

한글판 수정바텔지수 체크리스트의 평가자간 신뢰도 연구

이근호*, 구경협*, 신선희*, 안아라*, 손미남*, 조은희*, 김연희**

*삼성서울병원 재활의학과 작업치료실

**성균관대의대 삼성서울병원 재활의학교실

국문초록

목적 : 한글판 수정바텔지수 체크리스트를 고안하고 기존 한글판 수정바텔지수와 신뢰도를 비교 평가하기 위함이다.

연구방법 : 연구대상은 뇌병변 환자 30명을 대상으로 하였다. 한글판 수정바텔지수 체크리스트(이하 K-MBI 체크리스트)는 기존의 한글판 수정바텔지수(기존 K-MBI)와 수정바텔지수(MBI) 원본을 참고하여 각 항목별 평가 내용을 세분화 및 체크리스트화 하였다. 30명의 환자를 대상으로 K-MBI 체크리스트를 사용하는 그룹과 기존 K-MBI를 사용하는 그룹으로 나누어 각 그룹에 2명의 치료사를 배정하였다. 결과분석은 평가자간 신뢰도를 알아보기 위해 총점수준에서 급간내상관계수를 구하였으며, 세부항목 수준에서 카파검정을 시행하였다.

결과 : 총점수준에서의 신뢰도는 K-MBI 체크리스트(ICC= .981)와 기존 K-MBI(ICC= .973)를 사용한 두 그룹 모두에서 매우 높은 검사자간 신뢰도를 보였다($p<.001$). 세부항목별 일치도에서는 K-MBI 체크리스트를 사용한 그룹과 기존 K-MBI의 경우 모두 4개 항목에서 탁월한 일치도를 보였다. Mean K값은 K-MBI 체크리스트(.725)가 K-MBI(.699) 보다 높게 나타났다.

결론 : K-MBI 체크리스트와 기존의 K-MBI 모두 총점수준에서 높은 수준의 평가자간 신뢰도를 보였다. 세부항목수준에서 역시 두 그룹 모두 전 항목에서 상당한 수준 이상의 신뢰도를 확보하였으나, 항목 간에는 다소 차이가 있었다.

주제어 : 일상생활동작, 체크리스트, 평가자간 신뢰도, 한글판 수정바텔지수

I. 서 론

수정바텔지수(Modified Barthel Index; MBI)는 Mahoney와 Barthel(1965)이 일상생활동작 수행의 독립 정도를 평가하기 위해 개발한 바텔지수(Barthel Index)를 여러 차례에 걸쳐 수정 보완한 평가도구이다. 민감도가 0.9이상으로 환자의 일상생활동작 수행 능력을 평가하는데 매우 적절한 도구로 여겨져 널리 사용되고 있다(Shah, Vanclay, & Cooper, 1989). 이와 더불어 한글판 수정바텔지수(Korean version of Modified Barthel Index; K-MBI)는 Shah 등(1989)에 의해 개정된 5점 척도를 기반

으로 대한뇌신경재활학회에서 우리나라에 실정에 맞게 수정 번역하여, 이미 높은 신뢰도와 타당도를 확보하고 있으며, 현재 국내에서 기능적 독립성 평가(Functional Independent Measure; FIM)와 함께 일상생활수행도를 평가하는 가장 보편화된 평가도구로 사용되고 있다(박소연과 유은영, 2002; 정한영 등, 2007).

재활 치료를 받는 환자의 장애에 대한 기능평가는 환자의 일상생활과 사회활동을 진단하는 방법으로 매우 중요한 부분을 차지한다(Middleton, Truman, & Geraghty, 1998). Delisa와 Gans(2004)는 기능평가가 갖추어야 할 조건으로 표준화되어있고, 척도화할 수 있으며, 신뢰도가

검증되어야 하고, 이해하기 쉬워야 한다고 이야기 하고 있다. MBI는 여러연구를 통해 신뢰도 및 타당도를 검증받은 평가도구이며(Hocking, Williams, Broad, & Baskett, 1999; Fricke & Unsworth, 1997; Leung, Chan, & Shah, 2007; Loewen & Anderson, 1990), 이러한 점에서 바람직한 평가도구라 할 수 있다(남궁영, 이한석과 장기연, 1996).

K-MBI는 임상에서 장애등급을 판정하는데 기준이 되는 평가도구로서(보건복지가족부, 2011), 평가자간의 신뢰도가 그 어느 평가도구보다 중요한 부분이라 할 수 있다. 장애 등급은 환자들에게 경제적인 혜택과 연관되는 부분으로 매우 민감한 부분이라 할 수 있다. 따라서 환자의 기능 상태가 객관적으로 평가될 뿐 아니라 다른 검사자에 의해서도 같은 결과가 나와야 함은 당연한 일이라 할 수 있다(Donaldson, Wagner, & Gresham, 1973).

MBI와 관련된 연구들을 보면 대부분 환자의 수행을 직접 관찰하거나 녹화된 비디오를 통해 실제 수행을 바탕으로 한 평가가 이루어지고 있다(정한영 등, 2007; Fricke & Unsworth, 1997; Hocking et al., 1999; Leung et al., 2007; Loewen & Anderson, 1990). 하지만 실제 임상에서는 평가에 필요한 시간 및 장소가 제한되어있고, 환자의 주변환경 및 보조도구 사용여부 등 많은 부분을 고려해야 하는 등의 어려움 때문에 모든 항목을 직접 관찰하여 평가하기에는 제한이 있다. 결과적으로 짧은 시간 동안 면담 및 간단한 관찰을 통해 평가를 진행하고 그로인해 평가자의 주관적 판단이 개입되어 평가자간의 일치도가 떨어지는 것을 자주 경험한다. 외국 역시 이러한 문제점을 제기 하고, 그 대안으로 전화 면담을 통한 연구나(Heuschmann et al., 2005; Korner-Bitensky & Wood-Dauphine, 1995; Shinar et al., 1987) 우편을 통한 자기기입식 방법(Gompertz, Pound, & Ebrahim, 1994; Heuschmann et al., 2005)등을 개발하여 연구하고 있지만, 국내에선 이러한 MBI 변형 연구나 면담중심 연구는 미비한 실정이다.

따라서 본 연구팀은 임상에서 K-MBI의 신뢰도 및 하위 항목 간 일치도를 높이기 위하여, 체크리스트 형식의 한글판 수정바텔지수 체크리스트(이하 K-MBI 체크리스트)를 고안하게 되었다. 체크리스트 형식은 수행을 점검할 때, 세부적인 항목의 고려가 누락되는 오류를 예방하여 차후 수행을 개선 및 관리 하는 것에 효과적이라고 증명된 바 있다(Boorman, 2001; Helmreich, 2000; Pronovost

et al., 2006; Wolff, Taylor, & McCabe, 2004). 국내 연구 결과를 보면, 조남중(2008)은 RULA(Rapid Upper Limb Assessment)와 REBA(Rapid Entire Body Assessment)를 이용하여 조선업의 근골격계부담 평가를 위한 Quick Checklist를 개발하여 빠르고 정확한 평가가 가능하다고 보고하였고, 주혜선(2008)은 외상 후 위기를 체크리스트의 세부 항목을 토대로 객관적으로 평가 및 확인하여 외상 후 스트레스성장애로 악화되는 것을 예방할 수 있음을 보고하였다.

본 연구는 기존의 K-MBI를 수정 및 보완하여, 짧은 시간동안의 면담을 통해 신뢰성 있는 평가를 진행할 수 있는 K-MBI 체크리스트를 고안하고 K-MBI와 신뢰도를 비교 평가 하는데 목적이 있다.

II. 연구 방법

1. K MBI 체크리스트의 고안

K-MBI 체크리스트의 고안은 2010년 4월부터 비공식적인 합의 개발 프로세스(informal consensus development process)를 따라 시행하였다. 국내 종합병원에 재직 중인 재활의학과의 교수, 전문의, 전공의 및 작업치료사로 구성된 K-MBI 체크리스트 개발팀이 구성되었으며, MBI 관련 세미나 7회, 연구논문 고찰 2회, 시험판의 임상적용 후 수정 및 보완작업을 위한 토론 3회가 6개월에 걸쳐 이루어졌다. 타당도와 신뢰도가 검증된 제 5판 MBI 원문(Shah et al., 1989; Shah & Muncer, 2000)과 K-MBI(정한영 등, 2007)를 비교 분석하고, 등급별 주요 내용과 설명의 서술식 문장을 간결하게 나누어 해당 여부를 객관적으로 파악할 수 있도록 체크리스트화 하였다. 또한 항목 서두에 필수적으로 고려해야 하는 항목을 제시하여 평가 진행 시 기준이 되도록 하였다. 예를 들어 목욕하기의 경우 목욕의 단계를 '이동하기, 욕조 또는 의자에 앉기, 비누칠하기, 물로 씻기, 목욕 타월 세척하기, 수건으로 닦아 말리기'와 같이 일련의 과정으로 제시하여, 실제 일상생활동작 수행 평가 시 모든 부분을 고려할 수 있도록 하였다(표 1). 이렇게 제작한 시험판 K-MBI 체크리스트를 임상적용을 거쳐 수정 및 보완하였으며, 2010년 10월 최종적인 K-MBI 체크리스트를 완성하였다. K-MBI 체크리스트를 고안함에 있

표 1. 항목별 수정의 예

K-MBI	K-MBI 체크리스트
목욕하기	[목욕의 단계] 이동하기/ 욕조 또는 의자에 앉기/ 비누칠 하기/ 물로 씻기/ 목욕 타올 세척하기/ 수건으로 닦아 말리기 [목욕용품] 샤워기/ 욕조/ 목욕타월/ 샴푸/ 바디샴푸 등
샤워나 목욕을 위해 이동하거나 씻고 말리는데 도움이 필요하다; 환자 상태나 질환 등으로 목욕을 완벽하게 할 수 없는 것을 포함한다.	환자는 목욕의 모든 단계에서 부분적 도움이 필요하다.
3등급 샤워나 목욕 또는 씻기/말리기를 위해서 이동시 도움이 요구된다. 욕조에 앉기, 비누칠하기, 수건으로 닦기, 수건세척, 팔다리를 씻을 때 도움이 요구된다. 환자는 힌트, 유도 또는 감독을 필요로 한다. ☞ 적용범위가 넓음.	☐ 도움을 받아 가슴과 팔을 포함한 다른 부위를 씻을 수 있다. ☐ 목욕의 모든 단계에서 부분적 도움이 필요하다. · 이동/ 욕조 or 의자에 앉기/ 비누칠하기/ 물로 씻기/ 목욕 타월 세척하기/ 수건으로 닦아 말리기 · 환자 상태나 질환 등으로 목욕을 완벽하게 할 수 없는 것을 포함한다. ☐ 힌트, 유도, 감독을 필요로 한다.

여 기존 평가의 타당도를 유지하기 위해 가급적 원문의 문장을 최대한 반영하였다.

2. 연구 대상

본 연구는 2011년 6월부터 동년 7월 29까지 서울시 소재 ○○병원의 뇌병변 환자 중 본 연구 참여에 동의한 30명을 대상으로 하였다. 연구 대상의 포함 기준은 손상 부위와 횡수에 제한없이 K-MBI와 K-MBI 체크리스트를 평가하는데 언어적, 인지적 문제가 없는 뇌병변 환자로 하였다. 본 연구는 진행 전 재활의학과 내에 설치된 임상연구위원회의 심의를 거쳤으며, 모든 연구 대상자에게 연구에 대한 충분한 설명을 한 후 서면 동의를 받았다. 연구대상자의 일반적 특성은 표 2와 같다.

3. 연구 절차

서면으로 작성된 연구 참여 동의서의 설명을 듣고, 자발적으로 참여를 결정한 뇌병변 환자 30명을 모집하였다. 기존의 K-MBI를 사용할 그룹과 K-MBI 체크리스트를

사용할 그룹으로 나누어 각 그룹에 2명의 치료사를 배정하였다. 평가는 그룹 간 하루의 시간 간격을 두고 평가하였고, 그룹 내 치료사간 평가순서 및 그룹 간 평가순서는 역 균형화(counter-balance)하여 순서에 의한 영향을 배제하였다. 평가는 한명의 평가자가 환자 또는 보호자를 일대일 면담하는 것으로 하였으며, 평가가 끝나면 같은 그룹의 다른 평가자가 다음 치료시간에 평가를 진행하는 것으로 하였다. 한 그룹의 평가가 끝나면 하루의 시간적 간격을 두고, 다른 그룹의 평가가 같은 방법으로 진행되었다. 즉 한명의 환자 및 보호자는 총 4번의 평가를 받게 되며, 이러한 과정에서 오는 연구대상자에 대한 피로감을 고려, 하루의 시간적 간격을 두고 평가하게 되었다. K-MBI 체크리스트 그룹의 경우, 평가자가 처음으로 접하는 평가도구이므로 평가 시 충분한 숙지가 이루어질 수 있도록 K-MBI 체크리스트 개발에 참여한 치료사들로 구성하였다. 반면 K-MBI 그룹의 경우에는 K-MBI 체크리스트 개발에 참여하지 않은 치료사로 구성하였다. 이와 별도로 각 그룹의 치료사들에게 2시간의 평가도구에 대한 추가 교육을 실시하였으며, 두 그룹 모두 임상경력이 2년인 작업치료사로 하였다. 본 연구에서 기존의 K-MBI 및 K-MBI

표 2. 대상자의 일반적 특성

(N=30)

항목	세부항목	빈도	비율(%)
성별	남	11	36.7
	여	19	63.3
나이	10대	3	10.0
	20대	1	3.3
	30대	2	6.7
	40대	1	3.3
	50대	5	16.7
	60대	7	23.3
	70대	6	20.0
	80대	5	16.7
진단명	stroke	19	63.3
	tumor	5	16.7
	cerebral palsy	1	3.3
	traumatic brain injury	4	13.3
	hypoxic brain damage	1	3.3

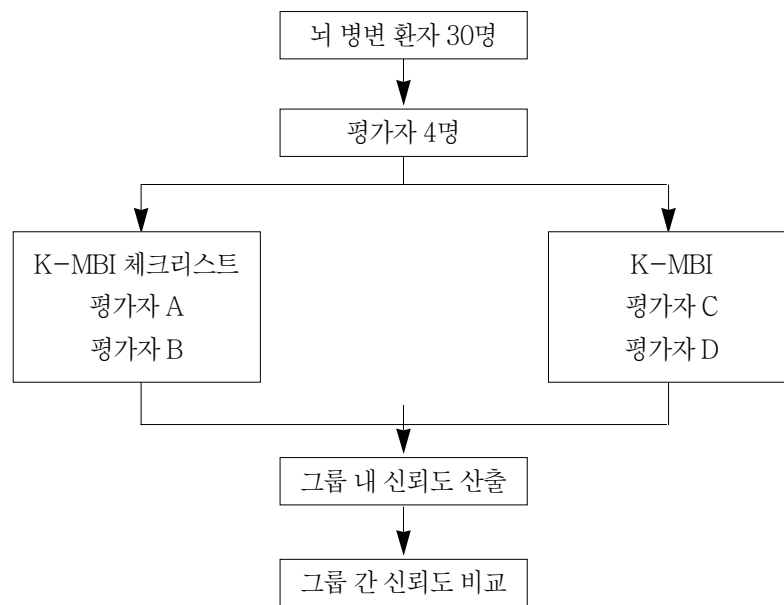


그림 1. 연구 절차

체크리스트 평가는 환자를 직접 면담하여 실시하였으며, 그렇지 못할 경우 환자를 돌보고 있는 가족이나 간병인의 면담을 통해 진행하였다. 환자를 직접 면담하지 못할 경우 K-MBI 및 K-MBI 체크리스트 모두 동일한 사람을 대상으로 평가하는 것을 원칙으로 하였다(그림 1).

4. 연구 도구

1) K-MBI

바텔지수는 1965년 호주의 의사 Mahoney와 물리치료사 Barthel이 고안한 평가도구로서 세계적으로 널리 사용되고 있으며, 신뢰도와 타당도가 검증된 평가도구이다(Mahoney & Barthel, 1965; Roden-Jullig, Britton, Gustafsson, & Fugl-Meyer, 1994). 하지만 낮은 민감도로 인해 여러 편의 MBI가 발표되어 혼용되어 사용되고 있다. K-MBI는 그 중 민감도가 가장 높고, 내적 일치도가 높으며, 사용 방법에 대한 자세한 주석이 보고된 제5판(Shah et al., 1989)을 원본으로 채택하여, 정한영 등(2007)이 한국판으로 표준화한 평가도구이다. K-MBI는 10가지의 구체적인 일상생활활동으로 구성되어 있으며, 신변처리 동작(self-care index) 7개 항목과 가동능력(mobility index) 3개 항목으로 구성되어 있다. 각 항목은 5등급으로 차등 점수화 되어있고, 항목별 내용의 비중에 따라 가중치가 부가되어 있다. 10개 항목을 합산하였을 때 100점을 만점으로 하고, 평가 세부항목은 개인위생, 목욕하기, 식사하기, 용변처리, 계단오르내리기, 옷입고벗기, 배변조절, 소변조절, 보행/휠체어, 의자/침대 이동으로 구성되어 있다(Shah 등, 1989).

2) K-MBI 체크리스트

K-MBI 체크리스트는 MBI 제5판 원문(Shah et al., 1989; Shah & Muncer, 2000)과 대한뇌신경재활학회에서 제작한 K-MBI(정한영 등, 2007)를 참고하여 제작하였다. K-MBI 체크리스트에서는 각 세부항목별 평가 항목을 세분화 하였으며, 세부항목의 동작을 분석하여 각 동작별 필수 평가 요소들을 가급적 원문의 문장을 최대한 반영하여 체크리스트화 하였다.

5. 자료 분석

수집된 자료는 PASW 18.0 network version을 사용하여 분석하였다. 통계적 유의성 검증을 위한 유의수준은 $p < .01$ 로 하였으며, 환자의 일반적 특성은 기술통계를 이용하여 분석하였다. K-MBI와 K-MBI 체크리스트의 총점수준 그룹 내 평가자간 신뢰도를 알아 보기위해 two-way mixed 모델을 사용하여 급간내상관계수(Intraclass Correlation Coefficient; ICC)를 구하였으며, 항목 간 일치도를 알아 보기위해 카과검정(Kappa statistics)을 시행하였다.

III. 연구 결과

1. K MBI 체크리스트 사용 그룹과 기존의 K MBI 사용 그룹의 검사자간 신뢰도

각 그룹 내 총점수준의 평가자간 신뢰도를 구하기 위해 급간내상관계수를 구하였으며, K-MBI 체크리스트가 .981, 기존의 K-MBI가 .973으로 두 가지 방법 모두 매우 높은 신뢰도를 보였다($p < .001$)(표 3). 세부항목별 내적 일치도를 알아보기 위해 Kappa value를 구하였으며 그 결과는 표 3과 같다. K-MBI 체크리스트 사용 그룹이 .565 ~ .905로 나타났으며, 기존의 K-MBI 사용 그룹이 .437 ~ .864로 나타났다.

2. 신뢰도 평가 기준에 따른 그룹 내 항목별 신뢰도

각 그룹 내 세부항목별 신뢰도 평가 기준은 Fleiss, Levin과 Paik(2003)가 제시한 대로 Kappa값이 .75 이상이면 탁월한 신뢰도(excellent), .40 ~ .75이면 상당한 신뢰도(fair), .40 미만이면 불충분한 신뢰도(poor)로 분류하였다. 본 연구에서는 두 그룹 모두 탁월한 신뢰도 항목이 넷, 상당한 신뢰도 항목이 여섯으로 항목은 상이하지만 기준에 따른 개수는 같은 수치를 나타내었다. Mean K값은 K-MBI 체크리스트가 .725로 기존 K-MBI .699에 비해 다소 높은 값을 보였다(표 3).

표 3. K MBI 체크리스트 그룹과 K MBI 그룹의 그룹 내 평가자간 일치도 분석결과

(N=30)

일반적 특성	구분	명수
개인위생	.716**	.437**
목욕하기	.621**	.764**
식사하기	.762**	.630**
용변처리	.688**	.687**
계단오르내리기	.720**	.858**
옷입고벗기	.565**	.712**
배변조절	.728**	.723**
배뇨조절	.781**	.835*
보행/휠체어	.759**	.481**
의자/침대 이동	.905**	.864**
ICCb	.981**	.973**

^aK: kappa value^bICC: 급간내상관계수 (intraclass correlation coefficient)** $p<.001$

표 4. 신뢰도 평가 기준에 따른 두그룹의 세부항목 개수

(N=30)

	K-MBI 체크리스트	기존 K-MBI
탁월한(excellent : .75~1)	4	4
상당한(fair : .40~.75)	6	6
불충분한(poor : 0~.40)	0	0
Mean Ka	.725	.699

^aK: kappa value

표 5. 두 그룹 간 총점비교

(N=30)

	K-MBI 체크리스트		기존 K-MBI	
평가자	A	B	C	D
$M \pm SD$	34.13 ± 2.21	34.90 ± 24.88	38.26 ± 25.16	40.93 ± 27.15

3. K MBI 체크리스트 그룹과 K MBI 그룹의 각 평가자의 총점 비교

두 그룹 평가자들의 MBI 총점 평균은 표 4와 같으며, K-MBI 체크리스트 그룹이 기존 K-MBI 그룹에 비해 다소 낮게 나타났다.

IV. 고 찰

본 연구는 면담중심의 MBI 평가 시 평가자간 신뢰도를 높일 수 있는 K-MBI 체크리스트를 고안하여 기존 K-MBI와 신뢰도를 비교하는 것을 목적으로 하였다. 연구결과 두 평가도구 모두 총점수준에서 높은 수준의 신뢰도를 보였으며, 항목 간 일치도에서는 K-MBI 체크리스트에서 다소 높은 일치도를 보였다.

MBI는 여러 나라에서 자국 문화에 맞게 수정 번역하여 사용하고 있으며(Kucukdeveci et al., 2000), 국내에서도 정한영 등(2007)이 우리나라문화에 맞게 수정 보완하여 통계적으로 표준화된 K-MBI를 개발, 보다 정확한 일상생활수행능력 평가를 가능하도록 하였다. 하지만 MBI를 사용하면서 느끼게 되는 대표적인 어려움은 일상생활의 모든 활동을 관찰하여 평가하기가 어렵다는 점이다. 가령 옷입고벗기와 같은 항목은 옷을 가져오는 동작부터 입는 동작 모두를 고려해야 하며, 계절별 옷의 종류에 따른 수행여부와 성별의 차이에서 오는 옷의 종류도 고려해야 하기 때문에(Shah et al., 1989) 평가에 많은 시간과 노력을 요하게 된다. 특히 환자의 주변 환경이 일상생활수행에 영향을 미치게 되므로, 외래 환자의 경우 병원에 내원한 상태에서 그 환자의 집안환경에서의 일상생활 수행정도를 평가하는 것은 많은 제약이 따르게 된다. 결국 관찰에 의한 평가는 극히 제한적일 수밖에 없으며 대부분 면담에 의해서 평가가 이루어지고 있다. 이러한 점들을 생각하였을 때, 기존 K-MBI의 사용에 있어서 관찰이 아닌 면담 중심의 평가에 대한 논의가 필요한 상황이다.

본 연구 결과, 기존의 K-MBI와 K-MBI 체크리스트 모두 총점수준에서 높은 평가자간 신뢰도를 보이고 있다. 두 그룹 평가자들의 총점 평균은 K-MBI 체크리스트 그룹(A=34.13, B=34.90)이 K-MBI 그룹(C=38.26, D=40.93)보다 낮게 나타났다. 특히 다른 항목에 비해 대변조절(A=4.66, B=4.66, C=6.93, D=6.63)과 소변조절(A=5.06, B=5.43, C=7.20, D=7.10)항목에서 두 그룹 간 큰 점수 차이를 보이는데, 이는 K-MBI 체크리스트 그룹은 배변조절 항목에서 실변빈도, 배변자세, 회음부 위생을 모두 고려한 반면, K-MBI 그룹은 주로 실변빈도에 비중을 두어 평가하였고, 배뇨조절 항목 역시 K-MBI 체크리스트 그룹은 실뇨빈도, 배뇨자세, 기구관리 여부를 모두 고려하여 평가한 반면에 K-MBI 그룹은 실뇨빈도에만 중점을 두어 평가하였기 때문이다.

MBI는 각 항목들에서 고려되어야 하는 동작들이 완전히 독립되어 있지 않고 중복되어있는 경우가 많다. 예를 들어 용변처리항목에 포함되는 회음부 위생은 ‘배변조절’에서도 고려되어지는 사항이며, Fricke과 Unsworth(1997)의 연구에서도 이러한 점을 혼란스러운 부분으로 언급하고 있다. 또한 용변처리 항목에서는 착탈의, 이동, 옷 추스르기, 지퍼열고 닫기처럼 ‘옷입고벗기’ 항목에서 고

려해야하는 동작과 ‘개인위생’ 항목의 손 씻기 부분이 중복됨을 알 수 있다. 본 연구의 평가자들 역시 이 부분에 대한 혼란스러움을 보고하였다. 이러한 문제점을 해결하기 위해 K-MBI 체크리스트에서는 항목 간에 중복되는 동작들을 체크리스트화 하여 모든 항목에서 필수적으로 평가되도록 하였다.

그룹 내에서의 평가자들의 총점을 보면 두 그룹 모두에서 상당히 일치하고 있음을 확인할 수 있다. 이는 면담 중심의 평가에서도 기존의 K-MBI와 K-MBI 체크리스트 모두 총점수준에서 높은 평가자간 신뢰도를 보인다는 것을 의미한다. 카파검정 기준으로(Fleiss, 2003) 두 평가 도구 모두 탁월한 신뢰도 항목이 넷, 상당한 신뢰도 항목이 여섯으로 같은 결과를 보였으나, Mean K값은 K-MBI 체크리스트가 .725로 기존 K-MBI .699에 비해 다소 높은 값을 보였다. 하지만 이러한 결과는 Leung 등(2007)의 중국판 MBI개발을 위한 연구보다 낮은 결과로서 Leung 등(2007)의 연구에선 10개 항목 모두 Kappa value가 .75 이상으로 탁월한 신뢰도를 보이고 있다. 이것은 본 연구가 면담 위주의 평가가 진행된 것에 반해, Leung 등(2007)의 연구에서는 한 명의 주 평가자가 환자의 실제 수행을 관찰하여 평가를 진행하면 나머지 3명의 평가자 역시 그 과정을 관찰하여 평가를 하였기 때문으로 판단된다.

다른 나라의 경우 관찰 중심의 평가가 가지는 시간적, 비용적 문제를 해결하기 위해 여러 MBI 평가 방법을 고안하여 연구하고 있다. Gompertz 등(1994)은 우편을 통해서 설문지에 답하는 형식의 MBI를 개발하였고, 연구 결과 기능장애가 심하지 않은 환자의 경우에 한해서 높은 신뢰도를 확보할 수 있다고 보고하였다. 또한 Shinar 등(1987)은 관찰과 전화면담 형식을 통해 MBI 평가를 하였을 때 두 형식 간 높은 상관성을 보고하였고, 많은 대상자를 평가하여야 하는 대규모 연구에서는 전화면담과 같은 형식의 평가가 시간과 비용 측면에서 효율적임을 제안하였다. 이와 더불어 Heuschmann 등(2005)은 독일판 Barthel Index 신뢰도 연구와 우편 및 전화 형식의 Barthel Index 개발연구에서, 직접관찰을 통한 평가 시 Mean K값이 .93, 우편을 통한 설문 형식은 .79, 전화면담을 통한 평가에선 .80으로 보고하였다. 이러한 결과는 3가지 평가 방법 모두 탁월한 신뢰도를 확보함으로써, 3가지 평가 모두 적절한 평가 방법임을 보여준다. 특히 직접 관찰을 통한 방법이 MBI평가에 있어 가장 신뢰도를 높일 수

있는 방법임을 보여주는 동시에 설문 형식이나 전화형식에 맞춰 수정한 Barthel Index 역시 높은 신뢰도를 확보하고 있음을 보여주는 결과이다.

본 연구의 세부 항목별 일치도를 살펴보면 K-MBI 체크리스트에서 가장 낮은 일치도를 보이는 항목은 옷입고 벗기(k values=.565)인 것으로 나타났다. 옷입고 벗기는 계절별 옷의 종류와 상, 하의 등 많은 요소를 고려하여 평가를 진행하여야 한다. Fricke와 Unsworth(1997)의 연구에서도 MBI의 옷입고 벗기 항목은 낮은 일치도(k values=.520)를 보이고 있다. 이러한 결과는 상, 하의를 구분하여 평가하는 FIM에서와는 달리 MBI의 경우 상, 하의 구분 없이 평가를 진행하여, 평가의 민감도가 떨어지는 것으로 보고하고 있다(Fricke & Unsworth, 1997). 반면 본 연구의 K-MBI 그룹의 경우 옷입고 벗기 항목에서 높은 일치도(k values=.712)를 보이고 있다. 이는 K-MBI의 경우 전반적인 수행 정도에 따른 도움의 정도를 근거로 평가를 진행 한 반면, K-MBI 체크리스트의 경우 모든 동작들이 세분화 된 체크리스트를 기준으로 평가하는 과정에서 평가자들이 서로 다른 등급을 주는 것으로 확인되었다. 결국 많은 부분을 고려해야 하는 K-MBI 체크리스트에 비해 전반적인 도움의 정도를 기준으로 하는 K-MBI 평가가 더 높은 평가자간 일치도를 보인 것이다. 이것은 K-MBI 체크리스트가 오히려 항목간 일치도를 떨어트릴 수 있음을 보여주는 것으로서 향후 K-MBI 체크리스트 수정 과정에서는 모든 동작을 고려하되 평가자들의 혼란스럽지 않도록 하는 방법에 대해서 논의 되어야 할 것이다.

신체적 장애를 가진 환자의 일상생활을 객관적이고 정확하게 평가 측정하고, 그들에게 얼마나 많은 서비스를 어떻게 제공할 것인가를 결정하는 것은 매우 중요한 일이다(Loewen & Anderson, 1990; Wade & Hewer, 1987). 더욱이 MBI는 환자들의 장애등급을 판정하는 기준이 되는 평가로서 임상에서 평가의 정확성과 신뢰성이 다른 어떤 평가보다도 중요한 평가도구이다. 본 연구결과에서 알 수 있듯이 총점수준에서 기존 K-MBI와 K-MBI 체크리스트 모두 높은 수준의 평가자간 신뢰도를 보이고 있음을 알 수 있다. 반면에 세부 항목 간 일치도에서는 K-MBI 체크리스트 사용이 좀 더 높은 일치도를 보이고 있음을 확인할 수 있었다. 이러한 결과는 일상생활 동작 수행 평가 시 새로 개발된 K-MBI 체크리스트가 평가자간 신뢰도를 높일 수 있는 방법임을 보여준다.

그러나 본 연구는 국내 종합병원 한 곳에서 이루어졌으며, 대상자를 뇌병변 환자로 제한하여 모든 환자에게 일반화하기에는 어려움이 있다. 추후 연구에서는 이러한 점을 보완하여 보다 많은 평가자와 환자 군을 대상으로 하는 연구가 진행되어야 할 것으로 판단된다.

V. 결 론

본 연구의 결과를 통해 총점수준에서 두 평가 모두 높은 신뢰도를 보이는 것을 알 수 있었으며, 일부 항목에서는 K-MBI 체크리스트가 평가자간 신뢰도를 높이는 데 도움이 될 수 있음을 알 수 있었다. 본 연구의 의의는 현재 임상에서 사용되는 면담 중심의 K-MBI 평가에 대한 문제점을 논의하여 K-MBI 체크리스트의 개발 및 평가자간 신뢰도와 세부항목간의 일치도를 구한 것에 있다. 하지만 몇몇 항목에서는 낮은 일치도를 보이고 있어, 앞으로의 연구를 통해 수정, 보완해야 할 것으로 판단한다.

참 고 문 헌

- 남궁영, 이한석, 장기연. (1996). 뇌졸중 환자에서의 Functional Independence Measure(FIM)와 5점 척도 Modified Barthel Index(MBI)의 상관성에 관한 연구. 대한작업치료학회지, 4(1), 45-49.
- 박소연, 유은영. (2002). 한국 작업치료사의 평가도구 사용에 관한 연구. 대한작업치료학회지, 10(2), 99-108.
- 보건복지가족부. (2011). 장애등급판정기준. 서울: 보건복지가족부.
- 정한영, 박병규, 신희석, 강윤규, 편성범, 백남중 등. (2007). 한글판 수정바텔지수(K-MBI)의 개발: 뇌졸중 환자 대상의 다기관 연구. 대한재활의학회지, 31(3), 283-297.
- 조남중. (2008). 조선업의 근골격계부담 평가를 위한 Quick Checklist 개발. 석사학위논문, 부경대학교, 부산.
- 주혜선. (2008). 외상 후 위기 체크리스트 개발 연구. 석사학위논문, 이화여자대학교, 서울.
- Boorman, D. J. (2001). Today's electronic checklists

- reduce likelihood of crew errors and help prevent mishaps. *International Civil Aviation Organization Journal*, 56(1), 17–22.
- Delisa, J. A., & Gans, B. M. (2004). Clinical Evaluation. In B. K. Ganter, R. P. Erickson, M. A., & Butters (Eds.). *Physical medicine, principle and practice* (4th ed., pp. 1–48). Philadelphia: Lippincott Williams & Wilkins.
- Donaldson, S. W., Wagner, C. C., & Gresham, G. E. (1973). A unified activity of daily living evaluation form. *Medical Rehabilitation*, 54(4), 175–179.
- Fleiss, J. L., Levin, B., & Paik, M. C. (2003) The measurement of interrater agreement. *Statistical methods for rates and proportions* (3rd ed., pp. 598–626). Hoboken: John Wiley & Sons.
- Fricke, J., & Unsworth, C. A. (1997). Inter-rater reliability of the original and Modified Barthel Index, and a comparison with the Functional Independence Measure. *Australian Journal of Physiotherapy*, 44(1), 22–29.
- Gompertz, P., Pound, P., & Ebrahim, S. (1994). A postal version of the Barthel Index. *Rehabilitation*, 8(30), 233–239.
- Helmreich, R. L. (2000). On error management: Lessons from aviation. *British Medical Journal*, 320(7237), 781–785.
- Heuschmann, P. U., Kolominsky-Rabas, P. L., Nolte, C. H., Hunermund, G., Ruf, H. U., Laumeier, I., et al. (2005). The reliability of the German version of the Barthel-Index and the development of a postal and telephone version for the application on stroke patients. *Fortschritte der Neurologie-Psychrie*, 73(2), 72–82.
- Hocking, C., Williams, M., Broad, J., & Baskett, J. (1999). Sensitivity of Shah, Vanclay and Cooper's Modified Barthel Index. *Clinical Rehabilitation*, 13(2), 141–147.
- Korner-Bitensky, N., & Wood-Dauphine, S. (1995). Barthel Index information elicited over the telephone. Is it reliable? *American Journal of Physical Medicine*, 74(1), 9–18.
- Kucukdeveci, A. A., Yavuzer, G., Tennant, A., Suldur, N., Sonel, B., & Arasil, T. (2000). Adaptation of the Modified Barthel Index for use in physical medicine and rehabilitation in Turkey. *Scandinavian Journal of Rehabilitation Medicine*, 32(2), 87–92.
- Leung, S. O., Chan, C. C., & Shah, S. (2007). Development of a Chinese version of the Modified Barthel Index validity and reliability. *Clinical Rehabilitation*, 21(10), 912–922.
- Loewen, S. C., & Anderson, B. A. (1990). Predictors of stroke outcome using objective measurement scales. *Stroke*, 21(1), 78–81.
- Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal*, 14, 61–65.
- Middleton, J. W., Truman, G., & Geraghty, T. J. (1998). Neurological level effect on the discharge functional status of spinal cord injured persons after rehabilitation. *Archives of Physical Medicine and Rehabilitation*, 79(11), 1429–1432.
- Pronovost, P., Needham, D., Berenholtz, S., Sinopoli, D., Chu, H., Cosgrove, S., et al. (2006). An intervention to decrease catheter-related bloodstream infections in the ICU. *New England Journal of Medicine*, 355(26), 2725–2732.
- Roden-Jullig, A., Britton, M., Gustafsson, C., & Fugl-Meyer, A. (1994). Validation of four scales for the acute stage of stroke. *Journal of Internal Medicine*, 236(2), 125–136.
- Shah, S., & Muncer, S. (2000). Sensitivity of shah, vanclay and cooper's Modified Barthel Index. *Clinical Rehabilitation*, 14(5), 551–552.
- Shah, S., Vanclay, F., & Cooper, B. (1989). Improving the sensitivity of barthel Index for stroke rehabilitation. *Journal of Clinical Epidemiology*, 42(8), 703–709.
- Shinar, D., Gross, C. R., Bronstein, K. S., Licata-Gehr, E. E., Cabrera, A. R., Fishman, I. G., et al. (1987). Reliability of the activities of daily living scale and

- its use in telephone interview. *Archives of Physical Medicine and Rehabilitation*, 68(10), 723–728.
- Wade, D. T., & Hower, R. L. (1987). Functional abilities after stroke: Measurement, natural history and prognosis. *Journal of Neurology, Neurosurgery and Psychiatry*, 50(2), 177–182.
- Wolff, A. M., Taylor, S. A., & McCabe, J. F. (2004). Using checklists and reminders in clinical pathways to improve hospital inpatient care. *The Medical Journal of Australia*, 181(8), 428–431

Abstract

An Inter Rater Reliability Study of K-MBI Checklist

Lee, Keun-Ho^{*}, M.S., O.T., Koo, Kyung-Hyup^{*}, M.S., O.T.,
Shin, Seon-Hee^{*}, B.H.Sc., O.T., Ahn, A-Ra^{*}, B.H.Sc., O.T.,
Son, Mi-Nam^{*}, B.H.Sc., O.T., Cho, Eun-Hee^{*}, M.S., O.T.,
Kim, Yun-Hee^{**}, Ph.D., M.D.

^{*}Dept. of Occupational Therapy, Samsung Medical Center

^{**}Dept. of Physical Medicine and Rehabilitation, Samsung Medical Center,
Sungkyunkwan University School of Medicine

Objective : The purpose of this study was to devise a checklist version of Korean version of Modified Barthel Index (K-MBI) and compare its reliability with the K-MBI.

Methods : The subjects were 30 brain injury patients. We subdivided each subitem of the original K-MBI and created a checklist (so-called K-MBI checklist). We divided it into 2 groups: a newly developed K-MBI checklist, and the original K-MBI, with two randomly assigned occupational therapists for each group. The collected data were analyzed to ascertain the inter-rater reliability through an intraclass correlation coefficient (ICC) for the total test score and Kappa statistics for the subitems of the test.

Results : The inter-rater reliability of the total test score was significantly high in both the K-MBI checklist group (ICC=.981) and the original K-MBI group (ICC= .973) ($p<.001$). For the internal constancy in the subitems of the test, both groups showed excellent reliability in four items. The mean Kappa value was higher in the K-MBI checklist group (.725) than in the original K-MBI group (.699).

Conclusions : Both the newly developed K-MBI checklist and original K-MBI showed high consistency in their total scores. In the subitems, the newly developed K-MBI checklist provided higher consistency than the original K-MBI except for dressing subitems. Meanwhile, the K-MBI showed low consistency in personal hygiene and ambulation subitems.

Key words : Activities of daily living, Checklist, Inter-rater reliability, K-MBI