한국사회복지학
*Korean Journal of Social Welfare* vol. 46 2001. 9, 210~235


# 한국판 일반정신건강척도(KGHQ)의 개발에 관한 연구 I *

## KGHQ-20과 KGHQ-30을 중심으로

신 선 인 **

본 연구에서는 비정신의료전문가를 위한 정신건강척도로서 세계적으로 널리 쓰이고 있는 General Health Questionnaire의 한국판(KGHQ)을 표준화하여 개발함으로써 사회복지사들이 보다 쉽고 간편하게 클라이언트의 정신건강수준을 측정할 수 있도록 돕고자 하였다. 표준화과정은 GHQ 20문항형과(KGHQ-20) 30문항형(KGHQ-30)에 대한 번역작업과, 이를 대학생 및 대학원생 916명과 정신과병원에 입원중인 정신질환자 102명에 적용하여 실시한 설문조사과정, 그리고 이들의 응답자료를 이용하여 문항내용의 타당도와 신뢰도를 검증하는 과정으로 이루어졌다. 신뢰도 검증에 있어서 두 문항형 모두 검사-재검사 상관도가 높았으며 문항-총점간 상관도와 Chronbach's $\alpha$ 값에 의한 내적 합치도도 높게 나타났다. 타당도 검증은 집단간의 평균점수 비교검증과 기준관련 척도로 설정된 SCL-90, CES-D, SDS, BAI, STAI 등 유사도구들과의 척도간 상관도, 민감도(*sensitivity*)와 명확도(*specificity*) 등에 준하였는데 두 문항형 모두 높은 타당도를 나타내었다. 정신질환 발병의 위험집단을 선별하기 위한 기준점인 분할점수는 KGHQ-30은 16점, KGHQ-20은 13점에서 각각 설정하였다.

---

* 이 연구는 2000학년도 대구대학교 학술연구비 지원으로 수행되었음.
** 대구대학교 사회복지학과 조교수.

## 1. 서론

최근 우리나라의 정신보건사업은 주요 서비스 대상자인 만성정신질환자들을 위한 치료와 재활뿐 아니라 지역사회 내의 일반 주민들을 위한 정신건강증진 및 예방 프로그램의 개발을 지향하기 시작했으나 이를 위한 기초자료는 매우 부족한 실정이다. 즉 심리상담이나 위기개입 등의 단기적 처치를 필요로 하는 정신질환 발병 위험 군에 속하는 사람들의 수나 문제의 유형도 제대로 파악되어 있지 않으며, 만성정신질환자의 수도 의료보험 및 의료보호 급여자료와 정신요양원 재원자를 근거로 한 유병률에 의존하고 있을 뿐 아직 정확하게 측정하지 못하고 있는 실정이다(남정자, 1997). 이는 신체적 질환과 달리 정신질환은 내담자의 주관적인 경험에 대한 보고에 근거하므로 객관적인 진단이 쉽지 않기 때문이며, 특히 정신의료기관의 서비스를 받은 적이 없는 지역사회 내의 주민들을 고려할 때 이들의 정신건강수준을 간편하게 측정하여 정신질환 발병의 위험 군에 속해 있는 사람들을 가려낼 수 있는 적절한 선별도구가 미비하기 때문이다.

이러한 실정을 감안하여 현재 보건복지부는 WHO가 개발한 CIDI(The Composite International Diagnostic Interview)를 표준화하여 전국실태조사를 실시할 계획이다(김덕중, 2001). CIDI는 ICD-10과 DSM-IV의 정신질환 진단분류에 근거하여 개발된 설문문항에 응답자가 직접 답하게 함으로써 진단과정에서 평가자의 주관적 개입을 배제할 수 있도록 고안된 도구이며(www.who.int/msa/cidi), 필기용뿐 아니라 컴퓨터용으로도 개발되어 있어서 매우 유용한 도구임이 분명하다. 그러나 각 진단분류별 문항수 및 전체 문항수가 매우 많아서 지역사회에서 다수의 인구를 대상으로 이를 적용하는 데는 단계적인 전략이 필요하다고 본다. 즉 이러한 구체적이고 임상적 진단도구를 사용하기 이전에, 타당도 높은 간이검사도구를 활용하여 보다 간단하게 문제를 가진 대상을 일차적으로 걸러낸다면 광범위한 지역조사를 실시할 경우에 인력과 비용이 절감될 수 있을 것이다. 특히 지역사회에서 활동하는 사회복지사와 같은 비의료계 전문가(*non-clinician*)들로서는 이러한 과정에 가장 밀접하게 관여하게 되는데, 간이검사도구의 활용을 통하여 그들이 접하는 클라이언트 중에서 정신질환의 위험이 있는 자들을 일차적으로 선별하여 필요한 서비스를 제공함으로써 전체적인 서비스의 질도 높일

수 있을 것이다. 이는 사회복지사들이 접하는 클라이언트들의 다수가 정신적인 문제를 겪고 있으며 긴급한 처치를 요하는 경우도 빈번하기 때문이다.

이러한 목적에 부합되는 간이정신건강수준 측정도구로는 80년대 초부터 우리나라 심리학분야에서 널리 쓰여온 간이정신진단검사(SCL-90)가 있으나, 이는 문항수가 90문항으로 지역사회 내에서 바쁘게 활동하는 사회복지사 등이 간편하게 사용하기에는 너무 많은 편이다. 그 외에는 우울증, 불안증, 대인관계척도 등의 문제특성별 측정도구들이 자주 쓰이고 있으며, 비의료계전문가를 위한 신뢰성있고 종합적인 정신건강 간이측정도구에 관해서는 별로 알려진 바가 없다.

따라서 본 연구에서는 그동안 영국을 비롯한 세계 각국에서 널리 사용되어 온 간편한 정신건강수준 측정도구인 The General Health Questionnaire(GHQ)의 한국판 표준화작업을 통하여 우리나라 지역사회주민의 정신건강수준 측정을 위해 활용될 수 있는 유용한 기초도구를 개발하고자 하였다. 이는 무엇보다도 정신보건사회복지사를 비롯한 사회복지사들이 이를 쉽게 적용하여 현장에서 접하는 클라이언트들의 정신건강수준을 간편하게 측정한 후 문제성이 높은 대상자들을 가려내어 관련 치료기관으로 연계하는 등 적절한 정신건강 서비스를 제공할 수 있도록 돕는 데 그 목적이 있다.

본 연구의 구체적인 내용은 첫째, 20문항형과 30문항형의 GHQ를 한글로 번역한 후 검사-재검사 방법과 문항 간 내적 합치도 측정에 근거하여 이들의 신뢰도를 측정하였으며, 둘째, 집단간의 평균차이, 척도간 상관도, 민감성계수(*specificity*) 및 명확성계수(*speciality*)를 측정함으로써 타당도를 검증하였고, 셋째, 정신질환의 위험 군을 선별하기 위한 기준이 되는 분할점을 선정하여 제시하였다.

## 2. 표준화에 관한 선행연구

The General Health Questionnaire(GHQ)를 한글로 직역한 '일반건강설문지'라는 표현은 이 도구의 목적 및 내용과 일치하지 않으므로, 본 연구에서는 GHQ를 '일반정신건강척도'라는 용어로 표현하고 한글 번역판의 영문은 'the Korean General Health Questionnaire'로 하였으며 약자는 KGHQ로 표기하였다.

GHQ는 1972년 영국에서 개발되어 홍콩 등 주로 영국문화권에서 많이 쓰여왔으며

Goldberg & Williams(1991)에 보고에 의하면 1991년까지 중국과 일본을 포함한 약 38개 이상의 언어로 번역되어 사용되었고, 타당도 및 신뢰도에 관한 연구는 50건 이상이 수행되었는데, 10년이 지난 현재에는 그 숫자가 더욱 늘어났을 것이다. 우리나라에서는 1993년 한국보건사회연구원(남정자·최정수, 1993)에서 20문항형의 GHQ가 번역되어 사용된 적이 있으나 이 때는 본격적인 표준화 작업을 거친지 않은 채 사용되었다.

GHQ는 원래 일차적 진료기관이나 종합병원외래 등에서 비정신의료 전문가들로 하여금 이들이 접하는 환자들이 정신질환에 걸렸을 가능성을 쉽고 간편하게 파악할 수 있게 할 목적으로 1972년에 영국의 Goldberg가 개발한 자가기재방식의 검사도구이며, 현재는 사회복지사, 임상심리사와 같은 지역사회 내의 대인서비스 전문인력들이 정신질환자나 정신질환의 위험 군에 속하는 대상자들을 일차적으로 선별하는 도구로 널리 쓰이고 있다. 즉 GHQ는 정신질환자의 선별을 위한 첫 번째 단계의 평점도구로서, 이를 통해 정신질환자의 가능성을 가려낼 수 있도록 고안된 것이다(Benjamin et al., 1982; Shek & Tsang, 1995). GHQ는 그러나 정신분열증이나 조울증과 같은 만성정신질환의 유형을 찾아내거나 진단적 의도로 고안된 것은 아니며, 의료처치를 받고 있는 일반환자의 현재 건강상태에서 정신적인 문제와 관련된 요소를 찾아내고 이를 유형화시켜서 개발한 도구이다. 즉 GHQ는 응답자의 나쁜 건강상태의 원인이 되는 심리사회적 경험의 정도를 측정함으로써 정신질환의 가능성을 추정할 수 있도록 개발된 도구인데, 문항들의 하위요인을 분석함으로써 질환의 유형을 유추하는 것은 가능한 것으로 알려져 있다(Goldberg & Williams, 1991). GHQ는 응답자의 지난 2~3주 동안의 심리적 상태를 그 이전에 갖고 있던 응답자의 평소 심리상태와 비교함으로써 변화된 현재 상태의 문제점을 파악하도록 고안된 도구이며, 비교적 적은 문항수와 응답자가 쉽게 답할 수 있고 조사자의 주관적인 평가를 배제할 수 있는 객관적인 설문문항으로 구성되어 있다.

GHQ의 원문은 우울증, 불안증, 사회적 부적응, 신체화 증상, 불면증 등의 하위요인으로 나누어 질 수 있는 60문항으로 이루어져 있으나, 30문항, 28문항, 20문항 등으로 간소화하여 사용되는 경우가 더 많으며 현재까지는 30문항형이 가장 널리 사용되어 왔다. 28문항형은 그 하위요인이 각 7문항씩 4가지 요인으로(신체화증상, 불안증, 사회생활 부적응, 심한 우울증) 뚜렷하게 나누어지므로 이에 근거하여 응답자의 정

신질환의 유형을 일차적으로 추정하기 위한 목적으로 주로 쓰이고 있다(Goldberg & Williams, 1991). 특히 최근에는 2분 이내에 응답을 끝낼 수 있어서 더욱 간편하게 정신건강수준을 측정할 수 있도록 개발된 12문항형의 높은 타당도와 신뢰도도 보고되었다(Goldberg et al., 1997). 한편 선행연구들에서 나타난 각 문항형 간 신뢰도와 타당도의 평균값은 〈표 1〉과 같다.

GHQ의 타당도에 관한 많은 연구에서는 주로 다양한 정신병리심리척도들과 높은 척도간 상관도를 보임으로써 동시적 타당도(*concurrent validity*)를 입증하였으며, 요인구조의 분석을 통하여 정신건강수준 척도로서의 개념구성타당도(*construct validity*)를 입증하기도 하였고, 또한 측정하고자 하는 대상자를 얼마나 정확하게 파악할 수 있는지를 나타내는 타당도 계수인 민감도(*sensitivity*)와 명확도(*specificity*)의 수준을 검증함으로써 정신질환 이환 여부 측정을 위한 이 도구의 타당도가 입증되었다(Goldberg & Williams, 1991).

특히 GHQ는 WHO에 의해서 개발된 the Self Reporting Questionnaire(SRQ)와 정신질환의 사례를 찾아내는 면에 있어서 동등한 효과가 있다고 보고되었으며(Mary & Williams, 1985), 간이정신진단검사(the Symptom Checklist: SCL-90), Beck의 우울증척도(the Beck Depression Inventory: BDI), CES-D(the Center for Epidemiologic Studies Depression Scale) 등과 GHQ의 하위 요인인 우울, 불안 점수간에도 높은 척도간 상관성을 보인 것으로 알려져 있다(Goldberg & Williams, 1991; Shek, 1989). GHQ의 동시적 타당도에 관한 선행연구들은 또한 PSE(Present State Examination)를 사용하는 경우가 많았는데, PSE 역시 GHQ와 높은 척도간 상관(r=.73~92)을 나타

〈표 1〉 선행연구에 나타난 GHQ 신뢰도 / 타당도 계수의 평균값 · 범위

| 계수 / 문항형 | Chronbach's $\alpha$ | 반분법 | 검사-재검사 | 임상적 면접 진단과의 상관 | Specificity(%) | Sensitivity(%) |
|---|---|---|---|---|---|---|
| GHQ-60 | 0.93 | 0.95 | 0.76 | 0.72~0.81 | 87.8 | 95.7 |
| GHQ-30 | 0.87 | 0.92 | 0.77 | 0.45~0.77 | 87.0 | 91.4 |
| GHQ-28 | - | - | 0.90 | 0.67~0.83 | 82.0 | 86.0 |
| GHQ-20 | 0.79 | 0.90 | - | - | 86.0 | 88.2 |
| GHQ-12 | 0.85 | 0.83 | 0.73 | 0.53~0.71 | 78.5 | 93.5 |

내었다(Banks, 1983; Koeter et al., 1989; Vazquez-Barquero et al., 1986, etc.). GHQ는 또한 임상적 진단도구와의 관계에 있어서는 우리나라에서도 자주 쓰여온 CIS(Clinical Interview Schedule)와 높은 척도간 상관도(r=.72~.81)를 나타내었으며(Goldberg & Williams, 1991; Hobbs et al., 1983), 〈표 1〉에 소개된 바와 같이, 정신과의사의 면접에 의한 임상적 진단 결과와도 비교적 높은 상관도를 나타내었다. 또한 GHQ의 개발자인 Goldberg와 WHO의 공동연구에서는 CIDI의 의해 정신질환자로 선별된 케이스의 83%~95 %가 GHQ-12문항형에 의해서도 질환의 가능자로 판정되어서 높은 척도간 상관도가 있음이 보고되었다(Goldberg et al., 1997).

GHQ의 문항이 높은 신뢰도와 타당도를 갖는 점 외에도 그 내용이 일반 조사대상자들의 혐오감이나 거부감을 유발시키지 않을 수 있게 구성되었음을 감안할 때, 본 연구에서 사용한 KGHQ-20, KGHQ-30의 두 문항형 외에도 28문항형, 12문항형의 타당성이 검증되고, 각각의 적절한 분할점이 정해진다면, 이들은 정신건강수준을 쉽게 측정할 수 있는 도구로서 필요한 경우에 따라 매우 적절하고 유용하게 쓰일 수 있을 것이다. 본 연구에서는 인력 및 재정의 한계로 인해, 현재까지 가장 널리 쓰여왔던 30문항형과 1993년 한국보건사회연구원에서 사용된 바 있는 20문항형을 우선적으로 표준화하였으며, 초단축형인 KGHQ-12와 하위요인 분석용인 KGHQ-28에 관한 연구는 본 연구에 이은 다음 연구과제로 수행할 예정이다.

## 3. 연구방법

### 1) 조사대상 및 자료수집

KGHQ 두 문항형의 타당도 및 신뢰도를 조사하기 위하여 대학 및 대학원 석사과정 재학생으로 구성된 학생집단과 만성정신질환으로 진단받은 환자집단을 대상으로 자가기재를 중심으로 한 집단조사를 실시하였다. 본 연구의 목적이 응답자의 특성 파악에 있지 않고 측정도구의 타당성 및 신뢰성 검증에 있으므로 편의표집에 의하여 조사대상자를 모집하였다. 조사대상자는 서울과 대구지역의 2개 대학에 재학 중인 대학생 843명과 대학원생 73명, 그리고 대구, 경북지역 소재 4개 정신과병원의 입원

혹은 외래 환자 102명을 포함하여 총 1,018명이었다. 이 중에서 결측값이 있는 203명을 제외한 815명의 응답내용만 분석자료로 사용하였다. 이 유효응답자 중에서 대학 및 대학원생 501명은 KGHQ-20 분석을 위한 조사대상자였으며, 나머지 227명과 정신질환자 87명 전원은 KGHQ-30을 위한 조사대상자였다.

정신질환자 집단은 증상의 정도가 설문 내용을 충분히 인지하고 자가기재를 할 수 있는 수준의 환자로 선발하였는데, 필요한 경우에는 조사자인 사회복지사들이 환자들의 설문지 작성에 도움을 주었다. 환자의 67%는 정신분열증환자였으며 16%가 조울증을 포함한 우울증환자였고 나머지는 알코올중독, 기질성뇌장애 등이었다. 학생집단은 전체인원 중 44.2%(360명)가 남자였고 나머지(455명)는 여자였다. KGHQ-20에 관한 자료수집은 2000년 6월과 11월에 걸쳐 실시하였으며, KGHQ-30에 관한 조사는 2001년 5월중에 실시하였다.

2) 조사도구

(1) 한국판 일반정신건강척도 : KGHQ-20과 KGHQ-30의 문항내용은 〈부록 1〉과 같이 구성되어 있다. 모든 문항은 본 연구자가 한글로 번역하고 이를 이중언어 사용자 두 명에게 영어로 역 번역하게 하여 그 내용을 원문과 대비한 후, 이에 근거하여 번역한 내용을 수정·보완하는 과정을 거쳐서 번역된 내용과 원문과의 차이를 최소화하였다. 이를 다시 대부분이 사회복지분야의 현직에 종사하고 있는 사회복지학 석사과정 대학원생 34명을 대상으로 예비조사를 실시하여 각 설문 문항 및 응답항목의 내용 타당성을 확인한 후 수정·보완하여 사용하였다. 두 문항형에 공통적으로 포함된 20문항은 한국보건사회연구원에서 사용하였던 GHQ-20의(남정자·최정수, 1993) 문항들을 기초로 하여 재번역하여 구성하였고, 이에 10문항을 추가하여 KGHQ-30의 문항을 완성하였다.

또한 각 문항의 내용은 영어 원문에서의 응답항목 내용을 감안하여 작성되었다. 영어 원문에서는 4점 척도로 된 응답항목이 각 문항의 내용에 맞추어서 문항마다 조금씩 다르게 표현되어 있는데 이와 같은 구성을 한글판에도 그대로 적용하기에는 문제점이 있었다. 한글판에서는 응답항목을 문항마다 다르게 표현해 본 결과, 응답항목이 다양해져서 도구 전체가 상당히 복잡하게 되었다. 따라서 한글판에서는 응답항

목에 포함된 어의를 설문문항의 내용에 추가해서 표현하고 응답항목은 모든 문항에 공통적으로 적용될 수 있도록 '매우 그렇다', '그렇다', '아니다', '매우 아니다'의 4점 척도로 설정하였다. 예를 들면, 필요한 경우에는 문항내용에 "~보다 ~한 적이 많습니까?" 등의 양적 측정을 위한 표현을 추가하고, 응답항목인 "more than usual", "much more than usual" 등을 "그렇다", "매우 그렇다"로 나타내었다. 참고로 영어판 응답항목의 공통적인 주 내용은 "평소보다 좋음"(*better than usual*), "평소와 같음"(*same as usual*), "평소보다 못함"(*worse than usual*), "평소보다 훨씬 못함"(*much worse than usual*)이었다.

GHQ의 채점방법은 연구목적에 따라 다양하지만 정신질환 위험군의 선별이나 정신건강수준의 측정을 위해 가장 일반적으로 쓰이는 방법은 'GHQ scoring'이다. 이 경우, 4가지 응답항목에 대한 배점은 0-0-1-1이며 총점의 최대점수는 GHQ-20인 경우 20점, GHQ-30인 경우에는 30점이고 점수가 높을수록 정신건강수준이 낮음을 의미한다. 또한 요인분석에 의해 설정된 하위요인을 별도로 분석할 때는 'Likert scoring'이 많이 쓰이고 있으며, 0-1-2-3에 의한 배점으로 최대총점은 GHQ-20의 경우 80점, GHQ-30의 경우에 90점이다. 본 연구에서의 점수처리는 GHQ scoring (0-0-1-1)의 전통적인 방법을 사용하였으며 GHQ문항중의 일부는 다른 문항들과 역방향을 이루고 있어서 채점과정에서 역점수 처리가 필요하다.

한편 GHQ로 만성정신질환자들을 대상으로 측정할 때는 이 두 종류의 배점방식에 의한 변별력이 매우 낮은 것으로 알려져 있다. 즉 응답자가 지난 2~3주간에 변화한 심리적 상태를 그 이전의 평상시와 비교하여 답하게 함으로써 현재의 정신건강 상태를 추정하기 위한 문항내용을 만성질환자들에게 적용할 경우에는 그 변화의 차이가 뚜렷이 나타날 수 없을 것이므로 만성정신질환자집단과 일반인집단의 GHQ총점 평균의 차이가 유의미하지 않게 나타난다는 것이다(Koeter et al., 1989). 이를 위한 대안으로 만성정신질환자의 경우에는 Goodchild와 Duncan-Jones(1985)가 제안한 'CGHQ(Chronic GHQ) scoring'방법이 쓰이고 있다. 이 방법에서는 긍정적인 문항에 0-0-1-1을 배점하고, 부정적인 문항에는 0-1-1-1을 배점함으로써 만성환자들이 경험하고 있는 부정적인 정신심리 상태를 최대한으로 평가하도록 하였다. 본 연구에서도 정신질환자 집단에 'GHQ점수'를 적용하였을 때는 총점평균의 분포가 대학생집단과 유의미한 차이가 없었으므로 CGHQ방식을 도입하여 점수를 환산하였다.

(2) 척도간 상관을 위한 도구 : KGHQ의 동시적 타당도(*concurrent validity*) 조사를 위해 사용된 측정도구들은 외국의 선행연구에서 사용된 심리척도들 중에서 한글판으로 번역되어서 쓰이고 있는 도구들을 선정하였으며, 특히 설문지의 전체 문항수를 줄이기 위해서 문항수가 적고 사용방법이 간단한 도구들을 사용하였다. KGHQ-20문항형의 기준관련 척도로는 그 주요 하위요인인 불안증, 우울증과의 관계를 평가하기 위해 자기평가우울척도(SDS), CES-D, Beck의 불안질문지(BAI), 상태특성불안척도(STAI) 등 네 가지 척도를 사용하였으며, 정신질환자 집단이 조사대상자에 포함된 KGHQ-30문항형의 기준관련 척도로는 전체 문항수를 줄이기 위해서 간이정신진단검사(SCL-90)를 선택하였다. SCL-90은 정신과 환자들이 경험하는 일반적인 증상의 수준을 측정하여 정신질환의 이환 상태를 측정하는 총 90문항의 자기보고형 척도이며, 임상적 진단의 보조수단으로 주로 사용된다. 하위요인으로 불안, 신체화, 우울, 편집, 정신증, 예민성 등으로 나누어지고 이들은 각각 소척도로 사용되기도 한다. 점수가 높을수록 정신건강이 나쁜 상태를 의미하며 신뢰도수준은 조사방법 및 연구내용에 따라 다양하게 나타났으나 비교적 납득할 만한 수준을 나타내었으며(Chronbach's $\alpha$=.67~.89), 타당도도 높게 알려져 있다(고려대부설 행동과학연구소, 1999).

CES-D(Center for Epidemiologic Studies Depression Scale)는 일반인들이 경험하는 우울증상을 측정하기 위해 자주 사용되는 척도이며 총 25문항으로 구성되어 있다. 비교적 높은 신뢰도를 나타내고(Chronbach's $\alpha$=.89), SCL-90의 하위요인인 우울증척도점수와의 상관관계계수(r=.69)와 비교적 높은 공존타당도를 갖고 있는 것으로 알려졌다(전겸구·이민규, 1992). 자기평가우울척도(Self-Rating Depression Scale: SDS) 역시 일상생활에서 경험하는 우울한 증상에 대한 자기보고형 평가척도이다. 혼동, 공허감, 자기비하 등의 심리적 증상에 관한 문항들과 체중감소, 변비, 피로감 등의 신체화 증상에 관한 문항 등 총 20문항으로 구성되어 있으며 신뢰도는 Chronbach's $\alpha$=.86 정도로 높게 나타나 있다(고려대부설 행동과학연구소, 1999). Beck의 불안질문지(Beck Anxiety Inventory: BAI)는 개인이 최근 경험하고 있는 불안의 심각성 정도를 평가하기 위한 측정도구로서 총21문항으로 구성되어 있으며, 높은 신뢰도(Chronbach's $\alpha$=.91)를 갖고 있는 것으로 알려졌다. 상태-특성불안척도(State-Trait Anxiety Inventory: STAI)는 정신장애가 없는 정상적인 성인의 불안상태를 측정하는 도구로 개발되었으나 정신과 환자의 불안상태나 불안증을 나타내는 집단을 판별해주는 임상적

도구로도 사용되고 있다. 총 40문항으로 구성되어 있으며 비교적 높은 신뢰도 (Chronbach's $\alpha$= .89)와 타당도를 나타낸 것으로 알려져 있다(고려대부설 행동과학연구소, 1999).

### 3) 분석방법

전체 조사대상자의 응답자료 중에서 분석에 사용된 자료는 대학생 및 대학원생과 정신질환자의 응답지 825부였다. KGHQ의 신뢰도 검증을 위해서는 내적 합치도 계수(Chronbach's $\alpha$)와 문항 - 총점간 상관계수 및 검사 - 재검사 신뢰도를 측정하였다. 타당도 검증은 네 가지 각도에서 실시하였는데 우선 대학원생 34명을 대상으로 예비조사를 실시하여 각 설문 문항 및 응답항목의 내용적 타당성(*content validity*)을 확인하였고, 도구가 측정하고자 하는 개념을 정확하게 측정하는 정도를 나타내는 개념구성타당도(*construct validity*)의 검증을 위하여 건강한 정신건강수준의 대학생집단과 만성정신질환자 간의 KGHQ값의 집단간 비교를 실시하였다. 개념구성타당도를 측정하기 위하여 일반적으로 사용되는 요인분석방법을 본 연구에서는 실시하지 않았는데, 이는 KGHQ가 영문의 번역판이어서 문항내용의 개념적 구성에 관한 외국의 선행연구 결과를 받아들일 수 있기 때문이다. 또한 기준관련타당도(*criterion validity*) 측정을 위해서는 KGHQ와 다른 척도들 간의 상관관계 측정에 의한 동시적(공존) 타당성(*concurrent validity*) 검증을 실시하였으며, KGHQ가 정신질환자와 정상인들을 제대로 선별해 낼 수 있는 정도를 나타내는 타당도 계수(*Validity coefficients*)인 민감도(*sensitivity*)와 명확도(*specificity*)를 측정함으로써 KGHQ의 예측적 타당도(*predictive validity*)도 검증하였다.

모든 상관도는 Pearson's r의 값에 준하였으며 집단간의 평균값 비교를 위해서는 t-test를 실시하였다. 척도간 상관도 측정을 위한 비교척도로서는 KGHQ-30의 경우에 SCL-90과의 상관관계를 측정하였으며, KGHQ-20은 앞에서 소개한 네 가지 비교척도들과의 상관관계분석을 위해 20문항의 하위요인을 추출한 후 그 주요 요인인 불안요인 및 우울요인과의 척도간 상관도를 측정하였다. 요인추출은 GHQ 사용안내서(Godberg & Williams, 1991)에서 하위요인 분석을 위해 일반적으로 추천하고 있는 방법인 varimax 회전에 의한 주성분 분석법을 사용하였다.

정신질환의 위험집단을 선별하기 위한 기준점인 분할점(*cut-off point*)을 설정하기 위하여 KGHQ-30형의 경우에는 정신건강에 전혀 문제가 없이 건강한 것으로 판별된 대학생집단[1]과 만성정신질환자집단으로 나눈 실제의 두 대조집단과, 분할가능점에서 건강집단과 문제집단으로 분류가 예측되는 두 예상집단 간의 교차분석을 실시하고 상관계수인 phi값을 측정하였으며, 그 값과 타당도계수인 민감도와 명확도를 동시에 고려하여 분할점수를 설정하였다. 또한 표준화는 거치지 않았지만 이미 한국보건사회연구원에서 사용된 적이 있는 KGHQ-20형의 경우에는 그 당시의 결과를 참고로 하여, 분할점을 총점평균보다 한 단위의 표준편차 높은 점수를 기준으로 설정하였다. 이 과정에 관해서는 다음 장에서 보다 자세히 설명하기로 한다. 모든 통계는 SPSS-WIN 10.0 version을 사용하여 처리하였다.

## 4. 연구결과 및 고찰

### 1) 신뢰도

검사-재검사 방법과 문항간 내적 합치도 측정에 근거한 KGHQ의 신뢰도는 매우 높게 나타났다(〈표 2〉). 대학생 57명을 대상으로 2주 간격으로 측정한 KGHQ-30의 검사-재검사 신뢰도 계수는 0.85였으며, 61명을 대상으로 1주 간격으로 측정한 KGHQ-20의 검사-재검사 신뢰도는 0.87로 비교적 높게 나타났는데 이는 외국의 선행연구(〈표 1〉 참조)에서 보고된 값의 범위(0.76~0.90)와 비교해 볼 때도 높은 축에 드는 값이다.

정신질환자 집단을 대상으로 한 검사-재검사 방법은 이들이 병원에 입원 또는 외래환자들로서 재검사를 실시하기가 불가능하였으므로 실시하지 않았다. 한편 대학생 501

1) 실제의 두 상이한 비교집단을 선정하는 데 있어서 정신적으로 전혀 문제가 없이 건강한 집단은 대학생 중에서 SCL-90의 일반적인 분할기준점인 T점수 70점(김재환·김광, 1984)보다 엄격한 기준을 적용함으로써 문제가 전혀 없다고 판단되는 대상을 선발하였다. 즉 남학생의 경우에는 SCL-90의 9개 하위 척도 중에서 하나라도 T점수가 68점 이상인 경우(여학생의 경우에는 65점 이상)를 모두 제외하였으며, 이렇게 하여 선발된 인원은 총 162명이었다.

명 227명의 두 집단 및 정신질환자 87명 등 세 집단을 대상으로 측정한 집단별 내적 합치도 계수(Chronbach's $\alpha$)는 0.89~0.91이었다. 이 결과 역시 선행연구에서 $\alpha$값이 0.85~0.93으로 나타났음과 비교할 때 높은 편이며, 20문항형의 경우에는 1993년 한국보건사회연구원의 연구에서 보고된 값($\alpha$= 0.77)과 비교해 볼 때 본 연구에서 수정·개발한 KGHQ-20의 $\alpha$값이 뚜렷하게 높았다($\alpha$= 0.89).

한편 각 문항별 내적 합치도를 보기 위한 문항-총점간 상관계수는 〈표 3〉과 같다. 대다수의 문항들은 총점과 '비교적 높은 상관'(이인재 등, 1997, p.104)을 나타내었으며 이 또한 KGHQ가 높은 내적 합치도를 갖고 있음을 나타낸다. 단지 11번 문항인 "그 전에 했던 만큼 외출을 자주 하였습니까?"가 공통적으로 가장 낮은 상관도를 보였으며 특히 정신질환자 집단에서는 그 값이 매우 낮았다(r=0.04). 이는 응답자들이

〈표 2〉 KGHQ의 신뢰도

| 종 류 | 대학생 집단 | | 정신질환자집단 |
|---|---|---|---|
| | KGHQ-20 (N=501) | KGHQ-30 (N=227) | KGHQ-30 (N=87) |
| Chronbach's $\alpha$ | . 89 | . 91 | . 90 |
| 검사-재검사 신뢰계수 | . 87 | . 85 | - |

〈표 3〉 문항-총점간 상관계수

| 문항 번호 | KGHQ -20 | KGHQ-30 | |
|---|---|---|---|
| | | 대학생 | 환자 |
| 1 | .47** | .52** | .48** |
| 2 | .44** | .40** | .41** |
| 3 | .6** | .54** | .44** |
| 4 | .4** | .52** | .59** |
| 5 | .40** | .41** | .47** |
| 6 | .54** | .45** | .50** |
| 7 | .50** | .51** | .48** |
| 8 | .54** | .44** | .39** |
| 9 | .57** | .62** | .58** |
| 10 | .62** | .60** | .62** |

| 문항 번호 | KGHQ -20 | KGHQ-30 | |
|---|---|---|---|
| | | 대학생 | 환자 |
| 11 | .29** | .27** | .03 |
| 12 | .39** | .48** | .65** |
| 13 | .46** | .56** | .50** |
| 14 | .62** | .42** | .42** |
| 15 | .57** | .40** | .62** |
| 16 | .42** | .39** | .28** |
| 17 | .55** | .52** | .57** |
| 18 | .54** | .64** | .51** |
| 19 | .54** | .41** | .35** |
| 20 | .51** | .55** | .52** |

| 문항 번호 | KGHQ-30 | |
|---|---|---|
| | 대학생 | 환자 |
| 21 | .56** | .35** |
| 22 | .33** | .60** |
| 23 | .39** | .63** |
| 24 | .42** | .53** |
| 25 | .56** | .38** |
| 26 | .43** | .46** |
| 27 | .30** | .68** |
| 28 | .45** | .53** |
| 29 | .37** | .58** |
| 30 | .45** | .59** |

** P<. 01

만성질환자들이고 특히 다수가 입원환자들이어서 이 문항의 변별력이 없었기 때문으로 해석되며, 따라서 만성질환자들뿐만 아니라 기타 입원환자나 노인 등 평소에 외출이 적은 대상자들에게 KGHQ를 적용할 때는 이 문항을 제외하여도 좋을 것으로 판단된다. 반면 대학생 집단에서도 이 문항의 상관계수가 0.27과 0.30으로 낮은 편이었으나 유의미한 상관관계를 보였으며, 이 문항들의 삭제 후 $\alpha$값(Alpha if item deleted)의 증가가 0.02~0.03 정도로 미미하였으므로 일반인들을 대상으로 할 때는 이 문항을 전체 문항에서 삭제할 필요는 없다고 본다.

2) 타당도

(1) 집단간 평균값의 비교

일반적인 정신건강수준을 가진 대학생집단과 정신질환자집단 간의 KGHQ-30 총점 평균값을 비교한 t-test 결과는 〈표 4〉와 같다. 대학생 집단의 평균값(11.04)이 환자집단(19.45)보다 뚜렷이 낮게 나타났으며 그 차이는 $p<.01$ 수준에서 유의미하였다. 이 결과는 KGHQ가 비정상적인 정신건강 상태를 정상적인 상태로부터 변별할 수 있는 개념적 구성을 이루고 있음을 나타내며 이로서 정신질환자 및 위험군을 건강한 집단으로부터 분류할 수 있는 측정도구로서의 실증적 타당성을 입증하였다고 볼 수 있다. 두 집단간 차이를 입증하기 위해 참고로 분석한 SCL-90의 집단간 총점 평균값의 차이도 유의미하게 나타났다.

(2) 척도간 상관도

KGHQ-20의 동시적 타당도 검증을 위한 측정도구로는 CES-D, SDS, BAI, STAI, SCL-90을 사용하였고, KGHQ-30의 경우에는 SCL-90만을 사용하였는데 이는 앞서 밝힌 바와 같이 전체 문항수를 고려하였기 때문이며 문항수가 많을 경우에는 환자집단에서 조사가 어려울 것을 감안한 때문이다. 본 연구의 응답자료에서 측정된 척도들의 신뢰도는 모두 높았으며 〈표 5〉, 이들과 KGHQ와의 척도간 상관도는 〈표 6〉에 나타난 바와 같다.

〈표 4〉 KGHQ-30과 SCL-90 총점 평균값의 집단간 비교

| 집 단 | 척 도 | N | 평 균 | SD | t | df | Sig |
|---|---|---|---|---|---|---|---|
| KGHQ-30 | 대학생 | 227 | 11.04 | 6.51 | -11.19 | 312 | .000 |
| | 정신질환자 | 87 | 19.45 | 4.19 | | | |
| SCL-90 | 대학생 | 220 | 82.65 | 47.19 | -3.89 | 269 | .000 |
| | 정신질환자 | 51 | 112.02 | 54.48 | | | |

〈표 5〉 비교척도들의 신뢰도(Chronbach's $\alpha$)

| 척 도 | 신뢰계수 |
|---|---|
| CES-D | .90** |
| SDS | .85** |
| BAI | .88** |
| STAI | .95** |
| SCL-90 | .86** |

** $p<.01$

〈표 6〉 KGHQ와 비교척도간의 상관계수

| KGHQ-20 | | | | | |
|---|---|---|---|---|---|
| GHQ-20 | CES-D | SDS | BAI | STAI | SCL-90 |
| 대학생 | .69** | .24** | .51** | .50** | .72** |

| KGHQ-30 | | |
|---|---|---|
| SCL-90 | 대학생 집단 | 정신질환자 집단 |
| | .68** | .68** |

** $p<.01$

KGHQ-20은 모든 척도들과 유의미한 상관도를 보여주었으며 SDS를 제외하고는 모두 0.5이상의 비교적 높은 상관을 보여 주었다. KGHQ-30과 SCL-90 간에도 대학생과 환자 두 집단 모두에서 0.68의 비교적 높은 상관관계를 나타내었다. 한편 KGHQ-20과 SDS 간의 낮은 상관은 SDS의 20문항 중에 신체화 증상을 나타내는 문

항의 수가 8문항으로 다수를 차지하고 있어서 KGHQ보다 신체화증상의 특성을 더 뚜렷하게 포함하기 때문으로 해석된다. 이를 좀더 살펴보기 위해서 KGHQ-20을 요인 분석하여 추출한 4개의 하위요인[2] 중 불안요인 및 우울 요인과 불안, 우울 척도들간의 상관관계를 〈표 7〉에서와 같이 살펴보았다. KGHQ-20의 우울요인은 CES-D와는 비교적 높은 상관도를 가지나(r=0.60), 역시 SDS와는 상관도가 약한 편(r=0.18)으로 나타났고 20문항 전체와의 상관도(r=.24)보다 더 낮게 나타났는데 이는 앞서 언급했듯이 신체화 증상을 포함하는 SDS의 문항내용과의 불일치성 때문임을 입증한다고 본다. KGHQ-20의 불안요인도 BAI와 STAI 두 불안 측정도구들과 유의미한 상관을 가지나 STAI와의 경우에는 20문항 전체와의 상관도(r=0.50)보다 낮은 상관도(r=0.26)를 보였는데 이는 STAI가 조사 당시 현 상태의 불안특성을 측정하는 도구인 반면 KGHQ는 현재를 포함하여 지난 몇 주간의 평균적 상태특성을 점검하기 때문으로 해석되나 이에 대하여는 추후조사가 필요하다.

지금까지의 신뢰도 및 타당도 검증 결과를 요약하면, KGHQ의 두 가지 유형 모두 높은 내적 합치도와 검사-재검사 신뢰도를 보여주었으며, 대학생과 정신질환자 간의 집단간 평균비교에 의해 정신건강수준 측정도구로서의 KGHQ의 개념구성적 타당도가 입증되었으며 비교척도와의 높은 상관관계를 통해 그 동시적 타당도가 입증되었다.

〈표 7〉 KGHQ-20 하위요인과 비교척도들 간의 상관계수

| 척 도 | 불안요인 | 우울요인 |
|---|---|---|
| | 문항번호 2, 5, 12, 19, 20 | 문항번호 6, 9, 10, 11, 16 |
| CES-D | - | .60** |
| SDS | - | .18** |
| BAI | .58** | - |
| STAI | .26** | - |

** $p<.01$

2) 본 연구에서는 KGHQ-20의 하위요인이 4가지로 분류되었는데 불안요인(문항번호 2, 5, 12, 19, 2), 우울요인(문항번호 6, 9, 10, 11, 16), 사회적 부적응(*social disfunction*) (문항번호 1, 3, 4, 7, 8, 13, 14, 15, 17)과 외출빈도(문항번호 11)로 나타났다.

(3) 민감도(sensitivity)와 명확도(specificity)

측정도구의 타당도는 앞에서 설명한 바와 같이 비교 척도간의 상관계수 값(pearson's r)에 의한 동시적 타당도로 주로 나타내지만, 이와 더불어 그 도구가 판명하고자 하는 대상을 제대로 판명하는가를 직접적으로 증명할 수 있는 민감도나 명확도 등으로 나타내기도 한다. 민감도계수는 실제적으로 정신심리적문제를 가진 대상자들 중에서 측정도구에 의해 문제를 가진 자로 판정된 사람의 비율, 즉 정신질환자를 제대로 선별해 내는 정도로 나타내며, 명확도 계수는 정상인들 중에서 도구에 의해 정상으로 판정된 사람의 비율, 즉 정상인을 제대로 선별해 내는 정도를 나타낸다. 본 연구에서는 정신질환자 중 응답의 기재내용이 충실했던 83명과 SCL-90에 의해 선발한 정신적인 문제가 없다고 진단된 건강한 대학생 162명의 KGHQ-30에 대한 응답내용으로 이를 분석하였다.

각 분할가능점에서의 KGHQ-30의 민감도와 명확도는 〈표 8〉에 나타난 바와 같다. 분할점에 따라 민감도와 명확도는 서로 역 방향으로 변화했는데 이는 일반적인 현상이다. 즉 분할가능점이 높아질수록 걸러지는 정신질환자의 수는 줄어드는 반면

〈표 8〉 KGHQ-30 분할가능점에서의 빈도, phi, 민감도 및 명확도

| KGHQ-30 분할점 | | 정신질환자 (N=83) | 정상인대학생 (N=162) | phi | sensitivity | specificity |
|---|---|---|---|---|---|---|
| 14 | ≥14 | >78 | 38 | .66** | >.94 | .77 |
| | <14 | < 5 | 124 | | | |
| 15 | ≥15 | 78 | 32 | .70** | .94 | .80 |
| | <15 | 5 | 130 | | | |
| 16 | ≥16 | 74 | 26 | .70** | .89 | .84 |
| | <16 | 9 | 136 | | | |
| 17 | ≥17 | 67 | 21 | .67** | .81 | .87 |
| | <17 | 14 | 141 | | | |
| 18 | ≥18 | 56 | 26 | .70** | .68 | .90 |
| | <18 | 27 | 146 | | | |
| 19 | ≥19 | 44 | 12 | .52** | .53 | .77 |
| | <19 | 39 | 150 | | | |

** $p<.05$

정상인으로 판명되는 수는 늘어나기 때문이다. 따라서 어디를 분할점으로 정하는가 하는 문제는 두 계수의 적정선과 다양한 측정방법에 의한 타당도 등에 근거하여야 하며 더불어 연구대상의 특성 및 연구목적 등을 고려하여 결정해야 한다. 본 연구 결과에서는 여러 분할가능점수 중에서 15점과 16점에서 phi값이 가장 높게 나타났으며, 16점에서는 민감도 89%와 명확도 84%를 나타내었고 같은 phi값을 나타낸 15점에서는 민감도는 94%로 증가하지만 명확도가 80%로 떨어져서 두 계수간의 차이가 더 컸기 때문에 16점을 분할점수로 선정하였다.

3) 분할점(Cut-off point 혹은 Threshhold point)에 관한 논의

KGHQ와 같은 도구를 사용하여 정신질환자 혹은 정신질환의 위험성을 가진 자(*the probable cases*)를 선별할 때 연구자가 결정해야 할 사항은 분할점을 정하는 문제이다. 조사대상자간 정신건강수준의 상대적 비교는 모든 문항 값의 총점 비교에 의해 측정이 가능하나, 정신문제를 가진 대상자를 선별하고자 하는 경우에는 그 기준이 되는 적절한 분할점수를 정하는 것이 중요한데 이는 매우 어려운 문제가 아닐 수 없다. GHQ의 경우에도 외국의 많은 선행연구에서 다양한 분할 점수가 소개되었으며 이들은 문항형뿐 아니라 조사대상자의 특성에 따라서도 다소 다르게 나타났다. 선행연구들의 공통점을 파악하여 GHQ의 개발자인 Goldberg가 'GHQ scoring' 배점방식에서 제시한 분할점의 기본값(*default*)은 GHQ-20의 경우에는 3/4(최대점수 20), GHQ-30은 4/5(최대점수 30)이며, 이때 각 분할점수와 같은 값을 나타낸 응답자(GHQ-30의 경우 5점)가 정신질환을 가졌을 가능성은 50%로 간주된다(Goldberg & Williams, 1991).

그러나 이 default값에 준할 경우, 본 조사의 대학생 응답자 중 81%가 정신질환의 위험 군에 해당되는 것으로 나타나서 이 값을 한글판 KGHQ에는 적용할 수 없는 것으로 판단되며, 이런 결과는 1993년 한국보건사회연구원에서 지역사회주민을 대상으로 실시한 조사에서도 유사하게 나타났다. 이런 결과에는 여러 요인이 영향을 미쳤을 것이며 액면 그대로 받아들이면 한국인의 정신건강 수준이 실지로 서구인보다 낮기 때문이라고 해석할 수도 있으나, 그보다는 무엇보다도 언어문화적인 배경과 관련이 있는 것으로 추론된다. 즉 모든 문항이 정신적, 육체적 증세를 직접적으로 나타내는

부정적인 내용으로 이루어진 SCL-90 등과는 달리 KGHQ는 부정적인 내용의 문항과 함께 건강한 상태의 여부를 파악하는 긍정적인 문항을 포함하고 있어서 이를 역점처리를 함으로써 불건강 상태를 측정하게 되는데, 이를 응답항목과 함께 한국어로 번역하면 영어원문보다 다소 중도적인 어의를 띠게 되기 때문일 뿐 아니라, 상대적으로 척박한 환경에서 살아가는 한국인들에게는 긍정적인 내용보다 부정적인 문항의 내용이 보다 보편적으로 받아들여질 수 있기 때문이다.

이러한 문제는 서구의 정신병리 측정도구를 도입하여 사용하는 경우 자주 발생할 수 있는 것으로 알려져 있다. 신승철 등(1991)은 한국판 CES-D의 분할점을 미국의 경우와 횡문화적으로 비교한 연구에서 한국인의 CES-D 총점평균(12~13점)이 미국인(2~3점)과 큰 차이가 있는 것으로 보고하였으며, 이는 행복감, 희망 등 긍정적인 감정에 대한 표현에 있어서 한국인이 미국인보다 훨씬 부정적인 때문이라고 해석하였다. 즉 타인을 많이 의식하는 한국인의 대인문화적 특성이 자신의 감정에 대한 긍정적인 표현을 자제하게 만든다고 하였으며, 이 사실의 근거로 아시아계 미국인들 사이에서도 한국계의 CES-D 점수가 가장 높게 나타난 현상을 제시하였다. 이 연구에서는 한국인을 위한 CES-D 분할점을 24점으로 잡았는데 이는 미국인의 16점보다 상당히 높은 것이었다.

한편 GHQ의 타당성에 관해 세계 각국에서 행해진 연구들 중 GHQ-30에 관한 31건의 조사에서 분할점을 default 값인 5점으로 잡은 경우는 22건이고, 나머지는 그보다 높았으며 11~13점으로 잡은 경우도 3건이었다(Goldberg & Williams 1991, p.64). 이러한 사실을 감안하여 본 연구에서는 default값과는 별개로 KGHQ의 분할점을 결정하고자 하였으며 KGHQ-30의 경우에는 이를 위하여 여러 단계의 방법을 사용하였다. 우선, 심리척도들의 분할점을 정할 때 일반적으로 이용되는 점수인 평균값보다 표준편차 한 단위 위의 점수(평균+ 1 SD)를 정하였으며, 이 점수를 기준으로 하여 상·하로 높은 점수들과 낮은 점수들을 KGHQ-30의 분할가능점들로 선정하였다. 각 측정도구의 총점평균 및 표준편차의 값은 〈표 9〉에 나타낸 바와 같으며, 전체 대상자의 빈도분포도를 파악하기 위한 참고로 SCL-90의 값을 측정하여 비교하였다.

〈표 9〉 KGHQ 및 SCL-90의 평균 및 표준편차

| 대상자 | 문항형 | N | 평균 | SD | A: 평균+1SD 값에서의 누적빈도(%) | 1-A (%) |
|---|---|---|---|---|---|---|
| 대학생 | KGHQ-20 | 506 | 9.18 | 4.12 | 83% | 17% |
| | KGHQ-30 | 227 | 11.04 | 6.51 | 84% | 16% |
| | SCL-90 | 227 | 82.65 | 47.19 | 84% | 16% |

마지막으로 이 분할가능점들을 경계로 건강집단과 문제집단으로 나누어질 수 있는 두 예상집단과 대학생 중에서 선발된 건강한 집단과 만성정신질환자집단의 두 실제 집단간의 교차분석을 실시하였다. 즉 실제로 구별되는 두 집단과 각 분할가능점에 의해 분류가 예상되는 두 집단의 상관도(phi)가 상대적으로 높은 경우를 KGHQ-30의 분할점수로 정하고자 하였는데 15점과 16점에서의 phi값이 0.7전후로 가장 높은 상관정도를 보였다(〈표 8〉). 다시 말하면 이 점수들을 분할 기준점으로 할 때, 두 예상집단간의 분류형태가 실제의 두 집단, 즉 정상인 대학생집단과 정신질환자 집단의 분류형태와 가장 일치도가 높은 것으로 나타난다는 것이다. 본 연구에서는 phi값이 같은 두 점수 중에서 명확도와 민감도의 수준을 같이 고려하여 16점을 KGHQ-30의 분할점수로 선정하였다.

KGHQ-30의 경우, 이와 같이 분할점 16점을 기준으로 하면 전체 조사대상 대학생 집단의 17.6%가 정신건강상의 문제가 있는 것으로 판명되었다. 이와 같은 비율은 다소 높은 것으로 보이지만, 1997년 한국보건사회연구원의 우울증 측정에 관한 연구(남정자·조맹제, 1997)에서 밝힌 우리나라 국민(15세~59세)의 25.3%가 경증 이상의 우울증을 갖고 있으며 15세부터 22세 사이 청소년의 34.45%가 경증이상, 13.91%가 중증의 우울증을 갖고 있다는 결과와 그 맥락을 같이한다고 볼 수 있다.

한편 정신질환자를 대상으로 조사를 수행하지 않았던 KGHQ-20은 우울, 불안 등의 비교 측정도구들의 분포도를 감안하여, '평균값+1 표준편차'인 13점을 분할점으로 잡았으며, 이 경우에는 조사대상자의 약 17%가 정신건강상 문제가 있는 집단으로 분류되어 역시 다소 높은 비율을 나타냈으나(〈표 9〉), 이 집단의 16%가 SCL-90의 총점분포에서 '평균+1 표준편차'인 123점보다 높은 점수를 받았음을 감안하면 받

아들일 수 있는 수치라고 하겠다. 1993년 한국보건사회연구원의 연구에서는 당시의 GHQ-20문항 중에서 11문항 이상에 응답한 경우를 잠재적 정신장애라고 정한 바 있는데, 이 점수는 그 20문항 중 내적 일관도가 약하게 나타난 두 문항(11번, 12번 문항)을 제외한 나머지 18문항을 기준으로 한 점수였다(남정자 · 최정수, 1993). 이 사실을 감안한다면 본 연구의 20문항형에서 분할점이 그보다 높은 13점으로 설정된 것은 타당하다고 본다. 그러나 두 문항형의 이와 같이 높은 분할점에 관한 입증은 GHQ의 다른 문항유형에 대한 추후의 연구과정을 통해 보다 더 명확하게 입증할 예정이다.

## 5. 결론 및 제언

본 연구는 1993년에 한국보건사회연구원에서 수행한 '지역사회정신보건 시행방안 연구' 과정에서 한국 최초로 GHQ-20을 지역사회주민의 정신건강수준 측정을 위해 사용하였으나, 사용한 설문문항에 대해 충분한 타당도 검증과정을 거치지 못했던 당시 연구과정에서의 제약성에 기초하고 있다. GHQ가 1970년대부터 영국을 중심으로 하여 세계적으로 널리 쓰여왔음을 감안하여, 우리나라 정신보건사업의 근간이 된 1993년 당시 연구에서 측정도구로서 채택되었으나 그 후 표준화 작업을 거치지 않은 채 사장되어 왔다. 그러나 지난해부터 서울의 모 정신보건센터와 전남의 정신의료기관 등의 정신보건사업현장에 있는 몇 실무자들이 1993년의 연구에 근거하여 GHQ를 임의적으로 사용하기 시작하여 그 사용이 확대되어 나갈 조짐이 보였으므로 1993년 당시 연구에 깊이 참여하였던 본 연구자로서는 이를 보완할 표준화 작업이 시급하다고 판단되었던 것이다. 이는 곧 아직도 우리나라에서는 지역사회주민의 일반적인 정신건강수준을 간단하게 측정할 수 있도록 표준화를 거쳐 개발된 도구가 충분하지 않다는 사실을 보여주는데, 본 연구자가 임의적으로 실무 현장의 사회복지사들로부터 파악한 바로는 이러한 측정도구 및 측정방법에 관한 욕구가 급속히 늘어나고 있는 것으로 보인다.

본 연구에서 한국판으로 표준화하여 개발한 KGHQ-20 및 KGHQ-30은 지역사회 내의 정신건강상의 문제를 지닌 대상자, 즉 '숨겨진 정신질환'의 사례들을 가려내는 일차적인 선별도구로서, 또한 어떤 집단의 정신건강수준을 측정하기 위한 평점도구

로서 간편하고 효과적으로 사용될 수 있다고 본다. 무엇보다도 GHQ는 비의료전문가(*non-clinician*)들에게 유용한 도구로 널리 알려져 있기 때문에 지역사회 내의 소외계층을 항상 접하는 정신보건사회복지사나 일반사회복지사들이 보다 쉽게 지역사회 주민들의 정신건강수준을 측정하기 위한 도구로서 광범위하게 사용될 수 있을 것이다. 특히 IMF이후 증가한 실업 노숙자들 중에서는 알코올이나 약물 중독뿐 아니라 심한 우울증 등의 정신질환을 가졌거나 그 위험 군에 속해있는 사람들이 다수 속해 있는데 현장의 사회복지사들이 이런 사람들을 일차적으로 가려낼 수 있는 방법의 부족으로 어려움을 겪고 있음을 감안할 때 KGHQ가 유용하게 사용될 수 있으리라 보며, 아동학대를 포함한 가정폭력의 피해자와 노인, 장애인 등 모든 사회복지 대상자에게도 이를 적용할 수 있을 것이다.

그러나 도구의 표준화는 매우 까다롭고 기간을 요하는 작업이다. GHQ는 세계 각국에서 번역되어 오랜 기간 다양하게 사용되면서 그 효율성을 인정받은 도구지만, 본 연구과정 중 한글로 번역하는 과정에서 언어적 표현의 뉘앙스가 서구와 많이 다른 것을 보완하면서도 원문의 내용을 그대로 살리는 작업이 가장 어려운 과정이었다. 정신질환의 위험수위 결정의 지침이 되는 분할점을 결정하는 데 있어서 서구에서 사용하는 default 점수를 그대로 적용하지 못하고 KGHQ를 위해 별도로 설정하는 일은 세심한 노력이 요구되었다. 이 문제는 앞으로 KGHQ를 활용하는 여러 연구자가 이 작업을 다양하게 수행하면서 점차 보완되어 가리라 본다. 또한 고려해야 할 사항은 분할점은 조사대상자의 특성에 따라서 어느 정도 달라질 수도 있다는 사실이다. 이는 설문조사를 통하여 정신건강상 문제를 가진 사람들을 선별하는 데 있어서 조사대상자가 속해있는 집단의 성격에 따라 대상자의 가치관과 행동관이 다르게 나타나고, 이에 따라 언어적 표현에 대한 반응 등이 다르게 작동할 수 있고, 이 사실이 곧 응답에 영향을 미치기 때문이다. 따라서 본 연구의 결과도 표집의 편의를 위해 대학생 및 대학원생을 조사대상으로 선택하여 얻어진 것이므로 이를 일반적으로 적용하기 어려운 한계점이 있다. 그럼에도 불구하고 본 연구에서 제시하였던 분할점들은 우리나라 인구의 비교적 건강한 표준 집단인 대학생을 기준으로 설정되었으므로 그 유용한 기준점은 되리라 본다. 본 연구자는 GHQ의 나머지 두 유형인 GHQ-12와 GHQ-28에 대한 추후 연구에서 이 사실을 보완하여 입증하고자 한다.

## ▪참 고 문 헌▪

김덕중. 2001. "정신보건정책의 변화와 동향에 따른 정신보건사회복지사의 역할기대." 《한국정신보건사회사업학회 춘계학술대회 자료집》. pp. 1~8.

김재환 · 김광일. 1984. "간이정신진단검사(SCL-90-R)의 한국판 표준화연구 III." 《정신건강연구》 제2집. pp. 278~311.

고려대학교 부설 행동과학연구소 편. 1999. "심리척도 핸드북 II." 서울 : 학지사.

남정자. 1997. "정신장애자의 현황과 사회복귀 촉진을 위한 정책방안." 《국정신사회재활협회 주최 정신보건토론회 자료집》. pp. 3~5.

남정자 · 조맹제. 1997. "정신건강 증진을 위한 접근책 : 우울증 실태분석을 중심으로." 《보건복지포럼》 제8호. pp. 25~34.

남정자 · 최정수. 1993. "지역사회정신보건 시행방안연구." 《한국보건사회연구원 정책보고서》 제93권, 제19호.

신승철 외. 1991. "한국에서의 the Center for Epidemiological Studies-Depression Scale (CES-D)의 사용 : 표준화 및 요인구조에 대한 횡문화적 검토." 《정신의학회지》 제30권, 제4호, pp. 752~767.

이인재 · 이선우 · 류진석. 1997. 《사회복지통계분석》. 나남출판.

전겸구 · 이민규. 1992. "한국판 CES-D 개발연구 1." 《한국심리학회지 : 임상》 제11호, 제1권, pp. 65~76.

Banks, Michael, H. 1983, "Validation of the General Health Questionnaire in a Young Community Sample." 《Psychological Medicine》 13, pp. 349~353.

Benjamin S., Decalmer., & Haran, D. 1982. "Community Screening for Mental Illness: A Validity Study of the General Health Questionnaire." 《British Journal of Psychiatry》 140, pp. 174~180.

Goldberg, D. & Williams, P. 1991. *A User's Guide to the General Health Questionnaire, Windsor*, England: NFER-Nelson.

Goldberg, D., Gater, R., Sartorius, N., Ustun, T., Piccinelli, M., Gureje. O., & Rutter, C. 1997. "The Validity of two versions of the GHQ in the WHO study of mental illness in general health care." *Psychological Medicine*. 27, pp. 191~197.

Mary, J. J & Williams, P. 1985. "A comparison of the validity of two psychiatric screening questionnaires in Brazil using ROC analysis." *Psychological Medicine*. 15, pp. 651~659.

Goodchild, M. E. & Duncan-Jones, Paul. 1985. "Chronicity and the General Health Questionnaire." *British Journal of Psychiatry*, 146, pp. 55~61.

Hobbs, P. & Ballinger, C. B. & Smith, A. 1. 1983. "Factors Analysis and Validation of the General Health Questionnaire in Women: A General Practice Survey." *British Journal of Psychiatry*, 142, pp. 257~264.

Koeter, M. W. J. & Van Den Brink, W. & Ormel J. 1989. "Chronic Psychiatric Complaints and the General Health Questionnaire." *British Journal of Psychiatry*, 155, pp. 186~190.

Shek, Daniel T. L. 1989. "Validity of the Chinese Version of the General Health Questionnaire." *Journal of Clinical Psychology*, 45(6), pp. 890~897.

Shek, Danel T. L. & Tsang, Sandra K. M. 1995. "Reliability and Factor Structure of the Chinese GHQ-30 for Parents with Preschool Mentally Handicapped Children." *Journal of Clinical Psychology*, 51(2), pp. 227~234.

Vazquez-Barquero, J., Diez-Marique, J., Pena, C., Quintanal, R., & Labrador. L. 1986. "Two Stage Design in a Community Survey." *British Journal of Psychiatry*, 149, pp. 88~97.

## • 부록 : KGHQ 설문문항

아래는 당신의 육체적, 정신적 상태를 포함한 일반적인 건강상태에 관한 질문입니다. 당신이 최근 몇 주 동안 느낀 상태를 그 이전과 비교하여 설명한 것 중에서 가장 가깝다고 생각되는 응답항목 하나를 골라서 그 항목에 표시(V)해 주십시오. 응답내용은 현재를 포함하여 최근 2~3주 동안 당신이 주로 느낀 바에 관한 것이어야 하며, 그 이전 과거의 상태에 관한 것이 아님을 명심하여 주시기 바랍니다.

| 도구 유형 | 문항 번호 | 최근 2~3 주일 전부터 현재까지의 상태를 그 이전과 비교하여 볼 때, 당신은 | ① 매우 그렇다 | ② 그렇다 | ③ 아니다 | ④ 매우 아니다 |
|---|---|---|---|---|---|---|
| KGHQ -20 문항 | 1 | 하고 있는 일에 잘 집중할 수 있었습니까? | | | | |
| | 2 | 걱정 때문에 잠을 잘 못 잔 적이 많았습니까? | | | | |
| | 3 | 자신이 여러 면에서 쓸모 있는 역할을 잘 하고 있다고 느꼈습니까? | | | | |
| | 4 | 매사에 올바른 결정을 잘 내릴 수 있었습니까? | | | | |
| | 5 | 계속해서 긴장감을 느낀 적이 자주 있었습니까? | | | | |
| | 6 | 어려움을 극복할 수 없다고 느낀 적이 있습니까? | | | | |
| | 7 | 일상적인 활동을 즐겁게 할 수 있었습니까? | | | | |
| | 8 | 문제가 생기면 그것을 피하지 않고 맞서서 해결하려고 하였습니까? | | | | |
| | 9 | 우울감과 불행감을 느낀 적이 많습니까? | | | | |
| | 10 | 자신감이 없어졌다고 느꼈습니까? | | | | |
| | 11 | 그 전과 같은 정도로 외출을 하였습니까? | | | | |
| | 12 | 인생이 절망적이라고 느낀 적이 있습니까? | | | | |
| | 13 | 신경이 쓰여서 힘들게 느껴진 일이 많았습니까? | | | | |
| | 14 | 밤에 잠을 잘 못 이루거나 많이 설치곤 합니까? | | | | |
| | 15 | 어떤 일을 할 때 대부분의 다른 사람들이 하는 만큼 그 일을 잘 처리할 수 있었습니까? | | | | |
| | 16 | 당신과 가까운 사람들에게 따뜻함과 애정을 느낄 수 있었습니까? | | | | |

| 도구 유형 | 문항 번호 | 최근 2~3 주일 전부터 현재까지의 상태를 그 이전과 비교하여 볼 때, 당신은 | ① 매우 그렇다 | ② 그렇다 | ③ 아니다 | ④ 매우 아니다 |
|---|---|---|---|---|---|---|
| KGHQ -20 문항 | 17 | 자신의 미래를 희망적으로 느꼈습니까? | | | | |
| | 18 | 부지런히 생활하며 일에 전념할 수 있었습니까? | | | | |
| | 19 | 다른 사람들과 잘 지내는 것이 어렵지 않게 느껴졌습니까? | | | | |
| | 20 | 전반적으로 자신이 일을 잘 하고 있다고 느꼈습니까? | | | | |
| KGHQ -30의 추가 문항 | 21 | 대체적으로 행복하다고 느꼈습니까? | | | | |
| | 22 | 사람들과의 대화시간을 많이 가진 편입니까? | | | | |
| | 23 | 인생이 항상 고생스러운 것으로 생각되었습니까? | | | | |
| | 24 | 자신이 가치 없는 사람이라는 생각을 자주 하였습니까? | | | | |
| | 25 | 자신의 일을 처리해 온 방법에 만족하십니까? | | | | |
| | 26 | 모든 일이 한꺼번에 당신을 힘들게 한다고 느낀 적이 있습니까? | | | | |
| | 27 | 언제나 초조감이나 긴장감을 느끼는 편이었습니까? | | | | |
| | 28 | 때때로 신경이 쇠약해져서 아무 것도 할 수 없던 적이 있습니까? | | | | |
| | 29 | 별 뚜렷한 이유도 없이 공포와 두려움을 느낀 적이 많습니까? | | | | |
| | 30 | 인생이 살 가치가 없게 느껴진 적이 있습니까? | | | | |

# The Validity and Reliability of the Korean Version of the General Health Questionnaire

## KGHQ-20 & KGHQ-30

Shin, Sun-In

(Assistant Professor, Department of Social Welfare, Taegu University)

This study purposed to standardize the Korean versions of the General Health Questionnaire, especially KGHQ-20 and KGHQ-30, through the evaluation of their reliability and validity. The two KGHQs, along with several other scales that measure psychological symptoms, were administered to 916 students and 102 chronically mentally ill patients at mental health hospitals in Korea. Both of the KGHQs were found to have high reliability based on test-retest correlation coefficients, internal consistency(Chronbach's $\alpha$) as scales, and item-total correlations for most of the items. The two KGHQ scores were shown to have high validity based on three methods of analysis; first, both KGHQ scores had highly significant correlations with most of other measures for psychopathology, such as SCL-90, CES-D, SDS, BAI, STAI; second, there were significant differences of the two KGHQ scores between the group of mentally healthy students and the group of chronically mentally ill patients; the validity of the KGHQ-30 was also derived from high sensitivity and specificity, which helped identify the best cut-off points: 15/16 for the KGHQ-30. The cut-off point for the KGHQ-20 was determined at 12/13 based on the mean scores and standard deviations of other scales as well as of the KGHQ-20.