

Brief Psychiatric Rating Scale - 평정방법에 따른 평정자간 신뢰도와 자기보고식 검사와의 상관연구

김 민 경[†]

가톨릭대 성모자애병원
신경정신과

이 봉 건

충북대 심리학과

전 양 환

가톨릭대 성모자애병원
신경정신과

본 연구에서는 임상가의 직접관찰과 면담을 통해 환자의 정신증적 상태를 평가하는데 널리 사용되는 Brief Psychiatric Rating Scale(BPRS)의 Anchored Version(Gabbard 등, 1986; Overall, 1962)의 평정자간 신뢰도를 조사하고, BPRS와 자기보고식 검사와의 상관을 알아보았다. 우선 21명의 입원 정신증 환자를 대상으로 두 명의 평정자와 한 명의 평정자가 각각 짝진면접과 개별면접을 통해 평정한 후 평정자들 간의 신뢰도를 비교하였다. 다음에는 환자가 자신의 상태를 직접평가한 SCL-90-R과 BPRS 유사문항과의 상관연구를 통해 평정방법의 차이에도 불구하고(삭제) 안정된 일치도를 얻을 수 있는지 알아보려고 하였다. 연구결과 짝진면접자간의 신뢰도가 18문항 중 15문항에서 통계적으로 의미있는 수준에서 산출되어, 개별면접자간의 평정신뢰도에 비해 안정된 결과를 얻을 수 있음이 확인되었다. 또한, BPRS의 불안, 적대감, 의심 문항이 자기보고식 검사인 SCL-90-R의 불안, 적대감, 우울, 편집증, 대인민감성, 정신증 척도와 통계적으로 의미있는 정적 상관을 나타내어 유사문항간의 관계성이 확인되었다. 본 연구의 의의는 국내에서 연구된 적이 없는 BPRS-Anchored Version(Gabbard 등, 1987)의 신뢰도를 조사한 것이며, 같은 증상들이 다른 평가방법으로 평가될 때의 결과의 일치도와 더불어 연구방법론상의 제한점과 장차 연구를 위한 보완점을 제시한 것이다.

주요어 : BPRS(Brief Psychiatric Rating Scale), 평정자간 신뢰도, 짝진면접, 개별면접, SCL-90-R, 자기보고식 검사와의 상관

[†] 교신저자(Corresponding Author) : 김 민 경 / 가톨릭대 성모자애병원 신경정신과 / 인천시 부평구 부평 6동 665
FAX : 032-505-8994 / E-mail : cpp2001@hanmail.net

환자의 증상을 평가하는 방법은 크게 임상가가 환자를 직접 평가하는 평가척도(rating scale) 방법과 환자 스스로 자신의 증상을 평가하는 자기보고식(self-report) 방법이 있다. 두 방법은 각각 문제점이 있는데, 임상가의 직접평가 방법은 전문가의 시간소요가 많다는 단점이 있는 반면, 환자 스스로의 보고에 근거한 자기보고식 평가는 사회적 바람직성에 답하려는 반응자의 속성으로 인해 결과가 왜곡될 수 있다는 것이다. 현재 국내에서 자주 사용되는 임상가의 직접평가용 도구에는 Brief Psychiatric Rating Scale(이하 BPRS; Overall & Gorham, 1962), Hamilton Rating Scale for Depression(Guy, 1976), Hamilton Rating Scale for Anxiety(Guy, 1976), The Positive and Negative Syndrome Scale(Kay, Fiszbern & Opler, 1987), Zung Self Rating Depression Scale(Zung, 1970), Nurses Observation Scale for Inpatient Evaluation(Homigfeld, 1970), Young Mania Rating Scale(Youn 등, 1978) 등이 있으며, 자기보고식 검사로 임상장면에서 흔히 이용되는 도구에는 MMPI, SCL-90-R, Beck Depression Inventory, Beck Anxiety Inventory, Child Depression Inventory, K-PIC, PAI(Personality Assessment Inventory) 등이 있다.

이중 BPRS는 약물치료에 따른 환자의 증상변화를 보기 위한 목적에서 개발된 것으로, Lorr가 개발한 다차원척도(Multidimensional Scale for Rating Psychiatry Patients, 1953)와 입원환자들을 위한 다차원적인 정신병척도(Inpatient Multidimensional Psychiatric Scale, 1960)를 대상으로 군집분석과 요인분석을 실시하여 추출한 14개의 대표적 요인에, 비정상적인 사고내용(unusual thought content)과 무딘 감정(blunted affect)의 두 요인을 추가하여, 총 16개 문항으로 만든 것이다(Overall & Gorham, 1962). 4년 후에는 증상분류의 효율성을 높이기 위해서, 흥분(excitement)과 방향감각상실

(disorientation)을 평가하기 위한 두 문항이 추가되었다(Overall, 1974). BPRS는 초기의 개발목적에 비해 더욱 폭넓은 용도로 사용되고 있다. 즉, 정신과 입원환자들의 약물치료에 따른 변화를 효율적이며 빠른 시간에 보기 위한 목적뿐만 아니라(Overall & Gorham, 1962), 치료반응의 예측이나 환자의 증상을 설명하기 위해서도 사용되며, 최근에는 심리치료 프로그램의 효과검증을 위해 이용되기도 한다. 이와 같이 BPRS가 지속적으로 사용되는 이유로서, Rhoades & Overall(1988)은 다음의 여섯 가지를 들고 있다: 첫째, 이 척도의 구성요인에는 통상 임상가들에 의해서 정신증 환자에게서 나타난다고 생각되는 대표적 증상들이 포함되어 있다. 둘째, 정신병리의 심각성에 대한 평가기준이 전문가의 임상적 경험이다. 셋째, 다양한 임상 및 연구 상황에 이용될 수 있는 유연성이 있다. 넷째, 단순하면서도 효율적인 방식으로 정신증에서 나타나는 다양한 증상을 평가하며, 임상적인 증상을 구별하는 데 유용하고 특별한 목표증상보다는 평가할 때의 상태에 초점을 두어 제반 증상에 대하여 전반적인 윤곽을 알려준다. 다섯째, 수년간에 걸쳐 요인구조나 점수, 프로파일 분류 등의 방법론적 연구가 많이 이루어졌다. 여섯째, 오랜 기간 지속적으로 사용되어 왔기 때문에, 연구시 중요한 참고준거가 된다. 이를 토대로 볼 때, BPRS는 임상 및 연구 장면 모두에서 융통성있게 적용될 수 있는 평가도구라 하겠다.

BPRS는 16문항으로 구성된 초기의 것과 후에 문항이 첨가된 18문항, 24문항 짜리가 있는데, 자주 사용되는 것은 16문항과 18문항으로 구성된 것이며, 본 연구에서는 18문항으로 구성된 BPRS를 사용하였다. 18문항짜리 BPRS는 환자의 언어적 보고만을 근거로 평가해야 하는 “신체적 관심, 불안, 개념적 혼란, 죄책감, 과대망상, 우울

한 기분, 적대감, 의심, 환각행동, 비정상적인 사고내용, 지남력 상실” 등 11문항과 행동관찰을 근거로 평가해야 하는 “감정절회, 긴장, 버릇과 자세, 운동지체, 비협조, 무딘 감정, 흥분” 등 7문항으로 구성되어 있다. 각 문항은 증상이 거의 없는 것에서부터 극심한 정도까지 Likert 방식의 7점 척도(0-6점, 1-7점)나 5점 척도상에 평정하도록 되어있다.

외국에서 수행된 BPRS의 신뢰도와 타당도에 대한 연구는 많은 편인데, 신뢰도의 경우 각 문항들이 개별 속성을 가지고 있기 때문에 전체문항에 대한 내적 일관성 신뢰도를 산출하는 것은 의미가 없으며 따라서 개별문항에 대한 평정자간 신뢰도 산출에 초점이 맞추어졌다. 그러나 기존 연구 대부분이 전체점수를 합산하여 산출된 신뢰도계수를 보고한 경우가 많으며, 평가자들이 개별면접을 했는지 짝진면접을 시도했는지조차 보고하지 않은 경우가 많다(Lachar 등 2001). 더구나 국내의 경우는 신뢰도와 타당도 연구가 거의 없는 실정이다.

Heldung와 Vieweg(1980)가 당시까지의 BPRS 연구결과를 개관한 논문을 보면, 검토된 8개 논문 모두에서 환자의 언어적 보고에 근거해 정보를 얻게 되는 환각행동, 의심, 과대망상, 개념적 혼란 문항의 신뢰도가 .80이상으로 높은 반면 행동관찰을 통해 정보를 얻게 되는 무딘 감정이나 감정절회 문항은 .60 수준으로 낮았다. 또 문항총점의 신뢰도는 .70이상이지만, 개별문항의 신뢰도는 .16에서 .95의 범위로 불안정하였다. 이는 개별 문항 모두에서 안정적인 신뢰도를 확보하는 것은 어려울 수 있음을 시사한다. 개별면접과 짝진면접의 결과를 제시한 Manchanda(1986)의 연구에서는 운동지체와 무딘 감정을 제외한 모든 문항에서 개별면접의 신뢰도가 낮았으며, 총점의 신뢰도는 짝진면접이 .88, 개별면접이 .57

이었다. Flemenbaum과 Zimmermann(1973)은 신뢰도에 관한 연구결과를 일반화하는데 따른 문제점을 지적하였다. 이들은 초기의 신뢰도 연구는 오랜 훈련을 거친 평정자들에 의해 이루어진 만큼 경험이나 지식, 임상적인 기술, 평가기법의 숙달정도에 다양한 차이가 있는 실제 평가상황에 적용하기가 어려움을 시사하였다. Crippa 등(2001)도 임상 실제에서는 한 명의 평정자에 의한 평가가 많으며 평정 상황도 상이한 만큼, 신뢰도를 높이기 위한 방법으로 사용된 짝진면접이나 녹화된 비디오테이프를 보고 평정해서 나온 연구결과를 실제 장면으로 일반화하는 것에는 문제가 있다고 보았다. Hafkenscheid(1991)는 1978년부터 1990년 사이에 출판된 항정신의약 연구 15편을 무선표집하여 분석하였는데 단지 2개의 연구만이 짝진면접법에 의해 평정된 것임을 밝혀냈다. 또 이들은 평정자간의 신뢰도 연구를 위해서는 짝진면접법이나 비디오테이프를 이용한 방법 등 신뢰도를 높일 수 있는 방법이 이용되지만, 약물치료의 효과나 환자의 증상변화를 평가하기 위한 보다 실제 상황에서는 짝진면접법이나 비디오테이프를 이용한 평정법은 시간, 인력, 장비 등이 소요되어 거의 사용하지 않고, 주로 한명의 환자를 한명의 평정자가 두 번에 걸쳐 평가하는 개별면접법이 일반적으로 사용됨을 지적하였다.

평정자간 신뢰도에 영향을 미칠 수 있는 요인으로서는 평정자의 면접방식, 평정자의 교육 배경 및 경험수준, 증상에 대한 평정자의 정의, 환자의 반응 등이 있을 수 있다. 그러나 이 모든 것을 연구자가 통제하는 데에는 한계가 있다. 평정척도가 실용적인 도구로 사용될 수 있으려면, 평정될 속성이나 특질을 명확하게 정의하여 평정자에 따른 차이를 극소화시켜야 할 것이다(Manchanda, 1986). 즉 증상에 대한 개념적 설명

이나 평정기준을 명확히 정의해 놓는 것이 평정자간 평가차이를 줄이는 데 도움이 될 수 있다. 이런 맥락에서, 반구조화된 질문을 개발하거나 각 문항에 대한 개념을 조작적으로 정의하여 기술하는 방법, 척도수준에 대해 구체적 설명을 추가하는 방법 등 평정자간 차이를 줄이기 위한 시도들이 이루어졌다(Crippa 등 2001; Gabbard, Kenndey, David, Deering & Rason, 1987; Overall & Gorham, 1988; Woerner, Mannuzza, & Kane, 1988). BPRS는 개발초기에는 훈련받은 숙련가가 실시할 수 있도록 구성되었다(Anderson 등, 1989; Overall & Gorham, 1962; Raskin, 1988). 관련된 연구로서, 경험이 많은 평정자들에 대한 신뢰도가 경험이 부족한 평정자들의 신뢰도에 비해 안정적이라는 보고가 있다(Bech, Larsen, & Anderson, 1988). 경험의 정도에 따른 신뢰도의 차이를 줄이기 위한 방법으로는, 구조화된 질문의 개발, 문항에 대한 개념적 정의의 구체화, 증상의 심각도에 대한 설명을 추가하는 것 등이 있을 수 있다. 신뢰도를 높이기 위한 일환으로, Gabbard(1987)는 개별 문항마다 증상의 심각도에 구체적 기술과 설명을 덧붙인 수정된 BPRS를 개발하였다(이하 BPRS-A). 또한 Hudlung와 Vieweg(1980)는 반구조화된 면접에 사용될 선도성 질문을 개발하여 평정자간 차이를 줄이고자 시도하였고, Crippa 등(2001)도 같은 목적으로 면접 안내용 질문을 개발하였다. 원래의 BPRS와 반구조화된 질문 및 척도점수에 구체적 설명이 추가된 수정된 BPRS간의 신뢰도를 비교한 연구들은, 후자의 경우 평정자간 신뢰도가 높음을 확인하였다(Gabbard, 1987, Woerner 등 1988). 따라서 환자의 반응을 유도하기 위한 구조화된 질문을 사용하거나, 척도 수준에 구체적 설명이 추가된 BPRS를 사용하는 것이 평정자간 신뢰도를 안정되게 확보할 수 있는 방법이라고 할 수 있다. 그러나 외국에서도 이를 이용한 연

구들은 적으며(Faustman, 1994), 특히 국내에서는 보고된 바가 없다. 이에 김민경(1993)은 Gabbard 등(1987)의 수정된 BPRS와 원래의 BPRS를 사용하여 평정자간 신뢰도를 비교한 연구결과를 보고하였다. 연구결과, 행동관찰을 토대로 평가해야 하는 '긴장' 문항과 '버릇과 자세' 문항을 제외한 16문항 모두에서, Gabbard 등(1987)에 의해 개발된, 수정된 BPRS-A를 사용한 경우의 신뢰도가 높게 나타났다.

평정자간 신뢰도를 높이기 위해 고안된 반구조화된 질문지(Crippa 등, 2001)와 수정된 BPRS-A(Gabbard, 1987)를 이용하는 것이 신뢰도를 높여줌이 이미 밝혀졌기 때문에(김민경, 1993), 본 연구에서는 우선, 개별면접과 짝진면접에 따른 평정자간 신뢰도를 비교하고, 그 다음으로는 평정방법이 상이한 자기보고식 검사와의 상관관계를 조사하였다. 임상장면에서는 숙련된 임상가보다 수련중인 비숙련가가 평가하는 경우가 많을 수 있고, 평정방법 또한 짝진면접법보다는 개별면접법이 주로 이용될 것이다. 실제 장면에서 주로 사용되는 개별면접에 의한 평정이 짝진면접과 마찬가지로 안정적일 수 있다면, BPRS는 검사자 변인이나 상황 요인의 영향을 적게 받는 안정된 평가도구라 할 수 있다. 이에 따라서 면접 유형에 따른 평정결과를 비교하고자 하였다. 다음에는, 유사한 증상을 측정하기 위해 개발된 자기보고식 검사의 소척도 점수와 BPRS의 하위요인과의 상관연구를 통해, 두 평가방법간의 일치도를 알아보하고자 하였다. 평정척도와 자기보고식 검사는 모두 정신과환자의 증상평가를 위해 자주 이용되는 도구이다. 두 도구간의 상관관계가 높게 나타난다면, 시간절약을 위해 한 도구만을 사용하는 것이 권장될 수 있을 것이다. 그러나 그 상관관계가 유의미하지 않다면, 부득이 환자에 대한 이해를 높이기 위해 두 검사방법 모두를 사

용하는 것이 권장될 수밖에 없을 것이다. 평정척도인 BPRS와 자기보고식 검사인 Brief Symptom Inventory(이하 BSI; SCL-90-R의 단축형)를 비교한 결과를 보면(Morlan & Tan, 1998), BPRS의 총점과 BSI의 GSI, PST, PSDI의 상관이 각각 $r=.55$, $p<.05$; $r=.56$, $p<.005$; $r=.50$, $p<.01$ 로 통계적으로 유의미했다. 또, 개별문항의 경우를 보면, 우울 척도간에는 $r=.69$, $p<.0005$, 불안 척도간에는 $r=.46$, $p<.02$, 적대감 척도간에는 $r=.49$, $p<.01$ 로 통계적으로 의미있는 관계가 나타난 반면, 신체화, 정신증, 환각행동, 의심 척도간에는 상관관계가 의미가 없었다. Moses(1989)의 MMPI 개별척도와 BPRS 문항간의 상관연구에서는 F척도와 환각행동이 $r=.41$, $p<.05$, 반사회성 척도와 적대 문항이 $r=.50$, $p<.05$ 의 상관을 나타내었고, 다른 문항간에는 상관관계가 의미가 없었다. 이처럼 측정방법을 달리했을 때의 상관관계는 기대만큼 높지 않다. 국내에서는 BPRS와 자기보고식 검사와의 관계에 대한 연구가 전무하기 때문에, 본 연구에서는 이를 조사하였다.

연구방법

예비연구

예비연구는 정신과 입원환자 3명을 대상으로 이루어졌다. 평정자는 본 연구자를 포함하여 3명이었고, 평정이 이루어진 직후 각 문항값을 점검하고 2점 이상 차이를 보인 경우 평정자 3명이 함께 토론하여 해당 문항에 대한 개념적 이해의 차이를 좁혔다. 예비연구에서 평정자간에 3점 이상 차이를 보인 문항은 개념적 혼란, 반복운동과 자세, 무딘 감정이었다.

본연구

평정자

BPRS는 원래 경험이 풍부한 숙련된 평정자가 이용하도록 고안된 것이다. 그러나 임상 실제에서는 경험수준이 다양한 평정자들이 평가하며 수련중인 평정자들도 평가에 참여하는 사례가 많음을 감안하여, 본 연구에서는 훈련경험이 상이한 평정자 3명이 참여하였다. 첫번째 평정자는 임상심리전문가로 8년간의 임상경험을 가지고 있으며, 두 번째와 세 번째 평정자는 임상심리학 석사학위를 취득한 후 10개월간의 임상경험을 가지고 있다.

도구

원래의 BPRS(Overall & Gorham, 1962)에 Gabbard 등(1987)의 7점 척도 중 1, 3, 5점에 구체적 설명을 추가한 BPRS-A를 김민경(1993)이 번안한 것을 이용하였다. Gabbard 등(1987)이 개발한 것은 BPRS가 소개된 이후 만들어진 수정된 BPRS 중의 하나이다. 평정방식은 0-6점의 Likert 식으로 되어 있다. 척도에 포함된 18개 문항 중 “신체적 관심, 불안, 개념적 혼란, 죄책감, 과대망상, 우울한 기분, 적대감, 의심, 환각행동, 비정상적인 사고내용, 지남력 상실” 등 11개 문항은 환자의 언어적 보고만을 근거로 평가해야 하며, “감정철회, 긴장, 버릇과 자세, 운동지체, 비협조, 무딘 감정, 흥분” 등의 7개 문항은 행동관찰을 토대로 평가해야 한다.

환자의 반응을 유도하기 위한 반구조화질문에는 Crippa 등(2001)이 소개한 “BPRS 평가를 위한 구조화된 면접가이드”가 이용되었다. 구조화된 질문은 언어적 보고를 유도하기 위한 것으로 신체적 관심, 불안, 죄책감, 우울한 기분, 의심, 적대감, 환각행동, 지남력을 알아보기 위한 45개의

질문으로 구성되어 있다.

자기보고식 검사로는 SCL-90-R이 이용되었다. SCL-90-R은 90문항으로 이루어진 5점 척도(0 '전혀 없다'에서 4 '아주 심하다'까지)로서, 신체화 강박증, 대인민감성, 우울, 불안, 적대감, 공포불안, 편집증, 정신증의 9개 증상을 평가하도록 되어 있으며, 전체 심도지수, 표출증상 합계, 표출 증상 심도지수 등 3개의 전체지표가 산출된다. 본 연구에서는 김광일, 김재환, 원호택(1984)이 표준화한 것을 사용하였다.

평정방법 및 절차

평정자는 2명과 1명(본 연구자)으로 구분하여, 각각 짝진면접과 개별면접을 실시하였다. 평정시에는 면접가이드를 이용하되 그 시간은 20분에서 30분 이내에 마치도록 하였다. 짝진면접과 개별면접간의 평정간격은 기본적으로 만 24시간을 넘지 않는 것을 원칙으로 하였다.

BPRS 면담의 구성 :

1. 라포 형성
관계수립 3-5분
2. 비지시적 방법
자발적인 반응을 얻고 행동관찰 10분
3. 구조화된 직접 질문
평가에 필요한 구체적 정보의 획득 5-10분

피험자

피험자는 인천의 00대학병원 정신과에 입원해 있는 환자 21명이었다. 진단별로는, 정신분열증 환자가 13명, 조울증환자 4명, 주요우울증 환자 4명으로, 첫 입원인 경우가 10명, 두 번째 입원인 경우는 6명, 3번째 입원인 경우가 4명, 네 번째 입원인 경우는 1명이었다. 성별은 남자 10명, 여자 11명이며, 이중 기혼 7명, 미혼 12명, 사별과

이혼이 각각 1명이었다. 평균학력은 12.43년(s.d.=3.36년), 평균 연령은 34.90세(s.d.=8.74세)이었다.

자료분석

각 문항에 대한 평정치의 평균과 표준편차를 산출한 후, 이를 토대로 평정자간의 신뢰도를 분석하는 데에는 피어슨 상관분석법이 이용되었으며, 평가척도와 자기보고식 검사와의 관계도 피어슨 상관관계로 분석되었다. 이 모든 분석에는 SPSS/PC 10.0 이 이용되었다

결 과

신뢰도

BPRS 18개 문항에 대한 각 평정자의 평정결과가 표 1에 제시되어 있고, 평정자간 신뢰도는 표 2에 제시되어 있다. 각 문항의 신뢰도 범위는 .23부터 .87까지 다양하며, 짝진면접을 실시한 평가자 2와 3의 신뢰도가 개별면접자인 1과의 신뢰도에 비해 매우 안정적임을 보여주고 있다.

Flemenbaum과 Zimmermann(1973)에서도 짝진면접의 신뢰도는 .70에서 .90의 범위로 안정적인 반면, 시간대를 달리하여 평가한 개별면접자간의 신뢰도는 .50이하로 낮음이 보고되고 있는 바, 결국 평정자간 차이를 줄이기 위해 구조화된 질문과 척도수준에 대한 구체적 설명을 추가한 수정된 BPRS를 사용한 경우라도 시간과 면접상황에 따른 차이를 줄이는 것은 쉽지 않음이 시사된다. 이러한 차이를 유발할 수 있는 요인에는 평정자의 의사소통기술이나 대인관계기술, 전문적 경험, 증상에 대한 개념적 이해 수준 등 평가자 문제와 시간에 따른 환자의 상태변화 등 다

표 1. 평정자별 평가점수의 평균과 표준편차

각 문항	평정자 1		평정자 2		평정자 3	
	평균	표준편차	평균	표준편차	평균	표준편차
1. 신체적 관심	2.29	.78	2.24	1.04	2.43	.93
2. 불안	2.86	.79	2.38	.92	2.52	.87
3. 감정절회	2.52	1.44	2.14	1.31	1.95	.67
4. 개념적 혼란	1.90	1.26	2.05	.86	1.81	.81
5. 죄책감	3.19	1.25	2.00	.89	2.19	.93
6. 긴장	2.19	.98	1.81	.60	1.95	.59
7. 반복운동과자세	1.57	.68	1.43	.68	1.38	.59
8. 과대망상	1.33	.65	1.14	.36	1.09	.44
9. 우울한 기분	2.38	.74	2.05	.74	2.43	1.21
10. 적대감	2.48	1.08	2.24	1.26	1.86	1.01
11. 의심	2.19	1.21	2.24	1.55	2.33	1.65
12. 환각행동	1.81	1.12	1.95	1.47	1.95	1.50
13. 운동지체	2.43	1.07	1.52	.60	1.71	.78
14. 비협조	1.95	1.07	2.09	1.18	1.33	.57
15. 비정상적 사고	2.67	1.28	1.90	1.18	1.90	1.14
16. 무딘감정	2.52	1.40	2.00	1.09	1.95	.80
17. 흥분	1.86	.79	1.62	.67	1.48	.68
18. 지남력 손상	1.19	.51	1.43	.59	1.62	.59
총 점	2.19	.41	1.90	.41	1.82	.32

양할 것이다. 따라서 평정오류를 줄이기 위해서 개념에 대한 정확한 이해와 숙련된 면접기술도 필요하겠지만, 1명의 평정자에 의한 평가방법보다는 2명 이상이 짝편 면접을 통해 환자상태에 대한 정보를 얻는 것이 실제 임상상황에서도 권고될 필요가 있어 보인다.

반복운동과 자세문항이나 운동지체, 긴장, 무딘 감정 등 행동관찰을 근거로 평가해야 하는 문항들간에 신뢰도가 언어적 보고에 근거한 것보다 대체로 낮은 것은 행동관찰에 근거한 문항의 신뢰도가 낮게 산출된다는 선행연구의 결과와 일치한다.

BPRS와 자기보고식 검사(SCL-90)와의 상관

상당한 측정방법에도 유사한 결과가 산출될 수 있다면, 환자의 증상평가를 위해 여러 검사를 사용하는데 따른 비용과 시간을 절약할 수 있을 것이다. BPRS와 측정기법을 달리하는 자기보고식 검사와의 관계에 대한 연구가 국내에는 전무함으로 본 연구에서는 임상장면이나 연구장면 모두에서 자주 사용되는 SCL-90-R(김광일, 1984)과 수정된 BPRS 유사문항간의 관계를 알아보았다. <표 3>에는 수정된 BPRS와 SCL-90-R의 유사문항간에 상관관계가 제시되어 있다. BPRS의 불안

표 2. 평정자간 신뢰도

개별문항	신뢰도범위	평정자 1.2	평정자 1.3	평정자 2.3
· 1. 신체적관심	.28-.77	.28	.31	.77***
· 2. 불안	.23-.61	.23	.27	.61**
· 3. 감정절회	.55-.70	.70**	.55	.63**
· 4. 개념적혼란	.40-.46	.46*	.40	.44
· 5. 죄책감	.36-.72	.36	.44	.72***
· 6. 긴장	.23-.54	.23	.45*	.54*
· 7. 반복운동과자세	.06-.53	.53*	.06	.20
· 8. 과대망상	.42-.55	.42	.58**	.55**
· 9. 우울한기분	.42-.76	.60**	.42	.76***
· 10. 적대감	.53-.77	.61**	.53*	.77***
· 11. 의심	.34-.87	.43	.34	.87***
· 12. 환각행동	.36-.86	.36	.38	.86***
· 13. 운동지체	.23-.49	.49*	.39	.23
· 14. 비협조	.43-.80	.80***	.43	.47*
· 15. 비정상적인사고	.32-.68	.68**	.32	.52*
· 16. 무딘감정	.29-.74	.36	.29	.74***
· 17. 흥분	.50-.75	.55*	.50*	.75***
· 18. 지남력손상	.25-.63	.54*	.25	.63**
병리의 총점	.39-.78	.58**	.39	.78***

p<.05* p<.01** p<.001***

· 문항은 언어적 보고에 근거해 평정되는 문항임.

표 3. 유사문항간의 상관

SCL-90-R 소검사	BPRS 소검사				
	SOM	ANX	HOS	DEP	SUS
SOM	.361				
ANX	.243	.620**			
HOS	.107	.371	.358		
DEP	.122	.561**	.477*	.102	
PAR	.058	.504*	.600**	-.103	.537*
I-S	-.060	.526	.431*	.114	.500*
PSY	.317	.610**	.658**	-.157	.480*

* p<0.05 **p<0.01 ***p<0.001

SOM=신체화; ANX=불안; DEP=우울; HOS=적대감; SUS=의심; I-S=대인민감성; PAR=편집증; PSY=정신증.

문항과 통계적으로 유의미한 정적상관을 보인 SCL-90-R의 척도는 불안($r=.620$, $p=.002$), 우울($r=.561$, $p=.007$), 편집증($r=.504$, $p=.017$), 대인민감성($r=.526$, $p=.012$), 정신증($r=.610$, $p=.003$)이었으며, BPRS의 적대감문항과 정적상관을 보인 SCL-90-R의 척도는 편집증($r=.600$, $p=.003$), 대인민감성($r=.431$, $p=.045$), 정신증($r=.658$, $p=.001$)이었으며, BPRS의 의심문항과 정적상관을 보인 SCL-90-R의 척도는 편집증($r=.537$, $p=.010$), 대인민감성($r=.500$, $p=.018$), 정신증($r=.480$, $p=.024$)인 반면, BPRS의 신체화와 우울문항은 SCL-90-R의 어떤 척도와도 유의미한 상관을 보이지 않았다. 이들 결과는 BPRS와 MMPI 유사문항과의 관계를 연구한 Faustman 등(1989)에서 BPRS의 환각행동과 MMPI의 F척도, BPRS의 적대감과 MMPI의 사회병질 척도만이 상관이 의미있던 것에 비해, 의미있는 관계성이 산출된 것이며, BSI(Brief Symptom Inventory, SCL-90-R의 단축형)와의 관계를 연구한 Morlan & Tan(1998)와도 유사한 결과가 산출되었다고 볼 수 있지만, BPRS의 우울한 기분 문항과 SCL-90-R 우울 척도는 의미있는 상관을 보이지 않았다. 이는 BPRS에서의 우울한 기분에 대한 개념적 정의와 SCL-90-R의 우울 척도에 포함되는 문항속성에 차이가 있기 때문일 수 있다. 예를 들어, BPRS의 우울한 기분에서 중등도에 해당되는 수준은 “슬프거나 우울한 감정에 대해 이야기하는데, 가끔씩은 울기도 하며 면접 시간 내내 슬픔을 억누르는 상태를 보인다”로 SCL-90-R이 환자 자신의 주관적 지각만을 평가대상으로 하는 것과는 차이가 있을 수 있겠다.

본 연구결과에서는 BPRS의 불안문항이 우울문항에 비해 자기보고식 검사에서 환자 스스로 인식하는 부적응적인 측면과 높게 상관됨을 보이고 있고, 적대감이나 의심문항 또한 신체화나 우울보다 환자 스스로 느끼는 심리 내적 부적절감

과 높게 상관됨을 보이고 있다. 그러나 일정 정도 두 평가결과간에 상관을 보이더라도 자기보고식 검사의 결과와 평정척도의 결과는 상호보완적으로 이용되어야 할 것인데, 자기보고식 검사의 경우 환자가 사회적 바람직성에 따라 답하려는 경향이나 지나친 방어경향, 자기인식의 결여 등으로 결과가 왜곡될 수 있고 이것은 임상가의 직접관찰과 면접에 의한 평가로 보완할 수 있을 것이라 기대되며, 환자가 반응하기를 꺼리는 부분이나 행동적으로 표현되는 부분에 대한 정보를 얻기 위해서는 직접평가가 더욱 유용할 수 있기 때문이다.

논 의

BPRS는 전문적으로 훈련된 임상가가 정신증 환자의 행동관찰과 면접결과를 토대로 평가하는 평정도구로 1962년 개발된 이후(Overall, 1962), 임상장면이나 연구장면 모두에서 자주 사용되었으며, 국내에서도 PANSS와 함께 자주 이용되는 평정도구중 하나이다. 자주 이용되는 주된 이유는 정신증 환자가 나타내는 증상 대부분을 포괄하고 있고, 전문가들의 경험에 기초해 비교적 쉽게 평가할 수 있으며, 그동안 축적된 연구문헌이 많은 것 등을 들 수 있다. 그러나 국내에서는 BPRS의 신뢰도와 타당도에 대한 연구가 전무한 상태인 바, 김민경(1993)은 Gabbard 등(1986)의 수정된 BPRS와 초기의 BPRS(Overall 등 1962)간의 평정자간 신뢰도를 비교연구하여 수정된 BPRS를 사용한 경우 각 문항의 신뢰도가 증가됨을 밝힌 바 있다. 수정된 BPRS가 개발되는 이유는 평정자간의 신뢰도를 높이기 위함인데, BPRS는 개별문항 하나하나가 증상을 대표하고 있기 때문에 각 문항에 대한 개념적이며 조작적인 정의가 관련

전문가들이 의사소통할 수 있는 수준에서 이루어져야 하며, 증상의 심각도에 대한 정의 또한 행동적인 용어로 기술될 필요가 있다. 이외에 언어적인 보고에 근거해 평가되는 문항에 대한 정보를 얻기 위해 구조화된 질문을 사용하는 것도 평정자간 신뢰도를 높이기 위한 방법이 될 수 있다. 김민경(1993)의 결과에서도 나타났듯 신뢰도를 높이기 위해 수정된 방법을 사용한 선행연구들은 원래의 BPRS에 비해 평정자간 신뢰도를 높인다. 하지만 신뢰도를 높이기 위한 여러 방법의 개발에도 불구하고 실제 평가에서 평정자와 환자간의 상호작용에 따른 오차를 줄이는 것은 쉽지 않아 보인다. 숙련자들간의 평정결과가 비숙련자들간의 평정결과에 비해 신뢰롭다는 보고가 있긴 하지만, 대부분의 실제 임상장면에서는 숙련자보다 비숙련자가 환자를 평가하는 경우가 많다. 또한 신뢰도 연구를 위해 이용되는 비디오 테이프면접이나 짝진면접 등도 비용이나 시간적 제약으로 인해 잘 이용되지 않으며, 주로 한 연구자가 시간간격을 두고 반복측정하는 개별면접법이 이용된다. 따라서 선행연구의 결과들을 실제상황에 일반화하는데는 문제가 있어 보인다. 본 연구에서는 첫째, 실제 장면에서 흔히 이용되는 개별면접과 연구시 흔히 이용되는 짝진면접간의 신뢰도차이의 비교를 통해 면접상황에 따른 평가오류를 알아보고자 하였고 둘째, 측정방법이 상이한 자기보고식 검사와의 관계를 알아보고자 하였다.

연구결과, 선행연구에서처럼 행동관찰을 근거로 평정되는 반복운동과 자세, 운동지체, 긴장 문항의 신뢰도가 상대적으로 낮았으며, 언어적 보고를 근거로 평정되는 개념적 혼란이나 비정상적인 사고내용 등의 평정자간 신뢰도도 낮게 산출되었다. 이는 BPRS가 환자의 현상적인 상태를 평정하는데는 유용할 수 있으나 짧은 시간의

면접과 행동관찰을 통해 잘 드러나지 않는 증상을 평정하기에는 부족할 수 있음을 시사하는 것으로 Krawiecka, Goldberg 그리고 Vaughan(1977)이 지적했던 것처럼 환자의 변화를 평가하기에 민감하지 않을 수 있고, Manchanda 등(1986)의 보고처럼 증상에 익숙해져 있는 만성 환자의 평가에는 적절하지 않을 수 있겠다. BPRS의 18문항 중 11개 문항이 환자의 언어적 보고를 근거로 평정하게 되어 있는데 환자가 증상에 대한 방어를 계속한다면 정해진 짧은 시간안에 평가를 위한 신뢰로운 정보를 얻지 못할 가능성이 있다. 또한 연구결과 짝진면접자간의 신뢰도가 개별 면접자와의 신뢰도보다 대체로 높았는데, 이는 평정자와 평정상황의 상호작용 효과가 높음을 의미한다. 김민경(1993)은 BPRS의 평정결과에 대한 일반화가능도 연구에서 평정자에 따라 환자의 증상정도가 다르게 해석될 가능성이 높은 문항으로 “신체적 관심, 죄책감, 긴장, 적대감, 비협조, 흥분” 등을 제시하였다.

다음으로 평가방법이 상이한 자기보고식 검사와의 상관연구에는 BPRS(Brief Psychiatric Rating Scale)의 18개 문항 중 5개 문항과 SCL-90-R의 12 척도 중 7척도가 이용되었다. 연구결과, BPRS의 불안문항과 통계적으로 유의미한 정적상관을 보인 SCL-90-R의 척도는 불안, 우울, 편집증, 대인민감성, 정신증이었으며, BPRS의 적대감 문항과 정적 상관을 보인 SCL-90-R의 척도는 편집증, 대인민감성, 정신증이었고 BPRS의 의심 문항과 정적 상관을 보인 SCL-90-R의 척도는 편집증, 대인민감성, 정신증인 반면, BPRS의 신체화와 우울 문항은 SCL-90-R의 어떤 척도와도 유의미한 상관을 보이지 않았다. 연구자의 판단으로 이는 불안, 적대감, 의심 등의 BPRS 평가문항이 우울 문항보다 환자가 주관적으로 지각하는 부적응의 심각성과 더욱 정적으로 상관되기 때문일 수 있고,

다른 한편으로 불안이나 적대감 등을 표면증상으로 우울이 가려져 있기 때문일 수도 있어 보인다. Morlan와 Tan(1998)의 연구에서는 BPRS의 불안 문항보다 우울 문항이 자기보고식 검사의 증상들과 더욱 상관됨을 참고할 때, 우리나라 사람의 경우 서구보다는 자신의 내적 감정상태에 민감하지 못하고, 막연한 불안이나 적대감 분노, 불신적 태도로 내적 부적응감을 표현할 가능성이 높아 보인다.

결론적으로, 구조화된 질문지를 이용하거나 증상의 심각도에 대해 구체적 설명을 덧붙인 경우라도 개별면접의 신뢰도는 짝진면접에 비해 의미있게 낮았다. 이는 시간의 경과에 따른 환자 상태의 변화를 민감하게 평가했기 때문일 수도 있지만, 평가시간이 24시간을 넘지 않았고, 증상 변화가 급격히 이루어지지 않는 안정기에 접어든 입원환자가 평가대상이었던 만큼 환자의 증상 변화가 평가결과에 영향을 미쳤을 가능성은 적어 보인다. 결국, 수정된 BPRS라도 평정상황이나 평정자에 따른 차이를 통제하는 데는 일정 정도 한계가 있는 만큼 두 명 이상의 평정자가 짝진면접을 통해 환자를 평가하는 것이 실제상황에서도 권고되어야 할 것 같다. 만약 두 명이 평정에 참여할 수 없는 부득이한 경우라면, 한 명의 평정자가 일관되게 환자를 평정하는 것이 환자의 상태변화를 이해하는데 유용할 것이다. 측정방법을 달리하는 자기보고식 검사와의 상관연구에서는 평정자의 평가에 기초한 불안, 적대감, 의심 문항이 환자가 자기보고식으로 평가한 불안, 우울, 편집증, 대인민감성, 정신증과 의미있는 정적 상관을 보인 반면, 신체화나 우울 문항은 환자의 자기보고와 상관되지 않았다. 이는 환자가 면담시 우선적으로 보고하는 내용이 우울보다는 불안이나 화남, 분노, 의심 등이기 때문인 것 같다.

본 연구의 한계점으로는 개별면접과 짝진면접의 신뢰도 비교시 짝진면접자 두 명은 임상경험이 적은 비숙련자인 반면, 개별면접을 실시한 한 명은 상대적으로 임상경험이 풍부한 숙련자로 평정자간 신뢰도 차이가 면접방법의 차이이기 이전에 평정자 수준의 차이였을 가능성이 고려된다는 점이다. 현실적인 제약이 있긴 하나 “숙련자, 숙련자”, “비숙련자, 비숙련자”, “숙련자, 비숙련자”로 짝지어 평가했다면 신뢰도의 차이가 임상경험에 따른 차이인지, 면접상황이 상이한데 따른 문제인지 보다 명확히 알 수 있었을 것이다. 하지만 이와 같은 방법은 한 환자에게 같은 질문들을 짧은 시간간격을 두고 세 번 반복해야 하는 문제가 있고, 이에 따른 현실적 제약으로 비디오테이프 평가법을 이용하는 것 또한 일반화의 문제가 있으므로 어떤 방법을 사용하든 평정자간 신뢰도연구에는 한계가 있을 것으로 보인다. 또한 BPRS의 구성문항들이 정신증 환자의 증상들을 대부분 포괄하고 있긴 하지만, 비협조나 의심 문항의 경우 정의가 겹치는 부분이 있고, 불안과 긴장 문항, 감정철회와 운동지체 문항, 비협조와 적대감 문항의 경우는 증상이 공변하는 경향이 있어 평정자가 개별문항에 대해 명확히 구분하여 평가하는 것이 어려울 수 있다. 따라서 증상에 대한 개념적 정의를 보다 구체화하기 위한 연구가 이루어질 필요가 있어 보인다.

연구결과를 토대로 BPRS 사용시 주의할 점을 제안해 보면, BPRS의 각 문항은 평정자의 이해수준에 따라 달리 평가될 가능성 높은 바, 각 문항에 대한 개념적 이해가 명확해야 할 것으로 보이며, 이를 위해서는 구체적인 훈련기간을 제시할 수는 없으나 일정 정도 임상경험이 있는 숙련자가 평가할 필요가 있다는 것이다. 또한 두 명 이상이 환자 평가에 일관되게 관여할 것이

권고되지만, 부득이한 경우라면 한 명이 일관되게 한 환자를 평가해야 증상변화에 대해 신뢰로운 정보를 얻을 수 있을 것으로 보인다.

본 연구의 의의로는 국내에서 시도된 적이 없는 BPRS의 개별면접과 짝진면접의 평정자간 신뢰도를 알아본 것이며, 측정방법이 상이한 자기보고식 검사와의 관계연구를 통해 평가자에 의한 평가와 환자의 자기보고에 의한 평가간의 일치도를 알아본 것이다.

참고문헌

- 김민경 (1993). 정신분열증 환자에 대한 Brief Psychiatric Rating Scale의 평정자간 신뢰도 연구. 충북대학교 대학원 석사학위청구논문.
- Anderson, J., Larsen, J. K. Schultz, V. et al. (1989). The Brief Psychiatric Rating Scale. Demension of Schizophrenia-reliability and construct validity. *Psychopathology*, 22, 168-176.
- Faustman, W. O. (1994). Brief Psychiatric Rating Scale. In: Maruish ME, ed. *The use of Psychological Testing for treatment planning and outcome assessment*. New Jersey: Hillsdale. 371-401.
- Faustman, W. O., Moses, J. A., Csernansky, J. G., & White, P. A. (1989). Correlations Between the MMPI and the Brief Psychiatric Rating Scale in Schizophrenic and Schizoaffective Patients, *Psychiatric Research*, 28, 135-143.
- Flemenbaum, A. & Zimmermann, R. L. (1973). Inter- and Intra-Rater Reliability of the Brief Psychiatric Rating Scale. *Psychological Report*, 32, 783-792.
- Grippa, J. A. S., Sanches, R. F., Loureiro, S. R., Zuardi, A. W. (2001). A Structured Interview guide increases Brief Psychiatric Rating Scale reliability in raters with low clinical experience. *Acta Psychiatrica Scand*, 103, 465-470.
- Hafkenscheid, A. (1991). Psychometric Evaluation of a standardized and expanded Brief Psychiatric Rating Scale. *Acta Psychiatrica Scandinavica*, 84(3), 294-300.
- Hafkenscheid, A. (2000). Psychometric measures of individual change: an empirical comparison with the Brief Psychiatric Rating Scale(BPRS). *Acta Psychiatrica Scand*, 101, 235-242.
- Heldung, J. L. & Vieweg, B. W. (1980). The Brief Psychiatric Rating Scale; A Comprehensive Review. *Journal of Operational Psychiatry*, 11, 48-65.
- Krawiecka, M., Goldberg, D., & Vaughan, M. (1977). A standardized Psychiatric assessment scale for rating chronic psychotic patients. *Acta Psychiatrica Scandinavica*, 55, 299-308.
- Lachar, D., Bailey, S. E., Rhoades, H. M. et al. (2001). New Scale for an Anchored Version of the Brief Psychiatric Rating Scale: Construction, Reliability, and Psychiatric in Acute Psychiatric Admissions. *Psychological Assessment*, 13(3), 384-395.
- Manchanda, R., Saupe, R. & Hirsch, S. R. (1986). Comparison between the Brief Psychiatric Rating Scale and the Manchester Scale for the rating of schizophrenic symptoms. *Acta Psychiatrica Scandinavica*, 74(6), 563-568.
- Morlan, K. K. and Tan, S-Y. (1998). Comparison of the Brief Psychiatric Rating Scale and the Brief Symptom Inventory, *Journal of Clinical Psychology*, 54(7), 885-894.
- Overall, J. E. (1976). The Brief Psychiatric Rating

- Scales. Psychological Report, 10, 799-812.
- Overall, J. E. (1988). The Brief Psychiatric Rating Scales; Recent development in ascertainment and scaling. Journal of Operational Psychiatry, 24, 97-100.
- Overall, J. E., Hollister, L. E., & Pichot, P. (1967). Major Psychiatric Disorders: A Four-Dimensional Model. Archives of General Psychiatry, 16(2), 146-151.
- Raskin, A. (1988). Discussion: recent developments in ascertainment and scaling of the BPRS. Psychopharmacol Bull, 24, 122-124.
- Rhoades, H. M. & Overall, J. E. (1988). The Semistructured BPRS Interview and Rating Guide. Journal of Operational Psychiatry, 24, 101-104.
- Woerner, M. G., Mannuzza, S. & Kane, J. M. (1988). Anchoring the BPRS: An aid to improved reliability. Psychopharmacol Bull, 24, 112-117.
- 원 고 접 수 일 : 2003. 3. 3
수정원고접수일 : 2003. 7. 1
게 재 결 정 일 : 2003. 7. 30

Reliability of Korean Brief Psychiatric Rating Scale(BPRS) - Comparison of interrater reliability between the two rating methods and correlation of BPRS and SCL-90 self-report test -

Min-Kyung Kim

Department of Neuropsychiatry
Our lady Mercy Hospital
the Catholic University of Korea

Bong-Keon Lee

Department of Psychology
Chungbuk National University

Yang-Whan Jeon

Department of Neuropsychiatry
Our lady Mercy Hospital
the Catholic University of Korea

This study purported to compare the inter-rater reliability of Korean Brief Psychiatry Rating Scale-Anchored Version(BPRS; Overall, 1962) between the joint interview and the separation interview and to study the correlation between BPRS and self-report test(SCL-90). The subjects consisted of 21 inpatients in University psychiatric hospital. In the joint interview, the two 10-month experienced clinical psychology trainees rated the BPRS 18 items. In the separation interview, another licenced clinical psychologist alone rated the same patients. Also SCL-90 scales were administered to the same patients. The scores of the clinician-rated BPRS were correlated with those of the SCL-90-R, a self-report measure. The results showed that joint interview produced higher reliability than separation interview. The analysis of correlations between the BPRS and SCL-90 indicated that Anxiety, Hostile, Suspicious items of the BPRS were significantly correlated with Anxiety, Hostile, Depression, Paranoid Ideation, Interpersonal Sensitivity, Psychoticism scales of the SCL-90-R respectively. The limitations and clinical implications of this study and the recommendations for the future study were suggested.

Keywords : Korean BPRS-anchored version, Korean SCL-90, inter-rater reliability, joint interview, separation interview, correlation study with self-report measure.