

한국판 BECK 우울 척도의 분할점과 분류 오류*

신 민 섭 김 중 술 박 광 배

서울대학병원 서울의대 충북대학교
소아정신과 정신과 심리학과

본 연구에서는 우울증 진단과 우울집단의 선별을 위한 한국의 BDI 분할점과 그러한 분할점의 적용시 발생할 수 있는 분류 오류율을 알아보았다. 분할점과 분류 오류율은 기저율에 따라 변화될 수 있으므로, 정상 집단 對 우울증 집단의 이론적인 기저율을 50:50으로 간주하였을 때와, 전집에서의 정상인과 우울증 환자의 실제적인 비율을 고려하여 기저율을 70:30으로 설정하여 두 가지 방법으로 분할점을 산출하였다. 89명의 우울증 환자 집단과 272명의 정상 성인 집단을 대상으로 프로빗 분석을 이용하여 두 집단을 구분하는 분할점을 산출한 결과, 기저율을 50:50으로 상정한 경우에는 BDI 16점이, 그리고 기저율을 70:30으로 상정하였을 경우에는 BDI 22점이 우울증과 정상 집단을 분류하는 분할점인 것으로 나타났다. BDI 16점은 우울증 집단의 71.6%를, 그리고 정상 집단의 77.3%를 정확히 분류한 반면에, BDI 22점은 우울증 집단의 60.2%를, 그리고 정상 집단의 94.8%를 정확히 분류하였다. 이러한 긍정 오류율과 부정 오류율과 관련하여 각각의 분할점의 적용이 바람직한 상황과 그렇지 못한 상황에 대해서 논의되었다.

Beck 우울 척도(Beck Depression Inventory: 이하 BDI로 약칭)는 우울 증상의 유무와 증상의 심한 정도를 평가하기 위한 목적으로 제작된 자기 보고형 척도이다(Beck, 1961). BDI는 우울증의 인지적, 정서적, 신체적, 그리고 동기

및 행동적 증상들에 상응하는 21문항으로 이루어진 검사로서, 각 문항에 대한 개인의 평정치를 합산한 총점이 우울 정도를 나타내게 된다. BDI는 피검자들이 각 증상들의 심한 정도를 단순히 Likert형 척도로 평가하기 보다는 개별 증상의 심한 정도에 따라 0점에서 3점까지 평정치가 부과된 구체적인 진술문에 응답하게 함으로써, 피검자들

* 이 논문은 1991년도 서울대학교병원 지정진료 연구비 보조로 이루어졌음.

이 자신의 우울정도를 수량화하는데 있어서 덜 어려움을 느낄 수 있다는 점(한홍무, 염태호, 신영우 등, 1986) 외에 우울 증상을 객관적으로 평가하는데 있어 실시하고 채점하기가 간편하다는 장점이 있으므로(Frewmow, De Perczel, & Ellis, 1990), 임상 장면에서 환자를 진단하거나 치료 효과를 평가하는데 뿐만 아니라, 연구 목적으로 피험자를 선발하는 도구로서 국내외에서 가장 많이 이용되고 있는 척도중의 하나라고 생각된다.

BDI의 심리측정적 속성에 대한 많은 연구 결과들은 BDI가 다양한 집단에 걸쳐 신뢰롭고 타당함을 입증해주고 있다(Beck, 1967; Frewmow et al., 1990). BDI의 내적 일관성 계수는 외국 연구에서 .73(Gallagher, Nies, & Thompson, 1982)에서 .94(Schaefer et al., 1985)로 보고되었으며, 이영호와 송종용의 국내 연구(1991)에서는 우울증 환자 집단에서 .95, 정상 집단에서는 .78로 보고되었고, 신민섭(1992)의 연구에서 산출된 내적 일관성 계수는 .90이었다. BDI의 검사-재검사 신뢰도에 대한 연구들(Gallagher, et al., 1982; Oliver & Burkham, 1979; Hatzembuehler, Parpal, & Matthews, 1983; 이영호와 송종용, 1991)에서 BDI의 검사-재검사 신뢰도 계수의 범위는 .60에서 .86으로 보고되었으며, 이러한 결과는 BDI로 측정된 우울 증상이 시간이 지남에 따라 변화될 수는 있으나 비교적 안정적임을 나타내준다. 또한 BDI는 기존의 다른 우울 척도 및 심리 장애를 측정하는 척도들과도 중등도 이상의 상관 관계를 보였으며, 이상과 같은 결과들은 모두 BDI가 내적 일관성이 높고, 비교적 시간적인 안정성이 있으며, 타당한 척도임을 나타내준다.

정신과를 방문한 환자들이나, 혹은 비정신과 환자들중에서 우울증 환자를 일차적으로 가려내는데 있어(screening) BDI 점수가 정신과적인 진단 면접을 하는 것과 비슷한 효과가 있는 것으로 보고되었다(Nielsen & Williams, 1980).

정신과 치료를 받고 있는 환자들 외에 다양한 신체 증상으로 내과나 신경과, 가정의학과 등을 방문한 환자들이나 대학 상담실을 찾는 학생들 중에는 위장된 우울증(masked depression)이나 반응성 우울증 등 우울 증상의 심한 정도와 증상 표현에서는 차이가 있으나 정신과적인 도움이 필요한 사람들이 많을 것이라 생각된다. 이러한 경우, BDI를 간편한 진단 도구로 사용하기 위해서는 BDI 점수상에서 몇점을 우울증 진단을 위한 분할점(cut-off score)으로 사용할 것인지에 대한 객관적인 기준이 필요하다. 무엇보다도 우울증 연구를 위한 피험자 선발을 위해 국내외에서 임상, 비임상 집단을 대상으로 BDI가 널리 사용되고 있는 바, 이러한 연구 결과가 임상 장면의 환자들에게 일반화되고, 우울증의 원인론과 치료를 위한 연구의 활성화에 기여하기 위해서는 엄격하고 체계적인 연구를 통해 BDI의 분할점을 확립하는 것이 우울증 연구 분야에서 시급한 문제라 할 수 있다.

외국에서는 다양한 집단을 대상으로 BDI에 대한 기준 연구가 활발히 이루어졌으며, BDI 점수에 입각한 우울 수준의 분류는 다음과 같다.

Beck 등(1985)은 BDI 10-15점을 경한 수준, 16-23점을 중등도, 24점 이상을 심한 수준의 우울증으로 분류하였고, Shaw, Vallis, 그리고 McCabe(1985)도 BDI 0-9점을 비우울 상태, 10-15점이 경한 수준, 16-23점을 중등도, 24-63점을 심한 우울증으로 분류하였다. 또한 Shaw등(1985)은 치료 후 측정된 BDI 점수에서 0-9점이 완전치유 상태, 10-15점이 부분적 치유(partial remission), 그리고 16-63점이 여전히 우울 증상이 남아있는 것으로 보고하였다.

위와 같은 결과에 입각해 볼 때, 외국의 경우 BDI 10점 이상을 우울증으로 간주함을 알 수 있으나, 한국의 경우에는 일반적으로 BDI의 평균치가 외국에 비해 높으므로(한홍무 등, 1986; 신민섭 등, 1991, 1992), 외국 준거를 그대로 적용하여 임상 진단을 내리거나 우울증 피험자를 선발한 경

우에 많은 진단 및 분류 오류가 발생할 수 있다. 즉, 실제적으로는 정상인데 우울증으로 분류될 긍정 오류(false positive)가 발생할 가능성이 높다. 한국 정상 성인의 BDI 평균치가 서구 문화권의 그것보다 일관성있게 높은 이유에 대해서는 우울 증상의 원인 및 표현에 대한 유교 문화적인 영향, 척도 번역상의 문제점 등 몇가지 가능성을 생각해볼 수 있으나, 이에 대한 자세한 고찰은 본 연구에서 다루고자하는 주제를 벗어나는 것이므로 그 점에 대한 논의는 후속 연구에서 다루고자 한다.

외국 준거를 적용시 발생할 수 있는 긍정 오류의 문제를 해결하기 위하여, 한홍무 등(1986)은 BDI에 대한 한국판 표준화 연구를 수행하였는데, 그들은 연구에 포함된 정상 집단의 평균에서 1 표준편차 떨어진 점수인 BDI 21점을 잠정적으로 우울증 진단의 분할점(cut-off score)으로 제시하였다. 그러나 이 점수는 외국의 연구들에서 보고된 중등도의 우울 수준에 상응하므로, 이러한 준거를 적용할 경우에도 역시 외국 준거의 적용시와는 상반되는 분류 오류가 뒤따르게 된다. 즉, 우울증 환자를 정상으로 잘못 분류할 부정 오류(false negative)의 가능성을 모면할 수 없다(이영호와 송종용, 1991). 따라서 이영호와 송종용(1991)은 이러한 문제에 대한 절충적인 해결책으로 외국의 기준과 한홍무 등이 제시한 기준 사이의 점수가 적절한 분할점이 될 것으로 생각하여, 그들의 연구에 참여한 우울증 환자집단의 하위 25%에 해당하는 16점을 분할점으로 제시하였다.

이와 같이 한국에서의 BDI 분할점 확립을 위한 노력은 여러 차례 있어 왔으나, 대개가 정상 집단만을 대상으로 하거나 우울증 집단의 BDI 사분위 점수에 입각한 것이며, 정상 집단과 우울증 환자 집단을 연구에 함께 포함시켜 두 집단을 구분하는 분할점을 산출한 방법론적으로 엄격하고 치밀한 연구가 시도된 바는 아직 없었다고 생각된다. 임상 장면에서는 BDI 점수만을 가지고 환자에 대한 결정적인 진단을 내리는 경우가 드물므로 덜 문제가 되나, 연구 목적으로 BDI 점수에 입각해서 우

울집단을 선별하는 경우에는 분할점을 확립하는 것이 최우선의 일이라 할 수 있다. 임상 장면에서 치료를 받고 있는 우울증 환자들을 대상으로 연구를 수행하기에는 피험자를 모집하기가 어려울 뿐만 아니라, 여러가지 변인들에 대한 실험적인 통제를 하기가 곤란하다는 실제적이고 방법론적인 문제점이 많으므로, 대다수의 연구들에서는 정상 성인이나 대학생 집단을 대상으로 BDI 점수의 사분점이나 평균치에 입각해서 우울한 피험자를 선별하고 있는 실정이다. 따라서 우울증 피험자를 선별하기 위한 BDI의 분할점을 확립하는 것이 우울증에 대한 연구 결과의 내적, 그리고 외적 타당도(internal and external validity)에 있어 매우 중요한 문제라고 할 수 있다.

본 연구의 목적은 우울증 진단과 우울 집단의 선별을 위한 한국의 BDI 분할점과 그러한 분할점을 적용시 발생할 수 있는 분류 오류를 알아보기 위한 것이다. Beck(1961)은 척도 제작시 임상 환자들을 표준화 집단으로 사용하였으므로, 본 연구에서는 정상 집단과 우울증 환자 집단을 대상으로 두 집단을 유의미하게 분류할 수 있는 분할점을 결정하고자 한다.

그런데 분할점과 분류 오류는 기저율(base rate)에 따라 변화될 수 있다(박광배, 1992). 우울증 환자 집단에서는 분할점을 낮게 잡아야 그들이 정상으로 잘못 진단되는 부정 오류(false negative)율이 감소되는 반면, 정상 집단에서는 분할점을 높게 잡아야 그들이 우울증으로 잘못 진단되는 긍정 오류(false positive)율이 감소하게 된다. 그러므로 본 연구에서는 이러한 문제점을 해결하고자 두 가지 방법으로 분할점을 산출하고자 한다.

첫번째 방법으로, 전집에서의 이론적인 우울증 기저율을 .5로 간주하여 정상 집단과 우울증 집단을 분류하는 분할점을 산출하고자 하며, 두번째 방법으로는 전집에서의 정상인과 우울증 환자의 실제적인 비율을 고려하여 분할점을 산출하고자 한다. DSM-III-R(1987)에 따르면 성인 전집에서

주요 우울증(major depression)의 유병율(prevalence)이 남자의 경우에 2.3-3.2%, 여자의 경우에 4.5-9.3%이므로, 본 연구에서는 주요 우울증외에 기타 우울 장애를 다 포함한 유병율을 감안하여 정상 집단과 우울증 집단의 비율이 7 對 3인 경우에 두 집단을 구분하는 보다 엄격한 분할점을 산출하고자 하였다. 이와 같은 두 가지 방법에 의해 산출된 분할점 중에서 어느 것을 사용하느냐는 BDI 점수를 준거로 진단 분류를 하는 상황에서 연구자나 임상가가 어느 종류의 오류를 더 중요시하느냐에 따라 임의로 선택될 수 있다. 만일 우울증 환자를 정상으로 잘못 진단하는 것이 정상인을 우울증으로 잘못 진단하는 것보다 심각한 오류인 상황에서는 부정 오류가 최소가 되는 점수를 분할점으로 사용하는 것이 바람직하며, 그 반대인 상황에서는 긍정 오류가 최소가 되는 엄격한 점수를 분할점으로 사용하는 것이 권장될 수 있다.

마지막으로, 본 연구에서는 산출된 두 가지 분할점에 입각해서 BDI 점수로 정상 집단과 우울증 환자 집단을 변별하는데 있어서 발생할 수 있는 분류 오류율을 알아보고자 하였다.

방 법

피험자

서울에 거주하는 일반 회사원과 주부 272명(남 198, 여 74), 우울증 환자 89명(남 49, 여 40)이 본 연구에 피험자로 참여하였다. 우울증 환자 집단은 서울에 있는 4개 대학병원과 3개 개인 신경

정신과 의원에서 입원 치료를 받거나 통원 치료중인 환자로, 정신과 전문의와 임상심리학적 평가를 통해 우울증으로 진단된 환자로 구성되었다. 학력이 중졸이상이고 연령은 20세에서 45세 범위에 해당되는 환자들이었으며, 기질적 장애가 있거나 알콜 및 약물 중독, 성격 장애 진단이 수반된 환자와 정신병적인 증상을 보이는 환자들은 연구에서 배제되었다. 정상 성인 집단은 임상 환자 집단과 연령과 학력상에서 대응되는 회사원들과 주부들로 구성되었다.

BDI에서 절측치가 2문항 이상인 3개의 자료를 제외한 358명의 자료가 분석되었는데, 각 집단의 평균 연령과 학력 분포는 표 1과 같다.

도구 및 절차

본 연구에서는 BDI(Beck, 1967; revised, 1972)를 이영호(1991)가 번안한 것을 사용하였다. 정상 성인 집단에서는 연구자가 개별적으로 BDI를 실시하거나 회사원 연수 교육 중에 연수 교육 담당자가 집단으로 실시하였으며, 임상 환자들에게는 심리학자와 정신과 의사가 개별적으로 실시하였다.

자료분석 방법

자료분석은 프로빗 분석(probit analysis)을 이용하였다. 프로빗 분석은 연속적인(continuous) 독립 변인을 이용하여 범주적인(categorical) 종속 변인을 예측하기 위한 확률 모형이다. 이때 종속 변인이 범주 변인이므로, 프로빗 분석은 주어

표 1. 각 집단별 평균 연령과 학력 분포

집 단	연 령		학 력(%)	
	평균(표준편차)	중졸	고졸	대재 혹은 대졸
정상 성인	33.65(12.89)	4.2	38.40	57.4
우울증	35.29(16.85)	17.9	37.50	44.6

진 독립 변인의 값에 의해서 종속 변인의 한 범주가 발현할 확률을 도출하는데 이용된다(박광배, 1992). 본 연구에서 범주적인 종속 변인은 임상 진단 집단(우울증 對 정상)이고, 연속적인 독립 변인은 BDI 점수이다. 프로빗 분석에서는 이 두 가지 변인들을 이용하여 각 BDI 점수에 해당하는 피험자가 우울증 집단에 속할 확률을 도출하게 되는데, 이 확률이 .50인 BDI 점수가 우울증과 정상집단을 분류하는 분할점으로 이용될 수 있다. 왜냐하면 이 분할점보다 높은 점수를 받은 사람은 우울증 집단일 확률이 정상 집단일 확률보다 높은 반면에, 이 분할점보다 낮은 점수를 받은 사람은 우울증 집단일 확률이 정상 집단일 확률보다 낮기 때문이다.

그런데 앞서 언급한 바와 같이 분할점은 기저율에 따라 변화될 수 있으므로, 자료 분석시 이 점을 고려할 필요가 있다. 즉, 우울증 환자 집단에서는 분할점을 낮게 잡아야 그들이 정상으로 잘못 진단되는 부정 오류(false negative)율이 감소되는 반면, 정상 집단에서는 분할점을 높게 잡아야 그들이 우울증으로 잘못 진단되는 긍정 오류(false positive)율이 감소하게 된다. 따라서 본 연구에서는 두 가지의 기저율을 상정하여 분할점을 산출하였다. 특정한 피험자가 우울증 환자일 기저 확률을 50:50으로 정하기 위하여 우울증 환자와 동수(88명)의 정상인을 정상 집단에서 무선 표집하여 우울증 집단을 예측하기 위한 프로빗 분석을 실시하였고, 전집에서의 정상인과 우울증 환자의 실제적인 비율을 고려하여 정상 집단의 피험자를 모두 포함시켜 정상 집단 對 우울증 집단의 기저율이 70:30인 경우의 프로빗 분석을 다시 실시하였다. BDI 점수의 사용 목적에 따라 적절한 기저율과 분할점이 취사선택되어야 하는 바, 각각의 기저율과 분할점이 적절한 상황과 부적절한 상황에 대해서는 논의에서 검토하고자 한다.

일단 분할점이 구해지면 그 분할점을 이용하여 피험자들을 우울증 집단과 정상 집단으로 분류하는 경우에 발생하는 두 가지 오류율(부정 오류와

긍정 오류)은 즉각적으로 산출될 수 있다.

결 과

정상 성인과 우울증 환자 집단의 BDI 평균과 표준편차는 표 2와 같다. 선행 연구 결과들(Beck 등, 1985; Shaw 등, 1985; 한홍무 등, 1986; 신민섭 등, 1992; 이영호와 송종용, 1991)에 입각해 볼 때, 본 연구에 포함된 정상 성인 집단의 BDI 평균은 정상 범위에 해당되는 반면, 우울증 환자 집단은 심한 우울 수준에 속한다고 볼 수 있다.

표 2. 각 집단별 BDI의 평균과 표준편차

척도	정상 성인	우울증
	M (SD)	M (SD)
BDI	7.54(7.53)	25.18(11.56)

M: 평균, SD: 표준편차

프로빗 분석결과, 각 BDI 점수를 받은 사람들이 우울증 집단일 확률을 도출한 것이 표 3에 제시되어 있다.

표 3. 각 BDI 점수에서 우울증 환자 집단일 확률

확률	BDI 점수	
	기저율 50:50	기저율 70:30
0.01	-5.872	-3.260
0.02	-3.357	-0.336
0.03	-1.761	1.519
0.04	-0.560	2.914
0.05	0.416	4.049
0.06	1.247	5.015
0.07	1.976	5.862
0.08	2.629	6.620
0.09	3.222	7.310
0.10	3.768	7.945
0.15	6.030	10.574

확률	BDI 점수	
	기저율 50 : 50	기저율 70 : 30
0.20	7.828	12.663
0.25	9.370	14.455
0.30	10.755	16.065
0.35	12.038	17.556
0.40	13.256	18.971
0.45	14.434	20.341
0.50	15.594	21.688
0.55	16.753	23.036
0.60	17.931	24.405
0.65	19.149	25.821
0.70	20.432	27.312
0.75	21.817	28.922
0.80	23.359	30.714
0.85	25.157	32.803
0.90	27.419	35.432
0.91	27.965	36.067
0.92	28.558	36.757
0.93	29.211	37.515
0.94	29.940	38.362
0.95	30.771	39.328
0.96	31.747	40.463
0.97	32.948	41.858
0.98	34.544	43.713
0.99	37.059	46.636
유도비		
카이자승 (L^2)	32.02	44.80
자유도 (df)	43	43
유의도 (p)	.89	.40

표 3을 보면 BDI를 이용하여 우울증 집단을 예측하는 경우, 예측된 확률과 실제 비율 사이의 합치도가 기저율이 50 : 50일 때와 기저율이 70 : 30일 때 모두 좋음을 알 수 있다. 기저율이 70 : 30일 때의 카이자승은 자유도와 거의 동일한 44.

80 (df=43)으로서 통계적으로 유의미하지 않으며 ($p=.40$), 기저율이 50 : 50일 때도 역시 카이자승은 자유도보다 작은 32.02 (df=43)로서 유의미하지 않다 ($p=.89$). 따라서 BDI 점수는 우울증 집단을 예측하는데 타당한 지표가 될 수 있다.

BDI점수의 타당성에 기초하여 기저율이 50 : 50일 때, 특정한 피험자가 우울증 집단일 확률이 .5에 해당되는 BDI점수는 15.594이며, 기저율이 70 : 30인 경우의 그것은 21.688이다. 따라서 기저율을 50 : 50으로 상정하는 경우에는 BDI 16점이, 기저율을 70 : 30으로 상정하는 경우에는 BDI 22점이 우울증과 정상 집단을 분류하는 분할점이 될 수 있다. BDI에서 0보다 작은 점수는 존재하지 않으므로, 표 3에서 이들 점수에 해당되는 확률들은 가상의 확률이다.

다음으로, BDI 16점과 22점을 각각 분할점으로 사용하여 본 연구에 참여한 피험자들을 정상 집단과 우울증 집단으로 분류하였을 경우에 발생하는 분류 오류율을 산출한 결과는 표 4, 표 5와 같다.

표 4. BDI 16점을 분할점으로 사용한 경우의 분류 오류율

	정상 집단	우울 집단
BDI < 16	68 (77.3%)	25 (28.4%) 부정 오류
BDI ≥ 16	20 (22.7%) 긍정 오류	63 (71.6%)
전체	88 (100%)	88 (100%)

표 5. BDI 22점을 분할점으로 사용한 경우의
분류 오류율

	정상 집단	우울 집단
BDI < 22	254 (94.1%)	35 (39.8%) 부정 오류
BDI ≥ 22	16 (5.9%) 긍정 오류	53 (60.2%)
전체	270 (100%)	88 (100%)

BDI 16점을 분할점으로 사용하여 정상 집단과 우울증 집단을 분류하였을 때 전체 분류 적중율 $[(1-\text{분류오류 빈도수}) / \text{합/전체 빈도수}] \times 100$ 은 74.4%였다. 표 4를 보면 정상 집단을 우울증 집단으로 잘못 분류한 긍정 오류율이 22.7%이고, 우울증 집단을 정상 집단으로 잘못 분류한 부정 오류율은 28.4%였다. 반면에 BDI 22점을 분할점으로 사용하였을 경우에는 전체 분류 적중율이 85.8%로 16점을 사용하였을 때보다 더 높은 편이었다. 이는 BDI 22점을 적용시 정상 집단을 우울증 집단으로 잘못 분류한 긍정 오류율(5.9%)이 현저하게 감소된 점에 기인된 것이나, 오히려 우울증 집단을 정상 집단으로 잘못 분류한 부정 오류율은 39.84%로 BDI 16점을 사용하였을 때보다 다소 증가하였다(표 5).

이러한 결과는 BDI 22점을 적용하면 정상인을 우울증 환자로 잘못 분류할 확률(긍정 오류)은 매우 낮은 반면, BDI 16점을 사용할 경우에는 상대적으로 우울증 환자를 정상인으로 잘못 분류할 확률(부정 오류)이 낮아진다는 것을 나타내준다. 그러나 두 가지 종류의 오류는 서로 상대적인 것이므로 한 종류의 오류율을 감소시키면 다른 오류율이 증가하는 바, 어떤 분할점을 사용하든 분류 오류

가 발생하는 것은 불가피하다는 것을 알 수 있다.

논의 및 결론

임상 장면에서 우울증 환자의 진단 및 치료 효과에 대한 평가뿐만 아니라, 연구 목적으로 피험자를 선발하는 도구로써 BDI의 신뢰도와 타당도는 이미 국내외의 여러 연구들(Beck, 1967; Frewmow et al., 1990; 이영호와 송종용, 1991)을 통해 일관성있게 입증되어져 왔다. 그러나 BDI 점수상에서 몇점을 우울증 진단을 위한 분할점(cut-off score)으로 사용할 것인지에 대한 한국의 기준 연구는 체계적으로 이루어진 바 없으므로, 본 연구에서는 우울증 진단과 우울 집단의 선별을 위한 한국의 BDI 분할점과 그러한 분할점의 적용시 발생할 수 있는 분류 오류율을 알아보았다.

전집에서의 이론적인 기저율을 50:50으로 간주하였을 때와, 전집에서의 정상인과 우울증 환자의 실제적인 비율을 고려하여 정상 집단 對 우울증 집단의 기저율을 70:30으로 설정하여 두 가지 방법으로 분할점을 산출한 결과, 기저율을 50:50으로 상정한 경우에는 BDI 16점이, 그리고 기저율을 70:30으로 상정하였을 경우에는 BDI 22점이 우울증과 정상 집단을 분류하는 분할점인 것으로 나타났다.

본 연구에서 얻어진 우울증 진단을 위한 두 가지 분할점은 일반적으로 BDI 10점 이상을 우울증으로 간주하는 외국의 기준에 입각해 볼 때 중등도의 우울 수준을 나타내는 비교적 높은 점수라 할 수 있으나, 이러한 결과는 국내의 선행 연구들에서 보고된 점수들과 일치하는 것이다. 한홍무 등(1986)은 BDI 21점을, 그리고 이영호와 송종용(1991)은 16점을 잠정적인 분할점으로 제시한 바 있다. 그러나 기존의 연구들이 정상 집단만을 대상으로 하여 정상 집단의 평균치에서 1 표준편차 떨어진 점수나 우울 집단의 하위 25%에 해당되는 점수를 분할점으로 제시한 것에 비해, 본 연구의

장점은 전집에서의 정상인과 우울증 환자 비율을 감안하여 두 가지 방법으로 분할점을 산출한 것이라 할 수 있다. 이는 BDI 점수를 준거로 진단 분류를 하는 상황에서 연구자나 임상가가 어느 종류의 오류를 더 중요시하느냐에 따라 부정 오류와 긍정 오류가 최소가 되는 점수를 융통성있게 선택하도록 함으로써, 특정한 분류 오류가 발생하는 것을 어느 정도 통제하는 것을 가능하게 한다.

물론 본 연구에서는 연구자가 임의로 기저율을 두 가지로 설정하여 분할점을 산출하였으므로, 이는 현상을 지나치게 단순화시킨 것이며, 상황에 따라 기저율은 각기 다르리라 생각된다. 그러나 BDI가 적용되는 상황은 임상 장면이거나 비임상 장면, 혹은 임상적인 진단이 주목적이거나 피험자 선발이 주목적인 경우로 크게 구분될 수 있는 바, 본 연구에서 산출된 두 가지 기저율에 기초한 분할점은 연구자나 임상가가 판단을 내리는데 도움이 될 수 있을 것이다.

특정한 검사 점수를 기준으로 피검자들을 분류하는 경우가 많은데, 예를 들면 한 검사를 실시하여 피검자의 점수가 기준 점수보다 높은지 낮은지에 따라 합격/불합격, 혹은 환자/정상 등으로 분류하는 것이 이에 해당된다. 이때 분할점을 설정하는 이유는 객관적인 판단 기준을 세우기 위한 것이며, 이것이 절대적인 기준이 될 수는 없다. 왜냐하면 특정 집단으로 진단 분류를 하기 위해 어떠한 척도 점수를 사용하는 경우에 산출된 분할점은 모든 집단에 걸쳐 불변하는 것이 아니며, 특정 장애의 기저율, 표집 분포, 혹은 사전에 특정 집단으로 분류된 진단의 정확성 등 여러 요인들에 따라 변화될 수 있기 때문이다.

그러므로 특정 검사의 분할점에 입각해서 집단을 분류하는 경우에는 분류 오류가 필연적으로 뒤따르게 된다. 본 연구에서 산출된 분류 오류란 BDI 분할점에 입각해서 실제로 정상인인데 우울증 집단으로 잘못 분류되거나, 반대로 실제로는 우울증 환자인데 정상 집단으로 잘못 분류된 경우를 말한

다. BDI 16점을 분할점으로 사용하여 정상 집단과 우울증 집단을 분류하였을 때 정상 집단을 우울증 집단으로 잘못 분류한 긍정 오류율이 22.7%이고, 우울증 집단을 정상 집단으로 잘못 분류한 부정 오류율은 28.4%였다. 이러한 부정 오류율은 이영호와 송종용의 연구(1991)에서 BDI 16점이 우울증 환자 집단의 하위 25%에 해당되는 점수였으므로, 이 기준을 적용시 우울증 환자의 25%가 정상 집단으로 잘못 분류되는 비율과 거의 유사한 결과이다. BDI 22점을 분할점으로 사용하였을 경우에는 정상 집단을 우울증 집단으로 잘못 분류하는 긍정 오류율(5.9%)이 현저하게 감소된 반면, 우울증 집단을 정상 집단으로 잘못 분류한 부정 오류율은 39.84%로 BDI 16점을 사용하였을 때(28.4%)보다 오히려 증가하였다.

두 가지 종류의 분류 오류는 서로 상대적인 것이므로 한 종류의 오류율을 감소시키면 다른 오류율이 증가되는 바, 연구자나 임상가가 어떤 종류의 오류를 더 방지하기를 원하는가에 따라 두 가지 분할점중 어느 하나가 선택될 수 있다. 일반적으로, 우울증 환자를 정상으로 잘못 진단하는 것이 정상인을 우울증으로 잘못 진단하는 것보다 심각한 오류가 될 수 있는 임상 장면에서는 부정 오류율이 최소가 되는 BDI 16점을 우울증 진단의 분할점으로 사용하는 것이 바람직하리라 생각된다. 특히, 우울증 연구를 위해 피험자를 선발할 경우에 우울증상을 주소(chief complaints)로 정신과를 방문한 환자들 중에는 실제로 우울증일 가능성(사전 확률)이 일반 전집에 비해 높으므로 BDI 16점이 피험자 선발을 위한 합리적인 기준으로 권장될 수 있다. 그러나 대학생이나 일반 성인 집단을 대상으로 피험자를 선발하는 상황에서는 실제로는 정상인데 우울증 집단으로 잘못 분류되는 긍정 오류율이 최소가 되는 BDI 22점을 분할점으로 사용하는 것이 연구 결과의 내적, 그리고 외적 타당성을 보장하기 위해 바람직하리라 생각된다. 그 이유는 만일 정상인이 우울 집단으로 잘못 분류된다면,

그러한 연구에서는 독립 변인 조작(피험자 분류)에 있어서 근본적인 문제점이 오염 변인으로 작용하게 되어 연구 결과의 내적 타당도가 위협을 받게 될 뿐만 아니라, 연구 결과의 일반화에도 심각한 영향을 주기 때문이다. 본 연구에 포함된 정상 집단의 피험자들 중에서 BDI 22점 이상을 받은 사람들의 비율은 5.9%에 불과하였으므로, 피험자 선발시 이 기준을 적용한다면 정상인을 우울증 환자로 잘못 분류한데서 기인된 결과 해석 상의 문제점은 크게 우려하지 않아도 되리라 생각된다.

그런데 BDI 점수가 높다고 해서 반드시 우울증으로 진단될 수 있는 것은 아니라는 점(Kendall, Hollon, Beck, Hammen, & Ingram, 1987)은 간과해서는 안될 사실이다. 그들 중에는 불안하거나 약물 중독이거나 기타 다양한 정신과적인 문제를 가진 사람들이 포함되어 있을 수 있다(이영호와 송종용, 1991). 알콜 및 약물 남용 환자 집단의 BDI 평균치가 18.10(표준편차 12.90)인 것으로 나타난 신민섭의 연구(1992) 결과도 이러한 가능성을 시사해준다. 그들이 정말로 우울증인지를 결정적으로 진단하기 위해서는 개별적인 정신과적 면담이나 다른 질문지를 병행하는 것이 필요하다. 그러나 BDI 분할점 이상의 점수를 받은 사람들이 정말로 우울증 환자인지를 정확하게 진단하기 위하여 개개인을 1:1로 면담하기에는 시간, 비용, 노력이 많이 요구되는 실질적인 문제점이 뒤따르므로, 이러한 방법은 가장 이상적이긴 하나 임상 장면에서만 어느 정도 가능할 뿐이며, 연구를 위해 집단적으로 피험자를 선발하는 경우에는 사실상 실행하기가 불가능하다. 이에 대한 대안으로 선행 연구들에서는 BDI외에 다른 우울척도를 함께 사용할 것을 제안한 바 있다(이영호와 송종용, 1991; Lewinson & Teri, 1982). 그러한 연구들에서는 BDI와 기타 다른 자기보고형 우울 척도를 함께 실시하여 우울 집단을 분류한 결과, BDI만을 사용하였을 때보다 우울증이 아닌 사람을 우울 집단에 포함시킬 가능성이 감소되는

것으로 보고되었다.

결론적으로, 본 연구 결과는 BDI 점수가 우울증 집단을 예측하는데 타당한 지표가 될 수 있음을 보여주었다. BDI의 가장 큰 장점은 우울 증상을 평가하고 우울증 환자를 일차적으로 가려내는데 있어 비교적 신뢰롭고 타당하며, 정신과적인 면접에 비해 객관적이고, 실시하고 채점하기가 용이하다는 점이다. 그러나 BDI 분할점은 객관적인 판단 기준은 될 수 있으나, 모든 집단에 걸쳐 불변하는 것이 아니며, 특정 장애의 기저율, 표집 분포, 혹은 사전에 특정 집단으로 분류된 진단의 정확성 등 여러 요인들에 따라 좌우될 수 있으므로, 이것이 절대적인 기준이 될 수는 없음을 유념해야 할 것이다. 따라서 BDI 점수만을 가지고 최종적인 임상 진단을 내리는 것은 유보해야 하리라 생각되며, 어떠한 분할점을 적용시 연구자나 임상가는 그러한 상황에서 발생하는 분류 오류를 항상 염두에 두고 이를 사용하여야 하리라 생각된다.

마지막으로, 본 연구에서 산출된 우울증 진단을 위해 분할점은 외국의 기준에 비해 상대적으로 높은 편이었고, BDI뿐만 아니라 다른 자기보고형 척도들에 대한 한국 피험자들의 평균치들도 일반적으로 외국에 비해 더 높은 경향이 있으므로(Han, 1992; 한홍무 등, 1986), 외국에서 제작된 척도들을 번역해서 사용할 경우에 외국의 기준을 그대로 적용하는 것은 지양되어야 하리라 생각된다. 번안된 외국 척도에 대한 비교문화적인 심리측정적 연구는 언젠가는 다루어져야 할 과제라 생각되는 바, 이에 대한 연구가 조속히 이루어지기를 기대한다.

참고문헌

- 박광배(1992). 심리검사 제작의 이론과 실제: 절대평가 기준과 상대평가 기준. 한국심리학회.
박광배(1992). 빈도 분석. 성원사.
신민섭(1992). 자살 기제에 대한 실증적 연구: 자

- 기 도피 척도의 타당화. 연세대학교 박사학위 논문.
- 이영호, 송종용(1991). BDI, SDS, MMPI-D 척도의 신뢰도 및 타당도에 대한 연구. 한국심리학회지: 임상, 10(1), 98-113.
- 한홍무, 염태호, 신영우, 김교헌, 윤도준과 정근재(1986). Beck Depression Inventory의 한국판 표준화 연구: 정상 집단을 중심으로. 신경정신의학, 25, 487-500.
- American Psychiatric Association(1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., revised). Washington, DC: Author.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., & Erbaugh, J.(1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Beck, A.T.(1967). *Depression: Clinical, Experimental, and Theoretical Aspects*. New York: Harper & Row.
- Fremouw, W.J., De Perzel, M., & Ellis, T.E.(1990). *Suicide Risk Assessment and Response Guidelines*. New York: Pergamon Press.
- Gallagher, D., Nies, G. & Thompson, L. W.(1982). Reliability of the Beck Depression Inventory with Older Adults. *Journal of Consulting and Clinical Psychology*, 50, 152-153.
- Han, K.H.(1992). Korean Version of the MMPI-2: Inventory, Adaptation, and Equivalence Evaluation Using Korean, American, and Japanese College Samples. 한국심리학회 92 연차학술대회 MMPI Workshop 발표 논문.
- Hatzenbuehler, L.C., Parpal, M.M, & Mathews, L.(1983). Classifying College Students as Depressed or Nondepressed Using the Beck Depression Inventory: An Empirical Analysis: *Journal of Consulting and Clinical Psychology*, 51, 360-366.
- Lewinsohn, P.M. & Teri, L.(1982). Selection of Depressed and Nondepressed Subjects on the Basis of Self-Report Data. *Journal of Consulting and Clinical Psychology*, 50, 590-591.
- Nielsen, A.C. & Williams, T.A.(1980). Depression in Ambulatory Medical Patients. *Archives of General Psychiatry*. 37, 999-1004.
- Oliver, J.M. & Burkhart, R.(1979). Depression in University Students: Duration, Relation to Calendar Time, Prevalence, and demographic correlates. *Journal of Abnormal Psychology*, 88, 667-680.
- Schaefer, A., Brown, J., Watson, C.G., Plemel, D., Demotts, J., Howard, M. T., Petrik, N., Balleweg, B.J., & Anderson, D.(1985). Comparison of the Validities of the Beck, Zung, and MMPI Depression Scales. *Journal of Consulting and Clinical Psychology*, 53, 415-418.
- Shin, M.S., Oh, K.J., Park, K.B., & Kim, J.S.(1991). *The Depression-Hopelessness-Suicide Relationship in Korean Adolescents and Young Adults*. A Paper Presented at the American Psychological Association Annual Meeting, San Francisco, California.

The Cut-Off Score for the Korean Version of Beck Depression Inventory

Min-Sup Shin

Zoung-Soul Kim

Kwang-B. Park

Div. of child Psychiatry
Seoul National University
Children's Hospital

Seoul National University
Medical College

Depart. of Psychology
Choongbuk National
University

We determined the cut-off score for the Korean version of BDI in order to establish empirical and clinical criteria for screening clinical depressives from the populations with two different prior probabilities of clinical depression. To this end, BDI was administered to 89 patients who were diagnosed by psychiatrists and clinical psychologists as depression and 272 normal adults aged from 20 to 45. The cut-off score assuming the prior probability of .50 was turned out to be sixteen, which was determined by randomly selecting 89 normal subjects and analysing their BDI scores along with the scores of the 89 patients by a probit analysis. The cut-off score assuming the prior probability of .30 was twenty-two, which was determined by subjecting the BDI scores of the entire subjects to the same analysis. The cut-off score of sixteen under the assumption of .50 prior probability classified 71.6% of the depressives and 77.3% of the normals correctly. That of twenty-two under the assumption of .30 prior probability classified 60.2% of the depressives and 94.8% of the normals correctly. The situations where each of the two cut-off scores would be appropriate or not were discussed.