33 Hanley JA, McNeil BJ: A method of comparing the areas under Receiver Operating Characteristic curves derived from the same cases. Radiology 1983;148:839–843.
34 Lee DY, Lee KU, Lee JH, Kim KW, Jhoo JH, Kim SY, Yoon JC, Woo SI, Ha J, Woo JI: A normative study of the CERAD neuropsychological assessment battery in the Korean elderly. J Int Neuropsychol Soc 2004;10:72–81.
35 Seo EH, Lee DY, Kim KW, Lee JH, Jhoo JH, Youn JC, Choo IH, Ha J, Woo JI: A normative study of the trail making test in Korean elders. Int J Geriatr Psychiatry 2006;21:844–852.
36 Pauker JD: Constructing overlapping cell tables to maximize the clinical usefulness of normative test data: rationale and an example from neuropsychology. J Clin Psychol 1988;44:930–933.
37 Nelson HE: A modified card sorting test sensitive to frontal lobe defects. Cortex 1976;12:313–324.
38 Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, Hall K, Hasegawa K, Hendrie H, Huang Y, Jorm A, Mathers C, Menezes PR, Rimmer E, Scazufca M: Global prevalence of dementia: a Delphi consensus study. Lancet 2005;366:2112–2117.
39 Jhoo JH, Kim KW, Huh Y, Lee SB, Park JH, Lee JJ, Choi EA, Han C, Choo IH, Youn JC, Lee DY, Woo JI: Prevalence of dementia and its subtypes in an elderly urban Korean population: results from the Korean Longitudinal Study on Health and Aging (KLoSHA). Dement Geriatr Cogn Disord 2008;26:270–276.