

한국판 CBCL 1.5-5의 변별타당도 및 임상적 유용성

이진김영아 오경자[†]

(주) 휴노컨설팅

연세대학교 심리학과

본 연구는 근래에 한국판으로 개발된 영유아 행동평가척도 부모용(CBCL 1.5-5)의 변별타당도를 검증하고 임상 장면에서의 유용성을 평가하기 위해 수행되었다. 이를 위해 종합병원 소아정신과에 내원한 1.5~5세 영유아 228명의 임상집단과 이 집단의 성별·연령과 동일하게 대응시킨 228명의 정상집단을 규준산출을 위해 수집된 2,433명의 모집단에서 무선적으로 선별하여 분석을 실시하였다. ANOVA 분석 결과 한국판 CBCL 1.5-5의 모든 하위척도 평균 점수는 임상집단이 정상집단에 비해 높게 나타났다. 판별분석과 ROC 분석에서도 신체증상 및 수면문제 소척도를 제외한 모든 하위척도를 통한 임상집단과 정상집단 변별이 효과적으로 이루어질 수 있음이 확인되었고, 특히 DSM 전반적 발달장애 및 주의집중 하위척도의 변별력이 우수한 것으로 나타났다. 승산비(odds ratio) 분석을 통해 CBCL 1.5-5에서 임상범위의 판단기준으로 제시하고 있는 T점수로 임상집단과 정상집단을 예측해본 결과, 역시 신체증상 및 수면문제 척도를 제외한 모든 하위척도의 승산비가 유의하였고, 하나의 하위척도라도 임상범위 점수에 해당하는 경우를 임상집단으로 분류하는 기준으로 평가했을 때 승산비가 더욱 높아져 각 하위척도들의 임상적 변별력이 지지되었다. 마지막으로 본 연구에서 확인된 한국판 CBCL 1.5-5의 변별 타당도 및 임상적 유용성과 그에 대한 한계점이 논의되었다.

주요어 : CBCL 1.5-5, 표준화, 임상군, 정상군, 변별타당도, 승산비, ROC curve 분석, 판별분석

[†] 교신저자(Corresponding Author) : 오경자 / 연세대학교 심리학과 / 서울시 서대문구 신촌동 134
E-mail : kjoh@yonsei.ac.kr

CBCL 1.5-5는 Achenbach와 Rescorla(2000)에 의해 개발된 영유아용 문제행동 평가 도구로, 부모의 보고를 통해 영유아의 불안/우울, 위축 등의 내재화 문제와 주의집중 문제, 공격적 행동 등의 외현화 문제 수준을 측정하는 객관적 평가 도구이다. 이는 영유아부터 노인에 이르는 전 연령대를 대상으로 한 정신건강 척도인 ASEBA(Achenbach System of Empirically Based Assessment, Achenbach & Rescorla, 2000, 2001)의 검사 가운데 하나이며, CBCL 2-3 (Achenbach, Edelbrock, & Howell, 1987)의 개정판이다. 국내에서는 근래 영유아 문제 행동에 대한 효과적인 평가와 이해에 대한 관심이 제고됨에 따라, 예비 연구(김영아, 오경자, 문수종, 김유진, 2005; Kim, Oh, & Lee, 2007)를 거쳐 전국의 영유아 집단을 대상으로 한국판 CBCL 1.5-5가 표준화 되었다(오경자, 김영아, 2008).

CBCL 1.5-5는 정신보건, 의료, 교육 현장 등 다양한 상황에서 사용될 수 있는데 특히 영유아가 보일 수 있는 다양한 문제행동의 수준을 평가함으로써 전반적인 발달 및 심리사회적 적응 수준 등에 대한 개괄적 이해를 제공하는 것을 목표로 하고 있다. 특히 영유아를 대상으로 한 문제행동 척도의 임상적 타당도는 척도의 실시를 통해 영유아 문제 행동에 대한 조기 개입을 가능케 하고, 해당 영유아의 발달 경로에 대한 예측을 가능케 한다는 점에서 중요한 의미를 지닐 수 있다(Anastasiow, 1996; 이미선, 조광순, 김주영, 강병호, 2001; 황해익, 정혜영, 1996). 실제로 영유아의 문제 행동은 연령이 증가하면서 계속 유지가 되는 경향이 있고(Hwang & St. James-Roberts, 1998; Patterson, Capaldi, & Bank, 1989) 이후 학업, 사회적 적응 등의 다양한 문제로 이어질 가능성이 있기 때

문에 영유아 시기의 문제 행동에 대한 적절한 평가를 통한 조기 개입이 효과가 있음을 보인 연구결과들이 국내외에서 보고된 바 있다(Hester, Baltodano, Gable, Tonelson, & Hendrickson, 2003; Qi & Kaiser, 2003; 이경숙, 신의진, 신동주, 전연진, 박진아, 2004).

CBCL 1.5-5이 영유아 문제행동에 대한 사전 판별 도구로서 활용되기 위해서는, CBCL 1.5-5의 하위 척도점수 및 총점을 통해 전문적인 임상적 개입이 필요한 영유아를 그렇지 않은 영유아 집단으로부터 효과적으로 구별해 낼 수 있음이 경험적으로 뒷받침되어야 할 것이다. 검사의 원저자인 Achenbach와 Rescorla(2000)는 CBCL 1.5-5의 하위척도가 보호자에 의해 문제 유무에 대한 평가 또는 치료가 필요하다고 여겨져 소아 정신과 등의 임상기관에 내원한 적이 있는 1.5-5세 영유아 집단(referred sample)과 임상기관에 내원한 적이 없는 집단(non-referred sample)을 효과적으로 구별해 내는데 유용한지를 다양한 방식으로 검증한 바 있다. 먼저 임상집단에 해당하는 집단을 구분하기 위한 최적의 척도 점수를 탐색하기 위해 수신기 작동 특성 분석(receiver operating characteristic; ROC 분석)을 실시한 결과, 내재, 외현, 문제행동 총점의 경우 백분위 80~84%에 해당하는 점수가 임상기관 내원 집단과 그렇지 않은 집단을 가장 효과적으로 구별해 낼 수 있는 것으로 나타났다. 이러한 결과 및 임상 장면에서의 활용 측면을 고려하여, 이들 하위척도에 대해서는 T점수 64점 이상을 임상 범위로 판단할 수 있는 기준으로, 백분위 83~90%에 해당하는 T점수 60~63점이 경계선 범위의 임상집단 구분 점수로 보면 효과적일 것으로 제안하였다. 그리고 나머지 하위척도들에 대해서는 문항수가 적은 것을 감안하여

좀 더 엄격한 기준인 T점수 65~69를 경계선 임상범위, T점수 70이상을 임상범위의 기준으로 제시하였다. 이후 승산비(Odds ratio) 분석을 통해 임상집단 판별 점수 기준이 임상기관에 의뢰된 적이 있는 집단과 그렇지 않은 집단을 실제로 얼마나 정확히 구별해 내는지 평가하였다. 그 결과 CBCL 1.5-5의 모든 하위척도는 통계적으로 유의미한 수준에서($p < .01$) 두 집단을 효과적으로 구별해낼 수 있는 것으로 나타났다. 특히, CBCL 1.5-5의 DSM 하위 척도 중 하나인 DSM전반적 발달장애 하위척도의 경우 승산비가 11 정도로 상당히 높았는데, 이는 DSM전반적 발달장애 척도에서 임상범위에 해당하는 점수를 받을 경우, 그렇지 않은 집단에 비해 실제로 임상기관에 의뢰되어 문제 유무에 대한 진단 또는 치료를 받았을 가능성이 11배에 달함을 알려준다. 또한 판별분석을 통해 각 하위척도 점수들로 임상집단 내원 집단을 얼마나 정확하게 구별할 수 있는지를 확인해본 결과, 74~78%의 수준으로 두 집단을 정확하게 분류할 수 있는 것으로 나타났다. 이와 같은 연구 결과는 CBCL 1.5-5를 통해 임상적 개입이 필요한 영유아를 효과적으로 구별해 낼 수 있음을 알려준다.

CBCL의 다양한 검사군은 53개 언어로 번역되어 다양한 문화권에서 유용성이 검증되어왔고 여러 문화권에서 아동 문제 행동 수준에 대한 비교 연구가 실시되어오고 있는데 (Crijnen, Achenbach, & Verhulst, 1997, 1999), 아동 청소년 연령대인 CBCL 6-18에 대한 변별타당도를 검증하는 연구들은 국내외에서 진행된 바 있으나(Biederman, Ball, Monuteaux, Kaiser, & Faraone, 2007; Hudziak et al., 2006; Aschenbrand, Angelosante, & Kendall, 2005; 이수진, 하은혜, 오경자, 2008), CBCL 1.5-5의 변별

타당도를 검증하는 연구는 아직 초기단계이다. 영유아를 대상으로 한 CBCL 1.5-5는 한국에서는 처음으로 표준화가 된 검사이므로 검사의 변별 타당도를 검증하고, 적절한 임상적 활용을 위한 제안이 필요할 것이다.

본 연구에서는 실제로 소아 정신과에 내원한 적이 있는 임상집단과 그렇지 않은 정상집단 영유아를 대상으로 한국판 CBCL 1.5-5의 각 하위척도들을 통한 임상군·정상군 분류 정확성을 평가하고자 하였다. 이를 위해 우선 한국판 CBCL 1.5-5의 하위척도 점수를 통해 임상군·정상군 문제행동 수준을 전반적으로 비교 평가하고, 각 하위척도들의 변별력을 검증하고자 하였다. 또한 임상적으로 의미가 있다고 판단을 내릴 때 사용할 수 있는 임상 T점수 기준은 미국과 같은 수준으로 사용해도 적절한지, 이러한 기준선을 사용한 임상군 예측이 실제로 임상집단 여부를 적절히 구별해낼 수 있는지를 검토하고자 하였다. 또한 한국판 CBCL 1.5-5의 개별 하위척도들에 대한 임상집단 변별력을 탐색함으로써 하위척도별 변별 타당도 수준 및 임상적 활용에 있어서의 시사점을 검토해 보고자 하였다.

연구 방법

연구 대상

본 연구의 연구 대상은 크게 임상집단과 그에 대응되는 정상집단 영유아로 구성되어있다. 임상 및 정상집단에 대한 상세한 설명은 아래에 제시하였으며 연령, 및 성별에 따른 분포는 표 1과 같다.

표 1. 임상군 및 정상집단의 연령, 성별 분포(명)

	임상군			정상군			총계
	남	여	계	남	여	계	
1	8	4	12	8	4	12	24
2	19	19	38	19	19	38	76
3	44	19	63	44	19	63	126
4	52	17	69	52	17	69	138
5	30	16	46	30	16	46	92
계	153	75	228	153	75	228	456

임상집단

서울 소재 종합병원 소아 정신과에 2007년 1월부터 2008년 10월 까지의 기간 동안 의뢰된 만 18개월~5세 유아사례 중 정신과 전문의와 임상심리전문의의 의학적 면담 및 종합심리평가에 의거하여 DSM-IV기반 진단 소견이 있는 유아 228명(남아: 153명, 여아: 75명)을 무선으로 표집하여 분석에 포함하였다. 주요 호소 문제는 발달지연, 전반적 발달장애, 정신지체, 영아 연축, 불안장애, 주의력 문제, 언어 문제 등 일반적으로 소아 정신과에서 내원하는 환아들에게서 발견되는 문제들로 구성되어 있었다. 임상집단의 평균 연령은 3.43세(표준편차 1.14)였다.

정상집단

한국판 CBCL 1.5-5의 표준화를 위해 대한민국 주택 총조사(2005)에 근거해 지역 및 성별을 맞추어 수집한 18개월부터 5세 사이의 영유아 2,433명의 자료 중 228명의 임상집단과 연령 및 성별을 대응시킨 동수의 아동을 무선적으로 선별하였다. 이들 집단은 소아 정신과 및 상담기관 등의 임상 기관에 의뢰된 적이

없었다. 정상집단의 평균 연령은 임상집단과 마찬가지로 3.43세(표준편차 1.14)였다. 임상집단에 대한 대응 자료로 무선적으로 표집된 228명의 자료가 전체 모집단(N=2,433)을 잘 반영하고 있는지 확인하기 위하여 규준집단별로(1.5~2세, 3~5세) 각 하위척도 점수를 *F* 검증을 통해 비교한 결과, 모든 하위척도에 대해서 통계적으로 유의미한 차이를 나타내지 않았다. 따라서 이번 연구에 사용된 정상군 228명의 자료는 그 자료가 추출된 모집단 2,433명의 한국판 CBCL 1.5-5 하위척도 점수 수준을 잘 반영하고 있는 것으로 볼 수 있을 것이다.

연구 도구

한국판 유아행동평가 척도-부모용(Child Behavior Checklist for ages 1.5-5; CBCL 1.5-5)

Achenbach(1992)가 처음 제작하고 이후 개정 작업을 거친 CBCL 1.5-5(Achenbach & Rescrola, 2000)를 오경자, 김영아(2008)가 한국판으로 표준화한 것을 사용하였다. 영유아들이 보일 수 있는 주요한 증상군에 대한 8가지 하위척도(정서적 반응성, 불안/우울, 신체증상, 위축, 주의집중문제, 공격행동, 수면문제의 7개의 증후군 척도와 단일 증후군으로 묶이지 않은 문항들을 포함하는 기타 척도)와 이들 하위척도들의 합으로 이루어진 내재화 척도(정서적 반응성, 불안/우울, 신체증상, 위축 하위 요인의 합), 외현화 척도(주의집중, 공격행동 하위 요인의 합), 문제행동 총점(모든 8개 하위 요인의 합)을 분석에 이용하였다. 그리고 DSM 진단체계를 반영하여 구성한 5개의 하위척도인 DSM 정서문제, DSM불안문제, DSM전반적발달장애, DSM ADHD, DSM반항행동문제 척도도

분석에 포함되었다.

분석방법

먼저 임상집단과 정상집단에 따른 한국판 CBCL 1.5-5의 하위척도 평균 점수 수준을 파악하고 두 집단 간의 차이가 나타나는지를 확인하기 위해 ANOVA 분석을 통해 차이 검증을 실시하였다. 그리고 한국판 CBCL 1.5-5의 전체 하위척도 및 개별 하위척도 점수를 이용하여 임상집단과 정상집단의 분류를 어느 정도 정확하게 해 낼 수 있는지 검증하기 위한 목적으로 판별분석을 실시하였으며, 개별 하위척도 점수의 변별력 수준과 점수의 구체적인 임상적 활용에 대한 세부적인 이해를 얻기 위해 ROC(Receiver Operation Characteristic Curve) 분석을 시행하였다. ROC 분석에서는 민감도와 특이도의 수준을 근거로, 척도 점수를 통한 가장 적합한 수준의 집단 분류 기준을 제시해 준다. 이와 같은 방식으로 척도들의 변별력을 확인한 후, 척도의 점수를 해석하고 활용하는데 지침이 될 수 있는 준거점들을 중심으로 승산비(odds ratio) 분석을 실시하였다. 승산비 분석에서는 집단을 구분하는 하나의 이분변수 기준(이를테면 검사 점수 60점 이상, 이하 구분), 그리고 동일한 목적으로 집단을 구분하는 또 다른 하나의 이분변수 기준(이를테면 병원에서 진단명 유, 무 구분)을 통해, 전체 집단을 2x2 로 구분하고, 이중 비교하고자 하는 기준이 연구자가 가정했던 대로 다른 준거 기준과 동일하게 집단을 구분하는 가능성이 그렇지 않을 가능성에 비해 얼마나 높고 낮은 수준인지를 검증하게 된다. 승산비는 다음과 같이 계산 된다: 승산비(odds ratio) = (true positive × true negative)/(false positive × false

negative). 승산비 분석에서는 T점수를 활용하였고, 특히 총점 내외현화 척도에서는 미국판 CBCL 1.5-5에서 임상적으로 의미가 있다고 판단하는 기준점으로 사용되는 60, 65 T를 중심으로 살펴보았다.

결 과

한국판 CBCL1.5-5의 하위척도 점수에 대한 임상·정상 집단별 차이 검증

ANOVA 분석을 통해 한국판 CBCL 1.5-5의 각 하위척도에서 임상·정상 집단에 따라 차이가 있는지를 확인해본 결과, 신체증상 및 수면문제를 제외한 모든 하위척도의 평균은 임상집단이 정상집단보다 통계적으로 유의하게 높게 나타났다(표 2). 신체증상과 수면문제 척도에서도 임상집단의 평균점수가 다소 높게 나타났으나 통계적으로 유의한 차이는 관찰되지 않았다.

판별분석

한국판 CBCL 1.5-5의 각 하위척도들과 전체 하위척도를 이용하여 임상집단과 정상집단을 어느 정도로 정확하게 분류할 수 있는지를 파악하기 위해 판별분석을 실시하였다. 표 3은 한국판 CBCL 1.5-5의 모든 개별적인 하위척도들 각각을 통해 정상군과 임상군을 판별해낼 수 있는지를 조사한 결과이다. 신체증상 및 수면문제를 제외한 모든 다른 하위척도들이 $p < .01$ 수준에서 유의미하게 임상집단과 정상집단을 분류해 낼 수 있는 것으로 나타났다. 이후, 한국판 CBCL 1.5-5의 모든 하위척도를 예

표 2. 한국판 CBCL1.5-5의 하위척도 점수에 대한 임상·정상 집단별 차이 검증(ANOVA)

	임상군(n=228)		정상군(n=228)		전체(N=456)		임상·정상 차이검증 F
	평균	표준편차	평균	표준편차	평균	표준편차	
정서적 반응성	3.90	3.05	2.54	2.58	3.22	2.90	26.64**
불안/우울	4.08	2.89	2.95	2.26	3.52	2.66	21.63**
신체증상	1.82	1.92	1.68	1.87	1.75	1.89	.59
위축	5.04	3.21	1.75	1.92	3.39	3.11	176.37**
수면문제	2.90	2.40	2.64	2.44	2.77	2.42	1.30
주의집중문제	4.45	2.42	1.68	1.74	3.06	2.52	197.03**
공격행동	10.17	6.03	7.16	5.59	8.66	6.00	30.53**
기타문제	14.43	6.84	8.14	5.44	11.29	6.93	118.27**
내재화	14.84	8.18	8.92	7.18	11.88	8.24	67.41**
외현화	14.61	7.40	8.84	6.78	11.73	7.65	75.59**
총점	46.79	21.01	28.54	19.35	37.67	22.15	93.10**
DSM 정서문제	3.93	2.95	2.47	2.25	3.20	2.72	35.09**
DSM 불안문제	4.82	3.16	3.73	2.71	4.27	2.99	15.70**
DSM 전반적 발달장애	8.04	4.12	3.04	2.96	5.54	4.37	220.94**
DSM ADHD	4.79	2.53	3.04	2.48	3.92	2.66	55.44**
DSM 반항행동문제	3.39	2.33	2.15	1.95	2.77	2.23	37.48**

주. ** $p < .01$

측변인으로 넣고, 임상·정상 집단 여부를 분류하도록 하였을 경우, 전체 사례의 82.89%에 대하여 정확한 분류가 가능한 것으로 나타났다(표 4).

또한 총 16개의 하위척도 가운데 어떤 소척도가 임상·정상 집단 여부를 가장 잘 판별할 수 있는지를 분석하기 위해 단계적 판별분석을 실시한 결과, 판별 정확도는 DSM 전반적 발달장애, 주의집중문제, DSM ADHD, 그리고 내재화 하위척도 순으로 나타났다(표 5). 이때 정준상관계수(canonical correlation)는 .67이었고 이 4개 하위척도 요인으로 구성된 판별함수의

설명력(Eigenvalue)은 79.37%로 나타났다.

ROC 분석

한국판 CBCL 1.5-5의 개별 하위척도들이 대부분 임상집단 변별력을 보이고, 개별 하위척도에 따라 변별력 수준의 차이가 있는 것으로 나타났다. 다음으로 ROC 분석을 통해 한국판 CBCL 1.5-5의 개별 하위척도에 대한 정상집단 및 임상집단 판별력을 확인하고자 하였다. ROC 분석에서는 민감도와 특이도를 통해 도출된 곡선의 아래 부분 면적을 의미하는 AUC

표 3. 정상·임상집단 여부를 종속변인으로 한 하위 척도 별 판별분석 결과

	Wilks' λ	Chi-Square	F
정서적 반응성	.95	25.86	26.64**
불안/우울	.96	21.11	21.63**
신체증상	1.00	.59	.59
위축	.72	148.84	176.37**
수면문제	1.00	1.30	1.30
주의집중문제	.70	163.47	197.03**
공격행동	.94	29.52	30.53**
기타문제	.79	104.93	118.27**
내재화	.87	62.78	67.41**
외현화	.86	69.84	75.59**
총점	.83	84.60	93.10**
DSM 정서문제	.93	33.76	35.09**
DSM 불안문제	.97	15.42	15.70**
DSM 전반적 발달장애	.67	179.82	220.93**
DSM ADHD	.89	52.25	55.44**
DSM 반항행동문제	.92	35.97	37.48**

주. ** $p < .01$

표 4. 한국판 CBCL 1.5-5의 모든 하위척도를 예측변인으로 사용하여 정상·임상집단 여부를 분류한 결과

실제집단	사례수	예측된 집단	
		정상	임상
정상	228	204 (89.47)	24 (10.53)
임상	228	54 (23.68)	174 (76.32)

주. () 안의 값은 %. 집단을 정확하게 분류한 비율은 82.89%임.

표 5. 한국판 CBCL 1.5-5의 모든 하위척도를 예측변인으로 사용한 단계적 판별분석의 결과

	Wilks' λ	F
1 DSM 전반적발달장애	.67**	220.93
2 주의집중문제	.60**	150.91
3 DSM ADHD	.57**	114.94
4 내재화	.56**	89.49

주. ** $p < .01$

(area under the curve) 값을 통해 판별의 적절성을 확인할 수 있다. Holmes(1998)나 Luoto와 Hjort(2005)의 연구에서 적용되었던 기준에 의하면 판별의 정확도 수준은 AUC가 .90~1.00일 경우 매우 우수, .80~.90일 경우 우수, .70~.80일 경우 적합, .60~.70일 경우에는 변별력은 지지될 수는 있지만 그다지 우수한 수준은 아니고, .60~.50일 때는 변별력이 의미가 없는 수준으로 볼 수 있다.

한국판 CBCL 1.5-5의 하위척도들에 대한 ROC 분석 결과 및 AUC 값은 표 6에 제시되어 있다. 한국판 CBCL 1.5-5의 하위 척도 가운데 가장 우수한 수준의 AUC를 보인 척도는 DSM 전반적 발달장애 척도였고, 그 값은 .84($p < .01$)였다. AUC가 .80~.90에 해당하는 하위척도는 DSM 전반적 발달장애, 주의집중문제, 위축 척도였고 이들 척도의 판별 정확도는 앞서 언급한 기준에 의거, 우수한 수준인 것으로 볼 수 있다. 기타문제, 총점, 내재화, 외현화 소척도의 AUC는 각각 .77, .75, .72, .72 수준으로 정확도가 적합 수준인 것으로 나타났다.

그 외에 나머지 척도들은 신체증상과 수면문제를 제외하고는 .60~.69 범위의 적합도 수준을 보여 우수한 수준은 아니었지만 변별력

표 6. 한국판 CBCL 1.5-5의 하위척도에 대한 ROC 분석 결과

	AUC	표준오차	95% 신뢰 구간	
			Lower Bound	Upper Bound
정서적 반응성	.64**	.03	.59	.69
불안/우울	.61**	.03	.56	.66
신체증상	.52	.03	.47	.58
위축	.82**	.02	.78	.86
수면문제	.54	.03	.48	.59
주의집중문제	.82**	.02	.78	.86
공격행동	.65**	.03	.60	.70
기타문제	.77**	.02	.72	.81
내재화	.72**	.02	.68	.77
외현화	.72**	.02	.67	.77
총점	.75**	.02	.70	.79
DSM 정서문제	.65**	.03	.60	.70
DSM 불안문제	.60**	.03	.55	.65
DSM 전반적 발달장애	.84**	.02	.81	.88
DSM ADHD	.69**	.02	.65	.74
DSM 반항행동문제	.66**	.03	.61	.71

주. ** $p < .01$

에 있어서의 통계적인 유의미성이 지지되고 있음을 확인할 수 있었다. 신체증상과 수면문제 척도의 경우 ROC분석에서도 판별의 적절성이 지지되지 않았다.

승산비 분석

판별분석과 ROC분석을 통해 한국판 CBCL 1.5-5의 하위척도들의 임상집단 변별력을 확인하였다. 변별력이 있는 척도들을 잘 활용하기 위해서는 각 점수대에 대한 해석도 중요한데, 특히 몇 점 이상의 점수를 받게 될 경우 임상

적으로 의미 있게 평가할 수 있을 만한 높은 점수로 볼 것인가에 대한 판단이 필요하게 된다. 미국판 CBCL 1.5-5의 경우에는 총점 및 내재·외현화 척도의 T점수 64(91.92%ile)일 때 임상범위, T점수 60(84.13%ile)를 정상집단과의 경계선 임상범위로 보고, 각 하위척도에 대해서는 보다 엄격한 기준을 적용하여 70T 이상을 임상범위, 65T를 경계선 임상 범위의 기준으로 제시하고 있다. 어느 점수대 이상을 임상범위로 규정하는 것이 효과적인가는 척도의 활용목적에 따라 달라지게 된다. 임상집단을 정상집단으로 잘못 평가할 위험성과 정상집단

을 임상집단으로 잘못 평가할 위험성 중 어디에 비중을 두는 가는 판별 기준선도 달라져야 하는데, 즉 1차 선별 과정에서 부정오류를 줄이고 행동 문제가 있는 유아를 최대한 선별해내는 데 초점을 두는 경우에는 좀 더 낮은 점수를, 부정오류를 최소화하여 보다 엄격하고 확실한 평가 기준으로 활용할 때 좀 더 높은 점수가 적합한 변별 기준선이 될 수 있다. 따라서 가장 변별력이 높은 기준선을 정하는데 있어 절대적인 기준이 있기는 어렵지만, 일반적으로 활용되는 기준선들이 보이는 변별력 및 긍정오류, 부정오류를 검토하여 판단하는 것이 최선일 것이다. 총점 및 내외현화 척도는 CBCL 1.5-5가 사전 판별 도구로 사용될 때

가장 많이 활용되는 상위 척도인데, 원저자가 미국판에서 사용하였던 경계선 임상, 임상 기준선인 T점수 60, 64를 한국판 CBCL 1.5-5에 적용하는 것이 타당한지를 재확인할 필요가 있다.

64, 60T보다 높거나 낮은 점수대를 포함하여 각 점수대의 분류 정확성을 함께 검토하기 위해, 55, 60, 64, 70T를 기준으로 한 승산비 분석을 한 결과가 표 7에 제시되어있다. 우선 총점의 경우, 분석에 포함된 4가지 T점수 기준에 대해서 승산비가 모두 유의하였고, 60T 기준의 경우 승산비는 3.89였으며 49.56%의 임상집단을 제대로 변별하였고, 정상집단의 20.18%를 임상집단으로 잘못 분류하는 것으로

표 7. 한국판 CBCL1.5-5의 경계선 임상 T 점수 기준에 의한 임상·정상군 분류 적합성 검증을 위한 승산비 분석 결과

	Pearson Chi-Square	승산비	한국판 CBCL 1.5-5의 임상 T점수 해당 비율(%)		95% 신뢰구간	
			임상군	정상군	Lower	Upper
총점 55T 기준	4.16**	4.16	66.67	49.56	2.82	6.15
총점 60T 기준	43.35**	3.89	49.56	20.18	2.57	5.89
총점 64T 기준	25.06**	3.24	32.89	13.16	2.02	5.19
총점 70T 기준	9.28**	5.66	7.02	1.32	1.63	19.70
내재 55T 기준	54.77**	4.25	66.23	48.90	2.87	6.29
내재 60T 기준	36.01**	3.51	45.61	19.30	2.31	5.34
내재 64T 기준	26.01**	3.30	33.33	13.16	2.06	5.29
내재 70T 기준	5.58*	4.17	5.26	1.32	1.16	14.97
외현 55T 기준	74.25**	5.54	70.61	50.44	3.71	8.27
외현 60T 기준	39.31**	3.81	44.74	17.54	2.48	5.85
외현 64T 기준	35.15**	4.37	32.89	10.09	2.62	7.29
외현 70T 기준	11.65**	14.85	6.14	.44	1.94	113.91

주. * $p < .05$, ** $p < .01$, 한국판 CBCL 1.5-5의 임상 T 점수 해당 비율 수치의 단위는 %.

나타났다. 64T의 경우 승산비는 3.24로 약간 감소하였으며 임상집단의 32.89%를 제대로 임상집단으로 변별하고, 정상집단을 임상집단으로 잘못 변별할 확률은 13.16%로 보고되었다. 55T 기준의 경우, 승산비는 60T, 64T 기준일 경우보다 높고 실제 임상집단을 정확하게 분류할 수 있는 가능성이 66.67%로 높아지긴 하지만, 정상집단을 임상집단으로 오분류 할 가능성이 49.56%에 달하였다. 또한 70T 기준을 적용한 경우에는 승산비가 5.66으로 가장 높아 정상집단을 잘 못 분류할 확률이 1.32%로 매우 낮지만, 임상집단의 경우에도 7.02%만을 임상집단으로 분류하기 때문에 사전 판별도구로 사용할 때는 지나치게 엄격한 기준으로 볼 수 있어, 한국판 CBCL 1.5-5에서도 미국의 경우와 마찬가지로 총점의 64, 60T점수를 각각 임상, 경계선 임상 점수의 기준선으로 사용하는 것에 무리가 없을 것으로 판단된다.

내재화 및 외현화 척도의 경우에도 55, 60, 64, 그리고 70T 점수의 승산비는 모두 통계적으로 유의하였고, 이와 같은 전반적인 점수대별 분류 정확성 수준과 임상적 활용에 있어서의 유용성을 고려할 때, 총점의 경우와 마찬가지로 내재화 및 외현화 척도에 대해서도 64, 60T 점수를 각각 임상, 경계선 임상 집단 구분의 기준 점수로 사용하는 것이 적절한 판단으로 보인다.

표 8에는 한국판 CBCL 1.5-5의 하위척도들(총점, 내재화, 외현화 척도는 표 7 참고)의 경계선 임상 기준선으로 제시되는 65T점 기준을 적용하였을 경우의 승산비 값들을 제시하고 있다. 승산비가 가장 높은 것으로 나타난 하위척도는 주의집중 척도였으며 승산비는 11.12였다. 이는 주의집중 척도에서 65T 이상의 점수를 받을 경우, 그렇지 않은 경우에 비해 실

제로 임상군에 해당할 가능성이 11.12배 정도 높음을 시사한다. 그밖에 상대적으로 높은 승산비를 보인 하위척도들은 DSM 전반적 발달장애 및 위축 하위척도였는데 승산비는 각각 9.76과 7.77이었다. 승산비가 유의하지 않았던 두 개의 하위척도는 신체증상 및 수면문제 척도이다. 이 중 신체증상 척도는 통계적으로 유의한 수준은 아니나 정상군에서 임상점수 이상을 보인 사례 비율이 약간 더 높은 역전현상까지 보이고 있어 65T이상의 임상 범위에 대한 점수에 대해 주의 깊은 해석이 필요함을 알려준다.

정서적 반응성, 불안/우울, 신체증상, 위축, 수면문제, 주의집중문제, 그리고 공격행동의 7개 하위척도 중 1개라도 65T 이상의 경계선 임상범위에 해당할 경우를 기준으로 삼고, 실제 임상집단 여부를 어느 수준으로 정확하게 변별할 수 있는지를 탐색해 본 결과 승산비는 8.52였고, 실제 임상집단을 정확하게 임상집단으로 구별해내는 비율이 79.39%, 정상집단을 임상집단으로 잘못 구분하는 비율이 31.14%였다. 그리고 DSM 정서문제, DSM 불안문제, DSM 전반적 발달장애, DSM ADHD, 그리고 DSM 반항행동문제의 5개 척도 중 1개라도 65T 이상의 경계선 범위에 해당할 경우를 기준으로 삼았을 때는 승산비가 7.50, 임상집단을 임상집단으로 올바르게 분류하는 비율은 74.12%, 정상집단을 임상집단으로 잘못 분류하는 비율은 27.63%였다. 이와 같은 결과는 한국판 CBCL 1.5-5의 하위척도들 중 적어도 하나 이상에서 경계선 임상집단 구분 기준인 65T를 넘는 경우 임상적으로 의미가 있는 것으로 평가하는 것이 타당하다는 것을 지지하는 결과라고 볼 수 있을 것이다.

표 8. 한국판 CBCL 1.5-5의 경계선 임상 T 점수 기준에 의한 임상·정상군 분류 적합성 검증을 위한 승산비 분석 결과

	Pearson Chi-Square	승산비	한국판 CBCL 1.5-5의 임상 T점수 해당 비율(%)		95% 신뢰구간	
			임상군	정상군	Lower	Upper
정서적 반응성 65T 기준	9.26**	2.30	19.74	9.65	1.33	3.98
불안/우울 65T 기준	10.91**	2.39	21.93	10.53	1.41	4.04
신체증상 65T 기준	.11	.90	8.33	9.21	.47	1.72
위축 65T 기준	79.82**	7.77	50.00	11.40	4.79	12.61
수면문제 65T 기준	.02	1.04	14.04	13.60	.61	1.77
주의집중문제 65T 기준	117.65**	11.12	61.84	12.72	6.93	17.84
공격행동 65T 기준	12.45**	2.63	21.05	9.21	1.52	4.56
기타문제 65T 기준	41.97**	4.74	36.84	10.96	2.89	7.77
증후군척도 ¹ ≥ 65T 1개 이상	107.33**	8.52	79.39	31.14	5.56	13.04
DSM 척도 ² ≥ 65T 1개 이상	98.59**	7.50	74.12	27.63	4.95	11.36
DSM 정서문제 65T 기준	14.16**	2.59	25.00	11.40	1.56	4.30
DSM 불안문제 65T 기준	6.83**	2.10	17.54	9.21	1.19	3.69
DSM 전반적 발달장애 65T 기준	106.43**	9.76	59.65	13.16	6.12	15.55
DSM ADHD 65T 기준	27.80**	3.79	29.82	10.09	2.26	6.35
DSM 반항행동문제 65T 기준	17.02**	3.83	17.54	5.26	1.95	7.52

주. ** $p < .01$, 한국판 CBCL 1.5-5의 임상 T 점수 해당 비율 수치의 단위는 %, 증후군 척도¹는 정서적 반응성, 불안/우울, 신체증상, 위축, 수면문제, 주의집중문제, 그리고 공격행동의 7개 하위척도를 의미, DSM 척도²는 DSM 정서문제, DSM 불안문제, DSM 전반적 발달장애, DSM ADHD, 그리고 DSM 반항행동 장애의 5개 척도를 의미.

논 의

한국판 CBCL 1.5-5의 임상적 타당도를 확인하고 임상 장면 활용 시의 시사점을 살펴보기 위한 목적으로 수행된 본 연구에서 나타난 결과 및 논의점은 다음과 같이 정리할 수 있다.

첫째, 한국판 CBCL 1.5-5의 하위척도들 중 신체증상과 수면문제를 제외한 나머지 척도들

은 모두 임상집단과 정상집단을 효과적으로 분류할 수 있는 기준이 될 수 있음을 확인하였다. 변별력을 확인할 수 있는 여러 차원의 분석에서 일관된 결과가 관찰되었는데, 먼저 ANOVA 분석을 통해 모든 한국판 CBCL 1.5-5의 하위척도에서 임상집단이 정상집단에 비해 높은 점수를 보고하였다. 그리고 승산비 분석을 통해 신체증상과 수면문제를 제외한 모든

하위척도에서 효과적인 판별이 이루어지고 있음을 확인할 수 있었다. 특히 판별분석 결과는 한국판 CBCL 1.5-5의 모든 하위척도들을 동시에 고려하여 임상군을 구별해 내는 정확성이 82.89%에 달하는 높은 수준임을 알려주고 있다.

둘째, 판별분석, ROC 분석, 그리고 승산비 분석의 결과를 통해서 한국판 CBCL 1.5-5의 모든 하위척도들이 동등한 수준의 임상적 변별력을 갖는 것은 아니며, 일부 하위척도의 경우 임상적 변별력이 상대적으로 낮은 수준임을 확인할 수 있었다. 주의집중문제 척도와 DSM 전반적 발달장애 하위척도는 가장 변별력이 뛰어난 것으로 나타났고, 신체증상과 수면문제 하위척도의 경우는 임상적 변별력이 통계적으로 유의미하지 않은 것으로 나타났다.

이러한 결과는 CBCL 1.5-5의 전반적인 임상적 변별력을 평가하기 위해, 다양한 행동 문제로 전문적인 도움을 찾은 실제 영유아 임상 집단의 점수를 정상집단의 점수와 비교한 결과이다. 문제 유아들을 사전에 평가하여 예방적 접근을 할 때 여러 다양한 문제행동 집단으로 발전할 수 있는 가능성을 예측하는 일반적인 지표를 파악하고 변별력을 검증하는 것은 우선적으로 필요한 일이다. 그러나 하위척도 간에 변별력 차이를 보이는 이유를 탐색하기 위해서는 궁극적으로 개별 하위 척도들의 변별력은 임상집단이 보이는 문제행동 특성과 함께 파악하는 과정이 필요할 것이다.

현재 CBCL 1.5-5에 대한 변별력 연구는 임상군과 정상군을 비교하는 초기단계이지만, 아동용 CBCL의 경우 Biederman 등(2007)에 의해 ADHD가 있는 아동 중 반항성 품행 장애가 있는 변별하기 위해 아동용 CBCL의 공격성 하위척도의 변별력을 평가한 연구가 이루어진 바 있다.

또한 Hudziak 등(2006)의 연구에서 강박장애 임상군의 변별을 위해 아동용 CBCL의 하위척도를 활용한 경우, Aschenbrand 등(2005)이 불안장애의 변별 진단력을 검증하기 위해 CBCL의 불안우울, 위축, 내재화 척도를 활용한 경우가 그 예이다. 국내에서는 이수진 등(2008)에 의해 K-CBCL(오경자, 이혜련, 홍강의, 하은혜, 1997)의 주의집중문제 하위척도를 통해 ADHD 아동 임상군의 변별에 관한 연구가 이루어진 바 있다. 이제 막 타당화 과정을 마치고 임상적 변별력에 대한 전반적 검증 작업을 마친 한국판 CBCL 1.5-5의 경우에도 각 하위척도와 세부 진단별 변별, 혹은 감별진단과 관련된 추후 연구가 필요할 것이다. 이와 같은 연구를 통해 DSM 발달장애 척도의 높은 변별력이 현재 임상집단에 영유아 발달과 관련된 장애가 많이 포함된 것에서 기인하는 것인지, 다른 행동 증상에 대해서도 포괄적인 변별력을 갖는 것인지 평가가 가능할 것이다. 또한 신체증상이나 수면문제 척도의 임상적 변별력이 특정 진단군에 따라 달라지게 되는 것인지, 아니면 이 하위척도들의 구성, 문항 번역 및 임상집단 변별 기준을 기타 척도들과 다른 방식으로 적용해야 하는지에 대해 시사점을 얻을 수 있을 것이다.

셋째, 이번 연구는 한국판 CBCL 1.5-5의 하위요인구조를 통한 임상집단 변별의 타당성을 지지함과 동시에 전국적인 표준화 과정을 통해 마련한 표준 점수에서 제시하는 임상적 범위 판단 기준선(내재화, 외현화, 총점에 대해선 60T, 나머지 하위척도들에 대해선 65T)이 타당성이 있음을 뒷받침한다. 즉, 한국판 CBCL 1.5-5의 하위척도들에서 이 기준 이상의 점수가 보고될 경우 실제 임상 집단에 해당할 가능성이 실제로 높음이 경험적으로 지지되었

으며, 영유아 임상집단 사전판별을 위한 도구로 활용될 수 있음을 시사한다. 다만 신체증상 수면문제 척도와 같이 이번 연구에서 임상적 변별력이 명확히 검증되지 않은 일부 척도의 경우에는 다른 나머지 척도의 수준에 대한 고려 없이 독립적으로 임상 집단 구분을 목적으로 사용되는 것에 대해 주의를 기울일 필요가 있다.

한국판 CBCL 1.5-5의 변별 타당도가 여러 가지 분석을 통해 지지되고, 효과적인 임상 장면에서의 활용 가능성을 확인하였으나, 이 척도를 임상 장면에서 활용함에 있어 영유아 문제 행동 수준에 대한 기타 연구 결과 및 임상가의 전문적 경험에 의거한 융통성 있는 활용이 이루어지는 것이 바람직할 것으로 생각된다. 특히, 한국판 CBCL 1.5-5의 각 하위척도들은 특정 영역의 문제 행동을 탐색할 수 있도록 하는 것이 목적적이고, 직접적인 진단을 위한 것은 아니며 아직까지 세부적인 진단별 변별 타당도가 검증된 것은 아닌 만큼 그 사용과 해석에 있어서의 주의가 필요하다. 한국판 CBCL 1.5-5를 활용한 다양한 임상 장면에서의 연구가 누적됨에 따라 기준 점수대의 조정 또는 하위 요인 구조에 대한 추가적 논의 등이 이루어질 수 있을 것이다.

참고문헌

김영아, 오경자, 문수중, 김유진 (2005). 한국판 유아용 문제행동 척도(K-CBCL preschool)의 표준화를 위한 예비연구, 한국심리학회 연차학술대회.
오경자, 김영아 (2008) CBCL1.5-5 한국판 유아행동평가 척도 -부모용. 서울: (주)휴노컨

설팅.

오경자, 이혜련, 홍강의, 하은혜 (1997). K-CBCL 아동·청소년 행동평가검사. 서울: 중앙적성출판사.
이경숙, 신의진, 전연진, 박진아. (2004). 한국 유아 행동문제의 경향과 특성: 서울지역을 중심으로. 한국심리학회지: 발달, 17(4), 53-73.
이미선, 조광순, 김주영, 강병호 (2001). 장애영·유아 조기발견 및 진단·평가방안 연구. 안산: 국립 특수교육원.
이수진, 하은혜, 오경자 (2008). ROC 분석을 통한 주의력결핍 과잉행동장애 아동의 변별: K-CBCL 소척도를 기준으로. 한국심리학회지: 임상, 27(1), 191-207.
황해익, 정혜영 (1996). 유아용 발달선별검사도구 개발을 위한 기초연구. 영유아보육연구, 1(1). 35-56.
Achenbach, T. M. (1992). *Manual for Child Behavior Checklist/2-3 and 1992 Profile*. Burlington, VT: University of Vermont, Dept. of Psychiatry.
Achenbach, T. M., Edelbrock, C., & Howell, C. (1987). Empirically-based assessment of the behavioral/ emotional problems of 2-3 year old children. *Journal of Abnormal Child Psychology*, 15, 629-650.
Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms and Profiles*. Burlington, VT: University of

- Vermont Research Center for Children, Youth, and Families.
- Anastasiow, N. J. (1996). *Implications of neurobiological model for early intervention*. In S. J. Meisels & J. P. Schnkoff (Eds.), *Handbook of Early Childhood Intervention* (pp. 196-216). New York: Cambridge University Press.
- Aschenbrand, S. G., Angelosante, A. G., & Kendall, P. C. (2005). Discriminant validity and clinical utility of the CBCL with anxiety-disordered youth. *Journal of Clinical Child and Adolescent Psychology, 34*, 735-746.
- Biederman, J., Ball S. W., Monuteaux, M. C., Kaiser, R., & Faraone, S. V. (2007). CBCL clinical scales discriminate ADHD youth with structured-interview derived diagnosis of oppositional defiant disorder (ODD). *Journal of Attention Disorders, 10*, 410-417.
- Crijnen, A. A. M., Achenbach, T. M., & Verhulst, F. C. (1997). Comparisons of problems reported by parents of children in 12 cultures: Total Problems, Externalizing, and Internalizing. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 1269-1277.
- Crijnen, A. A. M., Achenbach, T. M., & Verhulst, F. C. (1997). Comparisons of problems reported by parents of children in twelve cultures: The CBCL/4-18 syndrome constructs. *American Journal of Psychiatry, 156*, 569-574.
- Hester, P. P., Baltodano, H. M., Gable, R. A., Tonelson, S. W., & Hendrickson, J. M. (2003). Early intervention for children at risk for emotional/behavioral disorders: A critical examination of research methodology and practices. *Education and Treatment of Children, 26*, 362 - 381.
- Holmes, W. C. (1998). A short, psychiatric, case-finding measure. for HIV seropositive outpatients. *Medical Care, 36*(2), 237 - 243.
- Hudziak, J. J., Althoff, R. R., Stanger, C., Bejsterveldt, C. E. M., van Nelson, E. C., Hanna, G. L., Boomsma, D. I., & Todd, R. D. (2006). The Obsessive Compulsive Scale of the Child Behavior Checklist predicts obsessive - compulsive disorder: a receiver operating characteristic curve analysis. *Journal of Child Psychology and Psychiatry, 47*, 160-166.
- Hwang, H. J., & St. James-Roberts, I. (1998). Emotional and behavioral problems in primary school children from nuclear and extended families in Korea. *Journal of Child Psychology and Psychiatry, 39*, 973 - 979.
- Kim, Y. A., Oh, K. J., & Lee, J. (2007). A study for standardization and validation of K-CBCL 1.5-5, *Poster session presented at International Conference of Cognition, Consciousness, and Culture. Seoul, Korea.*
- Luoto, M. & Hjort, J. (2005). Evaluation of current statistical approaches for predictive geomorphological mapping. *Geomorphology, 67*, 299 - 315.
- Patterson, G. R., Capaldi, D. M., & Bank, L. (1991). *An early starter model for predicting delinquency*. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp.139-168). Hillsdale, NJ: Erlbaum.
- Qi, C. H., & Kaiser, A. P. (2003). Behavior problems of preschool children from low-income families: Review of the literature.

Topics in Early Childhood Special Education, 23,
188-216.

원 고 접 수 일 : 2008. 11. 25.

게 재 결 정 일 : 2008. 12. 13.

Discriminant Validity and Clinical Utility of the Korean Version of the Child Behavior Checklist for Ages 1.5-5

Jin Lee

Young Ah Kim

Kyung Ja Oh

Huno Consulting Inc.

Department of Psychology, Yonsei University

The purpose of this study was to verify the validity and clinical utility of the Korean version of the Child Behavior Checklist for Ages 1.5-5 (CBCL 1.5-5) by demonstrating that the total and subscale scores of the checklist provide valid criteria for classifying referred and non-referred groups. The study participants were a referred group of 228 preschoolers (153 boys, 75 girls, age range 1.5-5 yrs, mean age 3.3 yrs) and a non-referred group which was demographically matched for sex and age (n=228, mean age 3.3 yrs). An ANOVA, ROC curve analysis, discriminant analysis, and odds ratio analysis were performed on the syndrome scale data obtained on the participants. The results supported the efficiency of the Korean version of the CBCL 1.5-5 in referral status screening and its clinical validity. The overall results showed that, with the exception of the somatic complaints and sleep problem scales, the sub-scales of the Korean version of CBCL 1.5-5 were effective in discriminating between clinical and non-clinical groups, and the DSM pervasive developmental problems and DSM attention deficit/hyperactivity problems scales were the most successful. The clinical cutoff-score-criteria adopted for the Korean version of CBCL 1.5-5 were shown to be statistically valid. Further studies with a larger clinical sample and more detailed diagnostic information will be necessary in order to clarify why some sub-scales were more successful than others in discriminating the referral status of the pre-schoolers. The implications of using the CBCL 1.5-5 in Korea as a clinical and research tool were discussed.

Key words : CBCL 1.5-5, referred group, discriminant validity, odds ratio, ROC curve.