

검사-재검사 간격이 K-MMSE 점수 변화의 유의성 지표에 미치는 영향

송민지 · 김지향* · 유경† · 김지현†
이주일‡ · 강연욱*‡

한림대학교춘천성심병원 신경과
한림대학교성심병원 신경과*
한림대학교 고령사회연구소†
한림대학교 심리학과‡

Received: November 14, 2012
Revision received: December 18, 2012
Accepted: December 24, 2012

Address for correspondence

Yeonwook Kang, Ph.D.
Department of Psychology, Hallym University,
39 Hallymdaehak-gil, Chuncheon 200-702,
Korea
Tel: +82-33-248-1724
Fax: +82-33-256-3424
E-mail: ykang@hallym.ac.kr

*본 연구는 한국학술진흥재단의 지원(KRF-2007-411-J01902)에 의하여 이루어진 것임.

The Influence of Test-Retest Interval on the Significant Change Indices for the K-MMSE

Minji Song, M.A., Jihyang Kim, M.A.* , Kyung Ryu, Ph.D.† , Jihyun Kim, Ph.D.†,
Juil Rie, Ph.D.‡, Yeonwook Kang, Ph.D.*‡

Department of Neurology, Hallym University Chuncheon Sacred Heart Hospital, Chuncheon; Department of Neurology*, Hallym University Sacred Heart Hospital, Anyang; Hallym University Institute of Aging†, Chuncheon; Department of Psychology‡, Hallym University, Chuncheon, Korea

Background: The Mini-Mental State Examination (MMSE) has been commonly used to measure cognitive change over time. The aim of present study was to investigate the normative rates of change for the MMSE across test-retest intervals. **Methods:** We administered the Korean MMSE (K-MMSE) to 1055 community-dwelling middle aged and older adults three times over 6 years. Based on the Christensen's health screening criteria (1991), 234 middle-aged healthy adults (104 men, 130 women; mean age 55.95 ± 6.20 years; age range 45-64 years; mean education 7.63 ± 4.06 years) and 505 healthy elderly (200 men, 305 women; mean age 71.00 ± 4.62 years; age range 65-79 years; mean education 5.61 ± 5.12 years) were selected for the statistical analysis. Reliable change indices were computed using two different statistical methods, the Reliable Change Index adjusted for practice effects (RC_{PE}; Chelune et al., 1993) and the Standardized Regression-Based Change Index (SRBCI; McSweeney et al., 1993). **Results:** For the middle-aged healthy adult group, the 90% confidence intervals of the RC_{PE} and SRBCI were the same such as ± 4 in 2-year, ± 5 in 4-year, and ± 6 in 6-year test-retest intervals. For the healthy elderly group, the 90% confidence intervals of the RC_{PE} were -5 and +4 in 2-year interval and -7 and +5 in 4- & 6-year intervals. The 90% confidence intervals of the SRB change index were ± 4 in 2-year interval and ± 6 in 4- & 6-year intervals. **Conclusions:** The result provides the normative data of the reliable change scores for the K-MMSE for the middle-aged and older adults. It shows that the reliable change indices were varied across different age groups as well as test-retest intervals.

Key Words: MMSE, Test-retest interval, Reliable change index, Standardized regression based change index

서 론

노인 인구의 증가에 따른 노인병 환자의 증가는 최근 사회적 문제로 대두되고 있다. 노인병 중에서도 가장 심각한 질병 중 하나가 치매이다. 그러나 치매는 초기에 발견하여 적절한 치료적 개입을 하면 증상을 개선시키거나 치매의 진행을 늦추는 것이 부분적으로 가능하고, 이를 통해서 환자 및 보호자의 고통과 부담을 경감시키고 치매로 인한 사회적 비용을 절감할 수 있다. 따라서, 많은 나라에서 치매를 초기에 발견하기 위한 국가적 차원의 사업을 시행하고 있는데 국내에서도 2006년에 치매 조기검진사업이 시작되어 보건복지부 주관 하에 전국 각지의 보건소 및 의료기관에서 치매진단검사를 실시하고 있다.

치매의 조기검진사업에는 인지선별검사(cognitive screening test)

로서 전세계에서 널리 사용되고 있는 Mini-Mental State Examination (MMSE)이 사용되고 있다[1]. MMSE는 다양한 인지기능들을 짧은 시간에 측정할 수 있고 치매여부를 탐지하는 검사로서의 신뢰도와 타당도가 입증된 검사이다[2]. 국내에서도 한국판 MMSE (Korean version of MMSE; K-MMSE)가 개발되어 표준화되었고 치매선별도구로서 의료기관에서 널리 사용되어 왔다[3].

지역사회를 중심으로 치매 예방 및 조기 발견 사업이 활발히 진행됨에 따라서 시간 간격을 두고 MMSE를 반복 실시한 후 검사-재검사시 MMSE 점수의 변화를 치매의 조기 탐지와 치매의 진전이나 치료 효과를 평가하는 지표로 사용하는 경우가 증가하고 있다[4]. 그러나, MMSE 점수의 변화를 이런 목적으로 사용하기 위해서는 먼저 고려하여야 할 중요한 문제가 있다. 그것은 능력 검사가 반복적으로 시행되었을 때 나중에 얻은 점수를 첫 시행에서 얻은 점수

와 동일하게 해석하여도 되는지에 관한 문제이다[5]. 같은 검사가 반복적으로 실시되었을 때 그 결과에 영향을 미칠 수 있는 변인들은 “검사 자체와 관련된 변인”과 “검사 상황과 관련된 변인”으로 나누어 생각해 볼 수 있다. 전자에 해당되는 변인들로는 검사 자체의 신뢰도, 측정 오차, 연습 효과, 천정 효과, 바닥 효과 등이 있고 후자에 해당되는 변인들로는 1차 검사와 재검사의 간격, 평균으로의 회귀 현상 등이 있다[6]. 또한 나이와 교육수준 등 환자 개인의 인구학적 변인, 환자의 임상적인 상태, 과거의 검사 경험 등도 재검사 시의 점수 변화에 영향을 미칠 수 있다고 알려져 있다[7,8].

따라서 이러한 변인들로부터의 영향을 최소화하기 위해서 여러 가지 방법론적인 보완과 통계적인 기법이 제안되어 왔다. 방법론적인 측면에서 가장 자주 언급되는 방법은 동형검사를 사용하는 것이다. 그러나, MMSE처럼 동형검사가 개발되지 않았거나 개발된 동형검사의 타당도가 충분히 입증되지 않은 경우가 많고, 동형검사를 사용하여도 연습효과를 제거할 수 없다는 선행연구가 보고된 바 있다[9]. 또 다른 방법은 반복적인 검사 시 “정상적인 변화”를 나타내는 통제 집단의 자료와 연구 대상 집단(실험 집단)의 자료를 비교하는 것인데, 정상적인 변화가 어느 정도의 변화인지 그 기준이 명확하게 제시되어 있지 않은 검사들이 많고 실험 집단과 통제 집단의 개인차를 통제하는 것이 어렵기 때문에 이 방법 또한 한계가 있다[10]. MMSE의 경우, 정상 집단의 점수 변화에 대한 연구가 외국에서는 소수 수행되었으나[4, 11, 12], 국내에서는 아직까지 이에 대한 자료를 제공할 만한 연구가 수행되지 않았다. 마지막으로, 또 다른 방법은 검사-재검사 간격을 통제하는 방법이다. 그러나 검사-재검사 간격이 짧을수록 연습 효과가 나타날 가능성이 높아지고 그 간격이 길수록 연습 효과로부터 자유로울 것이라는 사실은 많은 연구자들이 인정하고 있지만, 연습효과를 배제할 수 있는 최적의 검사-재검사 간격이 어느 정도인지에 대해서는 합의된 연구 결과가 아직 보고되지 않고 있다.

이 같은 방법론적인 보완에 의해서도 해결되지 못한 문제들을 보완하기 위해서 다양한 통계적인 방법들이 개발되었다[6]. 검사-재검사 점수를 통계적으로 비교하기 위해서는 환자 집단의 1차 검사와 2차 검사의 점수 차이를 정상 집단에서의 점수 차이와 비교하는 방법과 표준편차를 이용하여 정상분포상에서 점수의 차이가 유의한지 알아보는 방법이 전통적으로 사용되어 왔으나 이 방법들은 반복적인 평가에 영향을 미치는 앞에서 기술된 다양한 변인들을 통제하지 못한다는 비판을 받았다. 특히, 검사-재검사 두 시점에서 집단의 평균 점수의 변화를 가지고 MMSE 점수의 안정성을 판단하는 방법은 각 개인들의 점수 변화를 잘 반영하지 못한다는 강한 비판을 받았다[11]. 이 같은 전통적인 통계치들의 단점들을 보완하여 검사-재검사 점수 변화의 유의성을 평가하기 위해서 새로 개발된 통계적인 방법들 중에서는 “변화유의지표(Reliable Change Index

Table 1. Reliable change scores and their formulas

Reliable change score	Formula
RCI _{PE}	$T_2 - (T_1 + D) / SED$
SRBCI	$T'_2 - T_2 / SEE$

RCI_{PE}, Reliable Change Index adjusted for Practice Effect; SRBCI, Standardized Regression-Based Change Index; T₁, score at Time 1; T₂, score at Time 2; D, mean of difference of score at Time 2 and score at Time 1; SED, Standard Error of the Difference; T'₂, predicted score at Time 2 based on the regression model; SEE, Standard Error of the Estimate of the regression model.

with adjustment for Practice Effects, RCI_{PE}) [14]와 “표준화 회귀에 기반한 변화 지표(Standard Regression-Based Change Index, SRBCI) [15]”가 경험적인 연구들을 통해서 신뢰할 만한 방법으로 추천되었다 [13] (Table 1). RCI_{PE}는 검사-재검사 결과를 바탕으로 실제로는 인지 기능의 변화가 없음에도 나타날 수 있는 검사-재검사 점수 변화의 분포를 예측하고, 이 변화가 일어날 수 있는 확률을 계산하여 변화의 유의성을 판단할 수 있도록 한다. 연습 효과로 수행이 향상된 부분을 보정할 수 있고, 실제 관찰된 재검사 점수와 예측된 값의 차이가 유의미한 수준인지를 판단한다. 반면, SRBCI는 인구 통계학적 변수들을 통제하고 다중회귀 방정식을 구하여 개인의 재검사 점수를 예측한다. 따라서 SRBCI를 사용하면 점수 변화에 대한 인구 통계학적 변수의 조절효과와 평균으로의 회귀 현상을 동시에 고려할 수 있게 된다.

본 연구는 재검사점수에 영향을 미치는 다양한 변인들을 통제할 수 있는 이 같은 통계 방법을 사용하여, MMSE를 반복적으로 실시하였을 때 어느 정도의 점수 변화를 정상적인 변산을 벗어난 유의미한 인지기능의 변화로 간주할 것인지 그에 대한 국내 기준을 발견하기 위하여 수행되었다. 정상적인 노화과정에서 인지기능 감퇴와 치매와 같은 병리적인 인지기능 저하를 변별하기 위해서는 일정 기간 동안 정상인들에게서 정상적으로 발생하는 인지기능 변화의 크기를 먼저 알아야 한다. 따라서 본 연구는 지역사회에서 살고 있는 건강한 장노년들에게 K-MMSE를 반복 실시한 후 기저검사와 비교하여 K-MMSE 재검사 시의 점수 변화가 유의미한 변화인지를 판단할 수 있는 지표가 되는 예측구간인 “변화유의구간”을 RCI_{PE}와 SRBCI에 대해서 산출할 것이다. 또한 이 두 지표가 어떤 차이가 있는지를 밝히고, 더 나아가서 이 예측구간이 기저검사 후 재검사까지의 시간 간격에 따라서 어떻게 달라지는지를 확인하고자 하였다.

연구 대상 및 방법

대상 및 도구

Hallym Aging Study [16]의 일환으로 다단계 층화표집을 통해서

지역사회에서 선발된 45세 이상 80세 미만의 장노년들에게 2003년 1-2월에 K-MMSE [3]를 실시하였고, 2년 후인 2005년 1-2월에 2차 검사를 실시하였으며, 그로부터 4년 후인 2009년 1-2월에 3차 검사를 실시하였다. 3회의 검사에 모두 참여한 피검자는 총 1,055명이었다. 이들 중 매 검사 때마다 시행된 Christensen 등[17]의 건강선별설문에서 3회 모두 건강기준을 충족한 피검자들이 739명(70.0%)이었고, 이들을 신경학적이거나 정신과적인 문제를 지니지 않은 인지적/정서적으로 건강한 사람으로 판정하였다. 정상적인 노화과정에서 나타나는 재검사 시 점수변화의 유의성 지표를 산출하고자 하는 것이 본 연구의 목적이므로 결과 분석에는 전체 피검자 1,055명 중 건강한 피검자 739명의 K-MMSE 자료만을 포함시켰다. 이들은 평균 연령이 66.24세(SD 8.71세, Range 45-79세), 평균 교육년수가 6.25년(SD 4.90년, Range 0-21년)이었고, 남자가 304명, 여자가 435명이었다. 피검자의 연령이 40대에서 70대까지 중장년층에서부터 노년층까지를 광범위하게 포함하고 있으므로 전체 피검자를 65세를 기준으로 하여 45-64세에 속하는 중장년집단과 65-79세에 속하는 노인집단으로 나누어서 분석하였다. 중장년집단에 속하는 피검자는 234명(남자 104명, 여자 130명)이었고, 평균 연령이 55.95세(SD 6.20세, Range 45-64세), 평균 교육년수가 7.63년(SD 4.06년, Range 0-18년)이었으며, 노인집단에 속하는 피검자는 505명(남자 200명, 여자 305명)이었고, 평균 연령이 71.00세(SD 4.62년, Range 65-79세), 평균 교육년수가 5.61년(SD 5.12년, Range 0-21년)이었다.

자료 분석

2003년도에 실시한 기저검사(1차 검사)와 2005년에 실시한 2차 검사, 그리고 2009년에 실시한 3차 검사 점수 간의 차이를 분석하기 위

하여 전통적인 방법인 반복측정 일원변량분석(repeated measures analysis of variance)을 실시하였다. 다음으로 개인별 점수변화를 살펴보기 위해서 검사-재검사 점수가 차이가 없거나 1점에서부터 5점 이상 차이가 나는 피검자들의 수(백분율)를 각 차이 점수별로 산출하였다. 마지막으로 RCI_{PE}와 SRBCI, 두 지표 각각의 90% 예측구간에 해당하는 점수범위를 산출하고 그 결과를 비교하였다. RCI_{PE}는 표준점수(z score)와 같이 해석될 수 있으며, RCI_{PE}의 90% 예측구간은 $[1.645 \times (\text{검사-재검사 차이값의 표준편차})]$ 로 계산하였다. 피검자의 재검사 점수가 예측구간에 포함되지 않는다면, 개인의 점수가 재검사에서 유의하게 변화하였다고 해석할 수 있다. SRBCI도 RCI_{PE}와 마찬가지로 표준점수처럼 해석할 수 있으나, SRBCI에서는 회귀방정식에 의해서 예측된 재검사 점수와 표준오차 값을 사용하여 계산한다. 이 값을 얻기 위해서 단계적 회귀분석(stepwise regression analysis)을 실시하였고, 이때 재검사 결과에 강한 영향을 미치는 것으로 알려진 기저검사 점수를 기본적인 예측변수로 포함시켰으며, 인구통계학적 변인인 성별, 연령, 교육년수를 후보 예측 변수로 포함시켰다. SRBCI의 90% 예측구간은 $[1.645 \times (\text{회귀 모델에 의해서 추정된 표준 오차})]$ 로 계산하였다.

결 과

K-MMSE를 2년과 4년 간격을 두고 총 3회에 걸쳐서 시행한 결과, 전체 집단, 중장년집단 및 노인집단 모두에서 K-MMSE의 평균점수는 6년에 걸쳐서 점차 저하된 양상을 나타내었고 그 점수차이는 통계적으로 유의한 수준에 속하였으나(중장년집단: $F(2, 466) = 3.69, p < .05$; 노인집단: $F(2, 1008) = 31.52, p < .001$) 중장년집단보다는 노인

Table 2. Difference of scores between the baseline test and retests

Participants	Test	Mean (SD)	Range	F		
< 65 years old (n = 234)	Baseline test (2003)	27.30 (2.44)	16-30	3.69*		
	1st retest (2005)	26.93 (2.72)	16-30			
	2nd retest (2009)	26.89 (3.16)	11-30			
≥ 65 years old (n = 505)	Baseline test (2003)	24.88 (4.38)	12-30	31.52*		
	1st retest (2005)	24.68 (4.27)	12-30			
	2nd retest (2009)	23.82 (5.13)	2-30			
Participants	Contrast Comparison	Mean Difference Score ^a (SD)	Range	Percent of Baseline Score ^b (SD)	Range	F
< 65 years old (n = 234)	Baseline test vs. 1st retest	-0.38 (2.10)	-13-4	98.82 (7.68)	55.17-116.00	7.54 [†]
	1st retest vs. 2nd retest	-0.04 (2.89)	-15-11	100.32 (12.01)	42.31-168.75	0.04
	Baseline test vs. 2nd retest	-0.41 (2.69)	-15-5	98.71 (10.32)	42.31-125.00	5.56*
≥ 65 years old (n = 505)	Baseline test vs. 1st retest	-0.19 (2.50)	-8-9	99.92 (11.22)	66.67-164.29	3.05
	1st retest vs. 2nd retest	-0.87 (3.47)	-19-11	96.77 (16.02)	11.11-191.67	31.35*
	Baseline test vs. 2nd retest	-1.06 (3.51)	-19-9	96.13 (15.82)	10.53-150.00	46.12*

^aMean difference score was calculated by subtracting the baseline (earlier) score from the retest (later) score; ^bPercentages of baseline scores was calculated by dividing the retest (later) score by the baseline (earlier) score.

* $p < 0.05$, [†] $p < 0.01$, [‡] $p < 0.001$.

집단에서 더 현저한 점수 저하가 관찰되었다(Table 2). 검사-재검사 평균점수의 차이(mean difference score)와 재검사 점수를 기저검사 점수와 비교한 백분율(percent of baseline score)에 대해서 대비 검정(contrast comparison)을 실시한 결과, 중장년집단에서는 2차 검사의 점수는 기저검사(1차 검사)의 점수보다($F(1, 233) = 7.54, p < 0.01$), 3차 검사의 점수는 기저검사(1차 검사)의 점수보다($F(1, 233) = 5.56, p < 0.05$) 유의하게 저하되었음이 확인되었으나, 3차 검사의 점수와 2차 검사의 점수 차이는 유의하지 않았다($F(1, 233) = .04, p = 0.84$). 반면 노인 집단에서는 3차 검사의 점수가 기저검사(1차 검사)와($F(1, 233) = 46.12, p < 0.001$), 2차 검사의 점수보다($F(1, 233) = 31.35, p < 0.001$) 유의하게

저하되었음이 확인되었으나, 기저검사(1차 검사)의 점수와 2차 검사의 점수 차이는 그 경향성만 관찰되었을 뿐 유의한 수준에 이르지 못하였다($F(1, 233) = 3.05, p = 0.08$). 그러나 절대값으로 볼 때 평균점수 차이는 중장년집단에서는 각각 0.38, 0.04, 0.41, 노인집단에서는 각각 0.19, 0.87, 1.06으로 매우 작은 차이였고, 백분율로 볼 때에도 중장년집단에서는 각각 1.18%, 0.32%, 1.29%. 노인집단에서는 각각 0.08%, 3.23%, 3.87%의 매우 작은 차이를 드러내었다.

피검자들 개개인이 실제로 재검사 시 어떠한 점수의 변화를 나타냈는지 구체적으로 살펴보기 위해서 변화된 각 점수대별로 피검자들의 빈도를 분석하였다(Table 3, Fig. 1). 그 결과, 재검사 시에 점수의

Table 3. Number (%) of participants on each difference of score

Participants	Test comparison	Difference of score										
		≥-5	-4	-3	-2	-1	0	+1	+2	+3	+4	≥+5
< 65 years old (n = 234)	Baseline test vs. 1st retest	8 (3.4)	4 (1.7)	17 (7.3)	35 (15.0)	37 (15.8)	54 (23.1)	41 (17.5)	23 (9.8)	11 (4.7)	4 (1.7)	0 (0.0)
	1st retest vs. 2nd retest	13 (5.6)	7 (3.0)	16 (6.8)	21 (9.0)	35 (15.0)	45 (19.2)	42 (17.9)	18 (7.7)	17 (7.3)	10 (4.3)	10 (4.3)
	Baseline test vs. 2nd retest	16 (6.8)	5 (2.1)	14 (6.0)	22 (9.4)	35 (15.0)	62 (26.5)	34 (14.5)	20 (8.5)	20 (8.5)	4 (1.7)	2 (0.9)
≥ 65 years old (n = 505)	Baseline test vs. 1st retest	25 (5.0)	16 (3.2)	50 (9.9)	47 (9.3)	76 (15.0)	94 (18.6)	76 (15.0)	58 (11.5)	31 (6.1)	19 (3.8)	13 (2.6)
	1st retest vs. 2nd retest	62 (12.3)	27 (5.3)	44 (8.7)	57 (11.3)	66 (13.1)	76 (15.0)	60 (11.9)	40 (7.9)	38 (7.5)	17 (3.4)	18 (3.6)
	Baseline test vs. 2nd retest	71 (14.1)	25 (5.0)	39 (7.7)	69 (13.7)	63 (12.5)	72 (14.3)	66 (13.1)	49 (9.7)	16 (3.2)	14 (2.8)	21 (4.2)

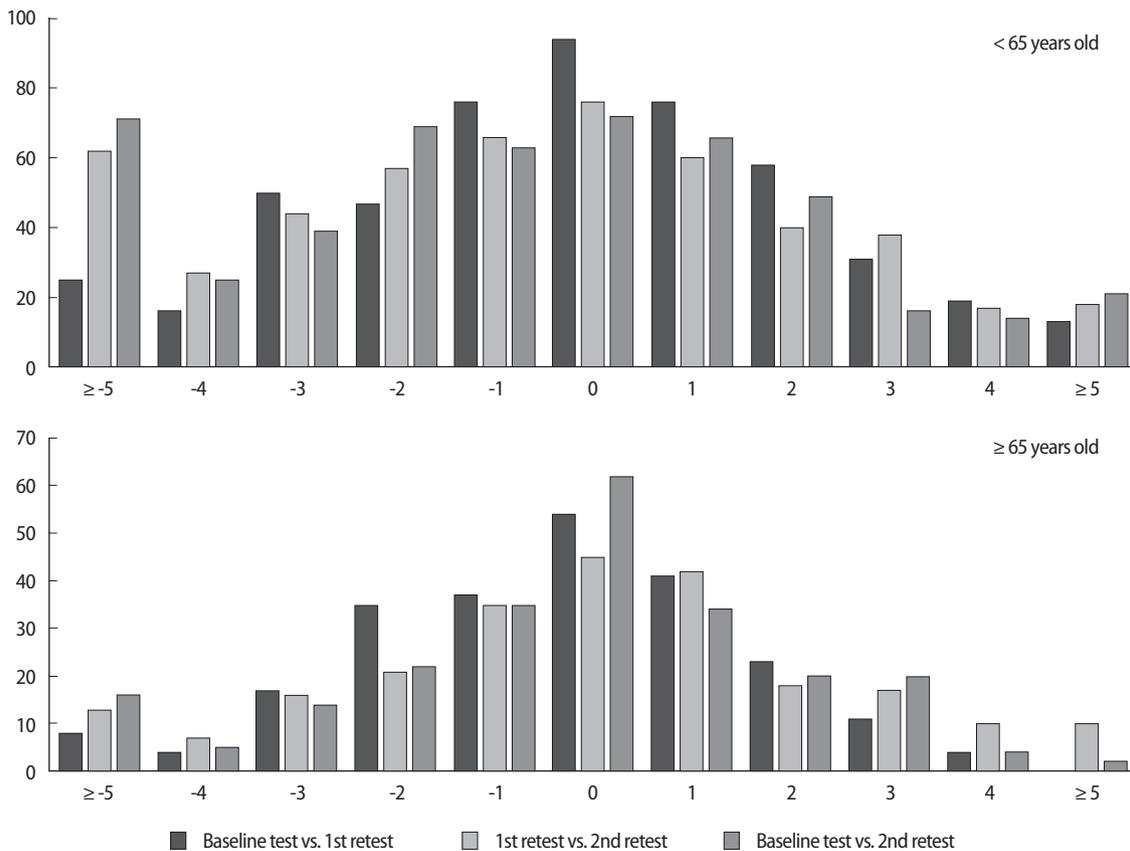


Fig. 1. Number of participants on each difference of score.

변화가 전혀 없거나 ± 1 점의 작은 변화를 보인 피검자들은 중장년 집단에서는 39-56%, 노인집단에서는 40-49%에 불과하다는 것을 알 수 있었다. 기저검사에 비해서 재검사 시에 ± 3 점 이상의 변화를 보인 피검자들은 중장년집단에서는 19-29%, 노인집단에서는 31-41%로 노인집단이 재검사시에 더 많은 점수 변화를 나타내는 경향이 있었다. 재검사 시 K-MMSE 점수가 5점 이상 저하된 피검자들은 중장년집단에서는 3-7%, 노인집단에서는 5-14%였고, 5점 이상 상승한 피검자들도 중장년집단에서는 0-4%, 노인집단에서는 3-4%에 해당하였다.

기저 검사와 재검사 점수의 RCI_{PE} 와 $SRBCI$ 변화유의구간을 검사-재검사 간격별로 각각 산출하였다(Table 4). 그 결과 중장년집단과 노인집단에서 RCI_{PE} 와 $SRBCI$ 는 유사한 변화 양상을 나타내었다. 즉, 두 집단은 각각 검사-재검사 간격이 2년인 경우보다 4년인 경우에 RCI_{PE} 와 $SRBCI$ 변화유의구간이 증가하였으나 검사-재검사 간격이 4년인 경우와 6년인 경우는 두 통계지표의 변화유의구간에 있어 차이가 없는 것으로 밝혀졌다.

K-MMSE 점수에 직접 반영될 수 있도록 RCI_{PE} 와 $SRBCI$ 변화유의구간을 정수로 만들면(Table 5), 중장년집단에서는 검사-재검사 간격이 2년일 때에는 재검사 점수가 4점이상 상승하거나 저하되었을 경우, 검사-재검사 간격이 4년과 6년일 때에는 재검사 점수가 5점이상 상승하거나 저하되었을 경우를 유의한 변화로 간주할 수 있다는

같은 결과를 두 통계지표에서 얻었다. 그러나 노인집단에서는 RCI_{PE} 와 $SRBCI$ 변화유의구간에 있어서 약간의 차이가 발견되었는데 검사-재검사 간격이 2년일 경우 RCI_{PE} 로는 재검사 점수가 5점 이상 저하되었거나 4점 이상 상승한 경우를, $SRBCI$ 로는 재검사 점수가 4점 이상 상승하거나 저하된 경우를 유의미한 변화인 간주하여야 한다는 결과를 얻었고, 검사-재검사 간격이 4년이나 6년인 경우에는 재검사 점수가 RCI_{PE} 로는 7점 이상 저하되었거나 5점 이상 상승한 경우를 $SRBCI$ 로는 6점 이상 저하되었거나 상승한 경우를 유의미한 변화로 간주하여야 한다는 결과를 얻었다. 이 결과를 토대로 본 연구에 참여한 지역사회 노인들 중 RCI_{PE} 와 $SRBCI$ 각 기준 점수에 해당하는 피검자들의 분포를 확인하였다. 그 결과, 중장년집단에서는 RCI_{PE} 나 $SRBCI$ 기준이 동일한 것으로 확인되었으므로 동일한 기준을 적용하였을 때 K-MMSE 점수가 유의하게 저하된 피검자들은 검사-재검사 간격에 따라 4-7명(2-3%)이었고 K-MMSE 점수가 오히려 유의하게 증가한 피검자들은 2-4명(1-2%)에 불과하였다. 그러나 노인집단에서는 RCI_{PE} 기준 점수를 적용하였을 때에는 5-12명(1-2%)이 재검사시에 K-MMSE 점수가 유의하게 저하되었고, 8-19명(2-4%)은 점수가 상승한 것으로 밝혀졌으나, $SRBCI$ 의 기준 점수를 적용하였을 때에는 15-16명(3%)이 재검사 시 점수가 저하되었고, 4-19명(1-4%)이 점수가 상승한 것으로 밝혀졌다.

Table 4. Prediction intervals of reliable change scores across test-retest intervals

Participants	Test-retest interval (years)	RCI_{PE} (90% prediction interval)	RC-	RC+	$SRBCI$ (90% prediction interval)
< 65 years old (n = 234)	2	± 3.45	-3.83	3.07	± 3.29
	4	± 4.75	-4.79	4.71	± 4.34
	6	± 4.43	-4.84	4.02	± 4.30
≥ 65 years old (n = 505)	2	± 4.11	-4.30	3.92	± 3.71
	4	± 5.71	-6.58	4.84	± 5.37
	6	± 5.77	-6.83	4.71	± 5.36

Table 5. Significant change scores across test-retest intervals and number (%) of participants who showed significant score change in the K-MMSE

Participants	Significant change index	Test-retest interval (years)	Decreased		Increased	
			Score	n (%)	Score	n (%)
< 65 years old (n = 234)	RCI_{PE}	2	4	4 (1.7)	4	4 (1.7)
		4	5	7 (3.0)	5	4 (1.7)
		6	5	7 (3.0)	5	2 (0.9)
	$SRBCI$	2	4	4 (1.7)	4	4 (1.7)
		4	5	7 (3.0)	5	4 (1.7)
		6	5	7 (3.0)	5	2 (0.9)
≥ 65 years old (n = 505)	RCI_{PE}	2	5	10 (2.0)	4	19 (3.8)
		4	7	5 (1.0)	5	9 (1.8)
		6	7	12 (2.4)	5	8 (1.6)
	$SRBCI$	2	4	16 (3.2)	4	19 (3.8)
		4	6	16 (3.2)	6	4 (0.8)
		6	6	15 (3.0)	6	9 (1.8)

고 찰

본 연구는 지역사회에서 치매의 조기발견을 위하여 중장년층과 노년층을 대상으로 K-MMSE가 반복 시행되었을 때, 어느 정도의 점수 변화를 유의한 변화로 인정할 수 있는지 그 기준을 제시하고자 수행되었다.

흔히 검사를 반복하면 연습효과로 인해서 점수가 상승할 것이라고 기대할 수 있고 실제로 3개월 간격으로 MMSE를 실시하였을 때 연습효과로 인해서 재검사 점수가 상승하였다는 연구 결과가 보고된 바 있다[11]. 그러나 본 연구에서는 2, 4, 그리고 6년 간격으로 검사를 반복할 때마다 K-MMSE 점수가 계속 저하되었고 45-64세의 중장년집단 보다 65세 이상인 노인들로 구성된 노인집단에서 그 경향이 더욱 두드러지게 관찰되었다. 이러한 결과는 중장년층보다 65세 이상의 노인집단에서 더 빠르게 진행되고 있는 정상적인 노화과정에 의해서 전반적인 인지기능도 중장년집단보다 노인집단에서 더 빠르게 저하된다는 사실을 반영하는 것으로 보인다. 또한 다른 한편으로는 2년, 4년 또는 6년이라는 비교적 긴 검사-재검사 간격이 연습효과의 영향을 충분히 둔화시켰을 가능성도 배제할 수 없다. 단, 본 연구에서 2009년에 실시된 3차 검사 결과를 2003년에 실시된 1차 검사(기저검사) 결과와 비교하면서 검사-재검사 기간이 6년인 것으로 간주하였으나 2005년에 2차 검사가 실시된 바 있으므로 중도에 재검사가 한번도 실시된 적이 없는 경우와 같은 결과로 간주할 수 있을지는 의문이다.

본 연구에서는 2, 4, 6년 후 K-MMSE 재검사 점수가 이전 검사의 점수보다 유의미하게 저하되었음을 검사-재검사 점수가 차이가 있는지를 결정하는 전통적인 방법인 집단 평균 점수의 비교를 통해서 확인하였다. 통계적으로는 유의한 차이가 있음이 확인되었지만 평균 점수들간의 실제 점수 차이는 1점 이하였고 기저점수와 백분율 비교에서도 0.1-3.9%의 차이만이 발견되었다. 그러나, 피검자들 개개인의 점수 변화를 살펴본 결과, 기저검사에 비해서 재검사 시에 ±3점 이상의 변화를 보인 피검자들이 중장년집단에서는 19-29%, 노인집단에서는 31-41%에 이르는 것으로 밝혀졌다. 이러한 결과는, 기저검사에서 얻어진 점수와 재검사 시에 얻어진 점수의 차이를 비교할 때, 평균 점수는 실제 피검자들의 점수 변화를 충분히 반영하지 못한다는 사실을 시사한다. 그러므로 같은 검사가 반복적으로 실시되었을 때 재검사 결과에 영향을 미치는 "검사 자체와 관련된 변인"과 "검사 상황과 관련된 변인" 및 인구학적 변인들을 비롯한 기타 변인들을 통제하고 피검자들 각각의 점수 변화를 반영한 통계적 기법이 필요하다.

따라서 본 연구에서는 개별 피검자들의 점수 변화를 반영하는 대표적인 통계지표인 RCI_{PE}와 SRBCI를 산출하고 두 지표를 비교하였다. 그 결과, 중장년집단과 노인집단에서 모두 검사-재검사 간격

이 2년인 경우보다 4년인 경우에 RCI_{PE}와 SRBCI 변화유의구간이 증가하였으나 검사-재검사 간격이 4년인 경우와 6년인 경우는 두 통계지표의 변화유의구간에 있어 차이가 없는 것으로 밝혀졌다. 두 지표의 변화유의구간을 K-MMSE 점수에 적용할 수 있도록 정수로 만들어서 비교한 결과 중장년집단의 경우에는 각각의 검사 간격에서 두 통계지표값이 동일한 것으로 확인되었으나, 노인집단의 경우에는 두 통계지표에서 다음과 같은 약간의 차이가 발견되었다. 즉, 재검사 점수가 저하된 경우에는 검사 간격에 관계없이 SRBCI의 변화유의구간이 RCI_{PE}보다 조금 더 넓은 것으로 밝혀졌는데 이는 SRBCI를 사용하면 RCI_{PE}를 사용할 때보다 더 많은 피검자들이 재검사에서 K-MMSE 점수가 유의하게 저하되었다고 판정된다는 것을 의미한다. 반면 재검사 점수가 증가하였을 때에는 저하되었을 때와는 반대로 RCI_{PE}의 변화유의구간이 SRBCI보다 조금 더 넓은 것으로 밝혀졌고 이는 SRBCI를 사용하면 RCI_{PE}를 사용할 때보다 더 적은 수의 피검자들이 재검사에서 K-MMSE 점수가 유의하게 증가되었다고 판정된다는 것을 의미한다. 그러므로 지역사회에 살고 있는 노인들을 대상으로 K-MMSE를 반복 실시하고 그 점수 변화를 통해서 치매를 조기에 탐지할 목적이라면 SRBCI가 RCI_{PE}보다 민감도가 더 높다고 할 수 있을 것이다.

따라서 45세 이상 65세 미만의 중장년층의 경우에는 검사-재검사 간격이 2년일 경우에는 4점 이상, 4년과 6년인 경우는 5점 이상 K-MMSE 점수가 상승하거나 저하된 경우를 유의미한 변화로 간주하여야 한다는 기준을 적용할 수 있다. 그러나 65세 이상의 노인층의 경우에는 SRBCI를 사용하는가 RCI_{PE}를 사용하는가에 따라서 약간의 차이가 발견되었는데 앞에서 기술한 바와 같이 K-MMSE가 지역사회에서 인지기능저하를 나타내는 치매위험군을 조기에 발견할 목적으로 반복 실시된 것이라면 두 통계지표 중에서 민감도가 더 높은 SRBCI를 사용하여 검사 간격이 2년 일 경우에는 4점 이상, 4년과 6년일 경우에는 6점 이상 K-MMSE 점수가 저하된 경우를 유의미한 변화로 간주하는 것이 더 타당할 것으로 사료된다. 여기에서 한 가지 주목할만한 결과는 검사 간격이 2년인 경우에는 중장년집단과 노인집단의 두 통계지표가 거의 차이가 없지만 검사 간격이 4년이나 6년일 경우에는 중장년집단보다 노인집단의 변화유의구간이 더 좁다는 것이다. 이 결과는 노인들의 경우에는 중장년층에 비해서 같은 기간 동안 K-MMSE 점수가 더 많이 감소하여야만 유의한 변화로 간주할 수 있다는 것을 의미하는데 이는 정상적인 노화에 따른 인지기능저하가 중장년층보다 노인층에서 가속화되는 것에 기인한 결과로 사료되며, 나이 범주에 따라서 다른 변화유의구간이 적용되어야 한다는 사실을 시사한다.

본 연구의 결과를 외국에서 정상 노인들을 대상으로 수행된 연구 결과들과 비교해 보면, SRBCI와 유사한 회귀에 기반한 변화지표를 산출한 Hensel 등의 연구에서는 75세 이상의 건강한 노인들 119

명에게 각각 1.5년의 간격을 두고 MMSE를 6회 반복 실시한 결과 1.5년 간격에서는 2-4점 이상을 유의한 변화로 간주하여야 한다고 보고하였고[12], Tombaugh는 5년 간격으로 실시된다면 3-4점 이상을 유의한 변화로 간주하여야 한다고 보고하였다[11]. 반면에 Schmand 등은 1년 간격으로 실시된 재검사에서 MMSE 점수가 5점 이상 저하된 경우에 치매를 의심할 수 있다고 주장하였는데 유독 Schmand 등의 연구에서 점수 기준이 높은 이유는 피검자들 중에 1년내에 치매로 진행된 사람들을 포함하고 있었기 때문으로 밝혀졌다[4]. 한편 Clark 등[18]은 알츠하이머병 환자들을 대상으로 MMSE를 반복 시행한 결과, 검사-재검사 간격이 1년일 때, 평균 약 2.8점 저하되었다고 보고하였으나 환자들 간에 점수 변화의 개인차가 매우 크고 연습 효과나 측정 오류 등을 고려하지 않았으므로 2.8점을 신뢰로운 변화 점수로 보기 어렵다고 자인하였다.

본 연구는 2년, 4년, 그리고 6년의 비교적 긴 시간 간격을 두고 지역사회의 거주하고 있는 대단위의 건강한 중장년층과 노년층에게 K-MMSE 검사를 반복 실시하여, K-MMSE 점수 변화의 유의성을 판단할 수 있는 기준을 제공하였고, 중장년층과 노년층에게 서로 다른 기준이 적용되어야 한다는 새로운 사실을 발견하였다는 데 의의를 지닌다. 그러나 본 연구에서 밝혀진 기준이 실제로 치매의 조기탐지에 유용하게 사용될 수 있는 임상적 가치가 있을지는 장기적인 종단 연구를 통해서 추후에 경험적으로 확인되어야 할 것이다. 또한 본 연구에서 밝혀진 기준은 정상적인 노화에 의한 K-MMSE 점수변화와 병리적이거나 비정상적인(또는 비정상이 의심되는) K-MMSE 점수변화를 구분할 목적으로 사용될 수 있는 기준이지 경도인지장애 환자나 알츠하이머형 치매환자와 같은 임상집단의 인지 기능 변화를 판단하는데 적용될 수 있는 기준이 아니다. 따라서 임상집단을 대상으로 병의 진행이나 치료 효과에 대한 정보를 얻기 위해서 인지 기능 변화 여부를 판단하는데 사용될 수 있는 기준은 해당 임상 집단을 대상으로 한 후속 연구를 통하여 새로 밝혀져야 할 것이다. 끝으로, 본 연구에 참여한 피검자들은 지역사회에서 별다른 어려움없이 생활하고 있는 사람들에게 건강선별설문을 실시하고 그 결과를 바탕으로 선정되었으므로 이들 가운데에는 일상생활에 지장을 주지 않을 정도의 경한 인지저하, 즉 경도인지장애에 해당하는 사람들이 일부 포함되어 있을 가능성을 배제할 수 없다. 후속연구에서는 좀 더 자세한 신경심리평가를 통해서 보다 정선된 피검자들을 대상으로 한 연구가 수행되기를 기대한다.

참고문헌

1. Folstein ME, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *J Psy-*

- chiat Resear* 1975; 12: 189-98.
2. Bondi MW, Salmon DP, Kaszniak AW. *The neuropsychology of dementia*. In Grant I, Adams KM (Eds), *Neuropsychological assessment of neuropsychiatric disorders*. New York: Oxford University Press, 1996.
 3. Kang Y. A normative study of the Korean-Mini Mental State Examination (K-MMSE) in the elderly. *Korean J Psychol* 2006; 25: 1-12.
 4. Schmand B, Lindeboom J, Launer L, Dinkgreve, Hooijer C, Jonker C. What is a significant score change on the Mini-Mental State examination? *Int J Geriatr Psychiatry* 1995; 10: 411-4.
 5. Heilbrunner RL, Sweet JJ, Attix DK, Krull KR, Henry GK, Hart RP. Official position of the American Academy of Clinical Neuropsychology on serial neuropsychological assessments: the utility and challenges of repeat test administrations in clinical and forensic contexts. *Clin Neuropsychol* 2010; 24: 1267-78.
 6. Duff K. Evidence-based indicators of neuropsychological change in the individual patient: relevant concepts and methods. *Arch Clin Neuropsychol* 2012; 27: 248-61.
 7. Dirks J. The effect of a commercial game on children's block design scores in the WISC-R IQ test. *Intelligence* 1892; 6: 109-23.
 8. Rapport LJ, Brines DB, Axelrod BN, Theisen ME. Full scale IQ as a mediator of practice effects: the rich get richer. *Clin Neuropsychol* 1997; 11: 375-80.
 9. Beglinger LJ, Gaydos B, Tangphao-Daniels O, Duff K, Kareken DA, Crawford J, et al. Practice effects and the use of alternate forms in serial neuropsychological testing. *Arch Clin Neuropsychol* 2005; 20: 517-29.
 10. Heaton RK, Temkin N, Dikmen S, Abitabile N, Taylor MJ, Marcotte TD, et al. Detecting change: a comparison of three neuropsychological methods, using normal and clinical samples. *Arch Clin Neuropsychol* 2001; 16: 75-91.
 11. Tombaugh TN. Test-retest reliable coefficients and 5-year change scores for the MMSE and 3MS. *Arch Clin Neuropsychol* 2005; 20: 485-503.
 12. Hensel A, Angermeyer MC, Riedel-Heller SG. Measuring cognitive change in older adults: reliable change in dices for the MMSE. *J Neurol Neurosurg Psychiatry* 2007; 78: 1298-303.
 13. Frerichs RJ, Tuokko HA. A comparison of methods for measuring cognitive change in older adults. *Arch Clin Neuropsychol* 2005; 20: 321-33.
 14. Chelune GJ, Naugle RI, Luders H, Sedlak J, Awad IA. Individual change after epilepsy surgery: practice effects and base-rate information. *Neuropsychology* 1993; 7: 41-52.
 15. McSweeney AJ, Naugle RI, Chelune GJ, Luders H. T scores for change: an illustration of a regression approach to depicting change in clinical neuropsychology. *Clin Neuropsychol* 1993; 7: 300-12.

16. Hallym Institute of Aging. *Quality of life in the Korean elderly*. Chuncheon: Hallym University Press, 2006.
17. Christensen KJ, Multhaup KS, Nordstrom S, Voss K. *A cognitive battery for dementia: development and measurement characteristics*. *J Consult Clin Psychol* 1991; 3: 168-74.
18. Clark CM, Sheppard L, Fillenbaum GG, Galasko D, Morris JC, Koss E, et al. *Variability in Annual Mini-Mental State Examination score in patients with probable Alzheimer Disease*. *Arch Neurol* 1999; 56: 857-62.