

## 국내 우울증 평가도구 타당화 연구의 체계적 고찰-진단적 타당성을 중심으로 : 한국형 우울 선별 도구 개발을 위한 예비 연구

고려대학교 심리학과,<sup>1</sup> 국립정신건강센터 심리건강과,<sup>2</sup> 한양사이버대학교 청소년상담학과,<sup>3</sup>  
인제대학교 의과대학 일산백병원 정신건강의학과<sup>4</sup>

정수연<sup>1</sup> · 김신향<sup>1</sup> · 박기호<sup>1</sup> · 제갈은주<sup>1</sup> · 이원혜<sup>2</sup> · 최윤영<sup>3</sup> · 이승환<sup>4</sup> · 최기홍<sup>1</sup>

### A Systematic Review of Validation Studies on Depression Rating Scales in Korea, with a Focus on Diagnostic Validity Information : Preliminary Study for Development of Korean Screening Tool for Depression

Sooyun Jung, MA<sup>1</sup>, Shin-Hyang Kim, MA<sup>1</sup>, Kiho Park, BA<sup>1</sup>, Eunju Jaekal, MA<sup>1</sup>, Won-Hye Lee, PhD<sup>2</sup>,  
Younyoung Choi, PhD<sup>3</sup>, Seung-Hwan Lee, MD, PhD<sup>4</sup> and Kee-Hong Choi, PhD<sup>1</sup>

Department of Psychology<sup>1</sup>, Korea University, Seoul,

Department of Clinical Psychology<sup>2</sup>, National Center for Mental Health, Seoul,

Department of Adolescent Counseling<sup>3</sup>, Hanyang Cyber University, Seoul,

Department of Psychiatry<sup>4</sup>, Ilsan Paik Hospital, Inje University College of Medicine, Goyang, Korea

#### ABSTRACT

Reliable and valid diagnostic screening tools in the primary care setting enable accurate estimation of depression in individuals at risk or in need of treatment, and provide patients with an opportunity to receive optimal treatments. Although there have been extensive studies on screening tools for depression used in domestic mental health settings, there is little consideration and lack of a thorough review of the diagnostic validity of screening tools. In the current review, we selected 13 representative screening tools for depression which were evaluated in a total of 19 validation studies conducted in Korea. We summarized DSM-5 target domains, diagnostic indices, sensitivity, specificity, cut-off scores, and diagnostic validity information for each tool. Finally, the depression measurement expert group was constituted to evaluate the current status of screening tools for depression, and their recommendations for a new screening tool were summarized. This study was conducted as part of the Mental Health Technology Development project to develop the Korean screening tool for depression (K-DEP). (Anxiety and Mood 2017;13(2):53-59)

KEY WORDS : Major depressive disorder · Screening tool · Diagnostic validity · Development of a scale · Evidence-based assessment.

#### 서론

우울증은 일반 의료 장면에서 가장 흔히 나타나는 정신 건강

강 이슈임에 반해,<sup>1</sup> 일차 진료기관에 방문한 우울증 환자 중 35~50%는 진단을 받지 못한다.<sup>2</sup> 이에 2002년 미국 예방 서비스 태스크 포스(The United States Preventive Services Task

Received : April 7, 2017 / Revised : May 26, 2017 / Accepted : June 7, 2017

#### Address for correspondence

Younyoung Choi, Ph.D., Department of Adolescent Counseling, Hanyang Cyber University, 220 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea

Tel : +82-2-2290-036, Fax : +82-2-2290-0600, E-mail : 1120008@hycu.ac.kr

#### Address for correspondence

Seung-Hwan Lee, M.D., Ph.D., Department of Psychiatry, Ilsan Paik Hospital, Inje University College of Medicine, 170 Juhwa-ro, Ilsanseo-gu, Goyang 10380, Korea

Tel : +82-31-910-7768, Fax : +82-31-910-7268, E-mail : lshpss@hanmail.net

#### Address for correspondence

Kee-Hong Choi, Ph.D., Department of Psychology, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Korea

Tel : +82-2-3290-2867, Fax : +82-2-3290-5390, E-mail : kchoi1@korea.ac.kr

본 연구는 보건복지부의 재원으로 정신건강 기술개발 사업단의 지원을 받아 수행된 기초연구 사업입니다(HM15C1169).

Force)는 정확한 진단, 효과적 치료, 추적이 가능한 체계를 갖춘 임상 실제에서는 일차 의료 장면에서 우울증의 선별은 우울증 환자를 정확히 확인하고 이들을 위한 치료를 제공하는 데 도움이 되므로, 모든 성인들에 대해 우울증 선별 검사를 실시하도록 권고한 바 있다.<sup>3</sup>

우울증의 선별에 사용되는 검사 도구를 개발 및 선택하는데 있어 고려해야 하는 첫 번째 요인은 평가도구가 우울증 환자를 감별해낼 수 있을 만큼의 양호한 진단적 정확성을 확보하였는지 여부이다. 선별 검사의 정확도는 검사 개발 단계에서 중점적으로 고려하는 요인으로, 민감도와 특이도를 지표로 판단할 수 있다. 민감도(sensitivity)는 실제 질병이 있는 환자를 임상군으로 분류할 수 있는 능력 정도를 수치화 한 것이다. 그러나 우울증 평가도구의 민감도가 100%라고 할 때, 이는 대상자가 우울증이 있거나 없거나 상관없이 모든 사람에게 양성 진단을 내린다는 것을 의미하기 때문에 특이도(specificity)를 진단을 위해서 민감도와 함께 고려된다. 특이도는 실제로 질환이 없는 일반인을 비임상군으로 분류할 수 있는 능력을 수치화 한 것으로 민감도와 특이도가 모두 1.0에 가까울수록 진단적 능력이 높은 평가도구로 판단이 된다. 우수한 민감도와 특이도가 확보된 선별검사는 잠재적 임상군의 정확하게 파악하여 일반군을 환자군으로 오진할 경우 발생할 수 있는 의료진, 연구진의 비용 낭비를 절감할 수 있다. 또한 잠재적 환자군을 일반군으로 오진할 경우 적절한 개입 시기를 놓쳐 우울증의 악화로 이어질 수 있는데, 진단적 타당성이 확보된 선별 도구는 조기 개입의 기회를 증진시킬 수 있다는 이점이 있다.

한편, 국내 의료 및 지역사회 등 다양한 임상 장면에서 사용되는 우울증 평가도구를 통합적으로 분석하려는 시도가 있어왔다. Jung 등<sup>4</sup>은 국내에서 사용되는 대표적인 우울증 평가도구를 소개하고, 국내 임상 현장에서 우울증 환자의 진료, 진단, 심각도 평가, 감별, 선별, 경과 관찰의 목적에 따라 전문가들이 선호하는 평가도구를 조사하였다는 의의가 있다. 또한 Yoon 등<sup>5</sup>은 국내에서 우울증 치료 장면에서 주로 사용되는 소수의 우울증 평가도구에 대해 척도의 구성과 국내 연구 동향, 점수의 해석과 평가도구 사용 장면에서의 이점을 보다 자세하게 보고하였다는 의의가 있다. 선행연구의 발견에서 나아가 본 연구는 한국형 우울 선별도구를 개발하기 위한 첫 단계로 진행된 연구이므로, 우울증 평가도구 개발의 관점에서 우울증 평가도구가 측정하는 구성 개념을 정리하고, 국내에서 시행된 표준화 연구에서 진단적 타당성을 보고한 연구를 중심으로 고찰하고자 한다.

즉, 본 연구는 국내에서 사용중인 우울증의 전문가 면담 평가도구와 자기보고식 평가도구의 속성을 고찰하고 DSM-5의 주요 우울장애 진단 영역을 기준으로 하여 단일한 분석

틀을 구성, 각 평가도구가 측정하는 각기 다른 조작적 정의의 영역을 동일한 기준으로 분류하고자 한다. 더불어 각 평가도구의 국내 대상 진단적 효용성을 분석한 선행 연구를 대상으로 민감도 및 특이도를 산출한 과정의 타당성을 분석하였다. 이를 통해 임상 현장에서 주로 사용되는 우울 평가도구의 효용성과 정확도를 가늠하는 양질의 근거를 제공하고자 한다. 마지막으로 국내 정신건강 의료 현장에서 일하고 있는 전문가 집단을 대상으로 국내 우울증 평가도구의 한계점을 보완하고 효과적으로 한국인의 우울증을 선별해 낼 수 있는 한국형 우울증 선별도구 개발을 위한 전문가 의견을 수집하였다. 본 예비연구는 한국형 우울증 선별도구(Korean Screening Tool for Depression, 이하 K-DEP) 개발의 1년차 성과이며, 추후 진행될 본 검사 개발의 초석이 될 것으로 기대한다.

## 단계 1 : 국내 우울증 평가도구의 측정 영역 및 진단적 타당성 고찰

첫번째로, 국내에서 주로 사용되는 우울증 평가도구를 선정하였다. 도구의 선정은 우울증을 측정하는 대표적인 임상가 면담 우울증 평가도구인 해밀턴 우울 평가 척도(Hamilton Depression Rating Scale, 이하 HDRS), 몽고메리-아스버그 우울 평가 척도(Montgomery-Asberg Depression Rating Scale, 이하 MADRS)와 자기보고식 우울증 평가도구인 벡 우울 검사 2판(Beck Depression Inventory-II, 이하 BDI-II), 역학연구센터 우울척도(Center for Epidemiologic Studies Depression Scale, 이하 CES-D), 환자건강질문지-9(the Nine-Item Patient Health Questionnaire, 이하 PHQ-9), 자가평가 우울 척도(Zung Self-Rating Depression Scale, 이하 SDS), 병원 불안-우울 척도(Hospital Anxiety and Depression Scale, 이하 HADS), 기분 및 불안 증상 질문지(Mood and Anxiety Symptom Questionnaire, 이하 MASQ), 에딘버러 산후 우울 척도(Edinburgh Postnatal Depression Scale, 이하 EPDS), 노인 우울 척도(Geriatric Depression Rating Scale, 이하 GDS)를 대상으로 하였다. 한편 국내 임상 및 지역사회, 연구 현장에서 사용되는 평가도구는 주로 국외에서 개발된 도구이므로, 저작권 및 문화적 차이를 이유로 국내에서도 자체적인 개발을 시도한 선행 연구가 있다. 이에 국내에서 개발된 자기보고식 우울증 평가도구인 한국 우울증 척도(Korean Depression Scale, 이하 KDS), 신 우울불안 척도(New Anxiety-Depression Scale, 이하 NADS), 한국형 노인 우울 검사(Korean form of Geriatric Depression Scale, 이하 KGDS)를 선정하였다.

두번째로, 선정된 우울증 평가도구가 측정하는 구성개념을 DSM-5에서 명시한 주요우울장애 진단기준의 주요 증상으

로 분류하여 각각 어떤 하위 차원을 측정하거나 측정하지 못하는지 여부를 확인하여 정리하였다(Table 1). 그리고 선정된 우울증 평가도구의 진단적 타당성 고찰을 위해 국내 논문 검색을 통해 선정된 평가도구의 표준화 연구 중 진단적 타당성 정보를 명시한 연구를 선정하여 표로 정리하였다. 각 평가도구의 타당화 논문에서 밝혀진 연구 대상(일반인과 임상군)과 준거타당도를 검증하는데 사용된 준거 검사, 검사 도구의 판별력 정보(민감도, 특이도, 절단점, 진단적 타당성)를 정리하였다. 그 결과 총 10개의 평가도구에 대한 19개의 연구 논문이 우울증 평가도구의 판별력 정보를 제공하여, 이를 정리하였다(Table 2).

## 단계 2 : 국내 우울증 평가도구의 평가 영역 및 진단적 타당성 고찰

기존 평가도구들이 측정하는 하위 영역 범위, 타당화 과정의 엄정성, 그리고 진단적 타당성 검증 여부를 중점으로 논의한 뒤, 이를 바탕으로 한국형 우울증 선별도구를 새롭게 개발하는 단계에서 얻을 수 있던 함의를 제안하고자 한다.

### 기존 평가도구 측정 영역 범위

DSM-5의 주요우울장애는 진단을 위해 우울 기분, 흥미저하, 체중 변화, 수면 변화, 정신운동성 초조 혹은 지연, 에너지 감소, 무가치감 혹은 죄책감, 주의집중력 감소, 자살 사고 및 시도 중 5가지 이상의 진단기준을 충족해야 하며, 충족된

기준 중 하나는 반드시 우울기분과 흥미저하 중 하나여야 한다. 그리고 이러한 증상이 개인의 사회, 직업, 일상 생활의 기능을 현저하게 손상하여 고통을 초래해야 한다. 더불어 주요 우울장애의 진단을 위해서는 조증 혹은 경조증 삽화의 이력이 없어야 한다는 조건이 있다. 이러한 진단 기준을 각각 하위 차원으로 명명하여 기존 우울증 평가도구의 측정 영역과 얼마나 겹쳐지는지 확인하였다. 그 결과 우울증 평가 도구 중 가장 많은 하위차원을 측정할 수 있는 평가도구는 PHQ-9이었다. 그러나 PHQ-9은 조증 이력 여부를 평가하는 문항이 없다는 점이 한계로 판단되었다. Jung 등<sup>4</sup>의 국내 우울증 평가도구 분석 연구에서 다양한 정신건강 장면에서 선별을 목적으로 하였을 때 가장 선호하던 평가도구인 BDI-II의 경우, 주요 증상기준 9개를 측정할 수 있지만 사회/직업/일상생활 기능 손상 영역과 조증 이력 영역을 평가하는 문항이 부재한 점이 눈에 띄었다.

### 연구 방법의 엄정성 및 진단적 타당성 검증

기존 우울증 평가도구의 타당화 연구는 내적 일치도 분석, 다른 척도와와의 상관 분석, 요인분석, 진단적 타당성 혹은 판별분석을 주요 분석 방법으로 채택하였다. 대부분의 연구에서 환자군 규명 기준을 진단 체계(DSM, SCID-I 등)를 기준으로 한 정신건강의학과 전문의의 진단으로 하였는데, 이 과정에서 타당화하고자 하는 척도의 점수를 알지 못하는 전문의가 진단을 내리는 이중맹 검사(double blinded test) 처치를 한 연구는 조사 대상이 된 표준화 연구 중 2건에 그쳤다. 이중

Table 1. DSM-5 domains that various depression scales measure

	1	2	3	4	5	6	7	8	9	10	11
HDRS	●	●	●	●	●	●	●		●		
MADRS	●	●	●	●	●	●	●	●	●		
BDI-II	●	●	●	●	●	●	●	●	●		
CES-D	●	●	●	●		●	●	●			
PHQ-9	●	●	●	●	●	●	●	●	●	●	
SDS	●	●	●	●	●	●	●	●	●		
HADS	●	●				●	●				
MASQ	●	●			●	●	●	●			
EPDS	●	●		●			●		●		
GDS	●	●			●	●	●	●			
KDS	●				●	●	●	●	●		
NADS	●	●	●	●	●		●	●	●		

1 : Depressed mood, 2 : Diminished interest or pleasure, 3 : Significant weight loss or gain, 4 : Insomnia or hypersomnia, 5 : Psychomotor agitation or retardation, 6 : Fatigue or loss of energy, 7 : Worthlessness or excessive or guilt, 8 : Diminished ability to think or concentrate, or indecisive, 9 : Recurrent thoughts of death, 10 : Impairment of functioning, 11 : Manic symptoms. HDRS : Hamilton Depression Rating Scale, MADRS : Montgomery-Asberg Depression Rating Scale, BDI-II : Beck Depression Inventory-II, CES-D : Center for Epidemiologic Studies Depression Scale, PHQ-9 : The Nine-Item Patient Health Questionnaire, SDS : Zung Self-Rating Depression Scale, HADS : Hospital Anxiety and Depression Scale, MASQ : Mood and Anxiety Symptom Questionnaire, EPDS : Edinburgh Postnatal Depression Scale, GDS : Geriatric Depression Rating Scale, KDS : Korean Depression Scale, NADS : New Anxiety-Depression Scale

맹 검사는 척도 타당화 연구에서 엄정한 연구 설계를 위해 사용되는 기법이다.

또한 본 연구진이 개발하고자 하는 한국형 우울증 선별도구는 일차 선별을 주 목적으로 하며 측정 도구의 진단적 타당성, 민감도 및 특이도의 개선을 주안점으로 두고 있다. 고찰의 대상이 되었던 13개 평가도구 중, 국내 타당화 연구에서 진

단적 타당성 혹은 민감도/특이도가 보고된 평가도구는 9개 (BDI-II, CES-D, PHQ-9, SDS, HADS, KDS, EPDS, NADS, GDS, KGDS)이며 관련 연구 논문은 20건이었다. 각 논문에서 제시한 민감도 및 특이도는 0.5~1.0 수준으로 다양하였다. 이중 민감도 및 특이도가 모두 0.9 이상인 척도는 BDI-II<sup>6</sup>와 EPDS<sup>7</sup>이다. BDI-II의 경우 임상군의 샘플 사이즈가 33명으로

**Table 2.** Summary of diagnostic validity information reported by validation studies of existing depression rating scales in Korea<sup>6-23</sup>

Scales	Year of issue	Clinical control	Healthy control	Index test	Sensitivity	Specificity	Cutoff score	Diagnostic validity information
CES-D	1991	30	50	Clinician's confirmation	-	-	24	Kappa=0.92
	1992	84 (MDD: 13)	220	Clinician's confirmation	-	-	16	-
	1993	164 (MDD: 46 rMDD: 27)	460	DSM-IV, CI(HDRS, DIS)	CD : 0.913 CS : 0.957	CD : 0.788 CS : 0.695	CD : 25 CS : 21	AUC : 91.40%
PHQ-9	2007	138	22	DSM-IV	0.818	0.899	10	AUC : 94.40%
	2009	109 (DEP : 61 ANX : 48)	58	DSM-IV	0.85	0.52	9.5	AUC : 92.20%
	2013	74	75	CI (M.I.N.I.-Plus)	0.885	0.947	9	AUC : 97.60%
KDS	2003	142	3,697	-	-	-	Male : 65 Female : 70	AUC : 73.10%
	2012	448	3,697	SCID-I	-	-	Male : 65 Female : 70	AUC : 77.70%
EPDS	2004	33 (MDD : 11 MnDD : 12)	125	CI(SCID, HDRS)	1.00 (MDD, MnDD)	MDD : 0.905 MnDD : 0.784	MDD : 12.5 MnDD : 9.5	-
	2008	30	209	CI(SCID-I)	-	-	9/10	AUC: 81.80%
NADS	1987	-	3,816	-	-	-	50	-
BDI-II	2008	31	77	DSM-IV	0.935	0.981	22	AUC: 99.20%
	2011	739	302	CI(SCID)	0.85	0.881	18	AUC: 92.80%
HADS	1999	ANX : 66 DEP : 74	189	DSM-IV, Clinician's confirmation (2)	ANX : 0.788 DEP : 0.892	ANX : 0.825 DEP : 0.825	ANX : 8 DEP : 8	-
GDS	1997	81	94	P : DSM-IV, Doctor HCI : DSM-IV, CI (CEG) R/O : MMSE-K	0.656	0.649	18	-
	1999	35	51	CI(DIS-III-R), Clinician's confirmation (independently) R/O : MMSE-K	0.943 0.914	0.745 0.824	17 19	Kappa=0.655 Kappa=0.718 AUC : 90.1%
	2004	62	92	CI(DIS-III-R), Clinician's confirmation (independently) R/O : MMSE-K	0.903 0.839	0.717 0.815	16 18	Kappa=0.590 Kappa=0.642 AUC : 87.5%
	2008	MDD : 61 MnDD : 45	782	CI(MINI, CERAD-K), DSM-IV	MDD : 0.872 MnDD : 0.820	MDD : 0.812 MnDD : 0.781	MDD : 17 MnDD : 16	AUC : 91% (MDD), 87% (MnDD)
K-GDS	1997	129	13	CI (DSM-IV), CEG R/O : MMSE-L	0.746	0.720	14	-

MDD : major depressive disorder, MnDD : minor depressive disorder, rMDD : recovered major depressive disorder, DEP : depressive symptoms, ANX : anxiety disorder, DSM-IV : diagnosis based on DSM-IV criteria, CI : Clinical interview, Clinician's confirmation : diagnosed by medical doctor or psychiatrist, MC : medical card, HC : healthy control, P : patients, CD : for clinical diagnosis, CS : for community screening

매우 작으며, 지표 검사 또한 정신건강의학과 전문의가 DSM-IV를 기준으로 진단하였으나 이중맹검이 아니라는 한계점이 있었다. EPDS의 경우 척도 시행 결과를 모르는 전문의의 진단으로 지표 검사를 시행하였지만 주요우울장애군 환자 수가 11명으로 매우 적다는 한계점이 있었다.

### 단계 3 : 국내 우울증 평가도구 사용 실태에 대한 전문가 견해 수집

국내에서 다양한 장면에서 사용되는 우울증 평가도구의 사용 실태에 대한 전문가 견해를 수집하기 위해 전문가 집단을 구성하여 회의를 개최하였다. 전문가 집단은 임상심리전문가 3명, 정신과 전문의 1명, 측정심리 전문가 1명으로 구성된 본 연구진과 외부 자문단(정신과 전문의 4명, 임상심리 전문가 2명, 심리측정 전문가 2명)으로 구성되었다. 기존 평가도구의 한계점 및 보완점에 대해 고찰하고, 이를 바탕으로 본 연구진이 제작하고자 하는 한국형 우울증 선별도구 개발의 의미와 지향점, 최종 평가도구의 구체적 결정 사항을 중점으로 토의가 이루어졌고, 자세한 내용은 Table 3에 정리하였다.

전문가 자문단의 의견을 토대로 본 연구진은 한국형 우울 선별도구의 용도를 임상 장면에서의 자세한 진단보다는 일차 의료현장 등에서의 간편하고도 정교한 선별에 방점을 두어

개발하는 것에 대해서 의견을 모았다. 이를 위해 주요우울장애의 선별을 위해서는 DSM-5에 명시된 아홉 가지 핵심 진단 기준을 모두 포괄할 뿐만 아니라 수검자의 신체적 호소, 불안 장애, 양극성 장애와의 감별을 위해 신체적 증상, 불안 증상, 조증 증상 등의 구성개념에 대한 질문이 평가도구에 포함되는 방안이 대해서 검토하였다. 또한 측정하고자 하는 증상의 기간 단서, 문항의 앵커링, 채점 방식, 문항의 수량 등의 심리 측정적 속성을 기존 우울증 평가도구를 참고하여 결정하였고, 최종 결과물을 배급하고 활용하는 과정에서 예측하고 고려해야 할 지점을 검토하였다. 이를 바탕으로 본 연구진이 개발한 한국형 우울증 선별도구 예비문항의 예시를 Table 4에 정리하였다.

## 결론

전통적으로 주요우울장애는 정신질환 평가 및 치료 장면에서 임상 전문가, 의료진, 보건 정책 담당자 등 다양한 구성원이 주목하는 정신질환이다. 우울 환자의 높은 유병률과 질환으로 인한 사회적 비용을 고려하였을 때, 주요우울장애의 선별에 대한 임상 현장과 학계의 관심이 높다는 점은 고무적이다. 기존 우울증 평가도구의 체계적인 고찰과 전문가 집단의 견해를 바탕으로 본 연구진은 한국형 우울증 선별도구

**Table 3.** Summary of Recommendations by Expert Committee for the Korean Depression Screening Tool (K-DEP)

The main purpose of development of K-DEP	<ul style="list-style-type: none"> <li>- The K-DEP development team needs to define the main purpose of the rating tool in order to screen out latent patients with depression related disorders or give diagnostic suggestions.</li> <li>- The committee recommends to set the main goal of the K-DEP as a primary screening tool, which gives latent depression patients a chance to get proper psychiatric services.</li> <li>- Time/Cost efficiency : The K-DEP needs to be short and cost-effective enough to simply rate the depression in the primary medical setting or local community setting for screening.</li> <li>- It would be critical to retain sufficient examinee size to be survived as a useful screening tool in Korean clinical and public health settings within the first several years from the distribution of K-DEP.</li> </ul>
Distribution of the K-DEP in various settings	<ul style="list-style-type: none"> <li>- It is needed to consider both the universal criteria to capture diagnostic depression and unique properties of Korean depression.</li> <li>- The item-driven scoring pattern will give us the information of how Korean depression looks like.</li> </ul>
Define what 'the Korean' screening tool means	<ul style="list-style-type: none"> <li>- Domains : To check whether the examinee has had any manic episodes, it would be recommended to add items rating manic symptoms.</li> <li>- Anchoring : The united anchoring system would be recommended, as the people with depression tend to show the diminished ability to concentrate.</li> <li>- Item numbers : Questions of K-DEP driven by the item response theory will assess the depression with the minimum number of items.</li> <li>- Sensitivity/Specificity : The K-DEP is needed to be improved the diagnostic efficiency better than the existing screening tools.</li> </ul>
The psychometric properties of K-DEP	<ul style="list-style-type: none"> <li>- The K-DEP will be expected to be the successful replacement of the existing depression rating scales which have a royalty issue, as the K-DEP will be a free test for public health settings.</li> </ul>
The Final product of K-DEP	<ul style="list-style-type: none"> <li>- Based on the bi-factor model, the compulsory items, which function as linking items between online-offline tests, will be included in both on-line and off-line version of K-DEP.</li> </ul>

**Table 4.** DSM-5 domains that the K-DEP preliminary items measure

No	DSM-5 domains	Additional domains	Examples
1	Depressed mood	-	지난 2주간 어느 누가 도와주어도 우울한 기분을 떨쳐버리기 어려웠다. 지난 2주간 이유 없이 눈물이 줄줄 났다.
2	Diminished interest or pleasure	-	지난 2주간 나의 일상에는 흥미로운 일이 거의 없었다. 지난 2주간 이전에 비해서 다른 사람에 관한 관심이 줄었다.
3	Significant weight loss or gain	-	지난 2주간 입맛을 잃어, 다른 사람의 강요 없이는 밥을 잘 먹지 않았다. 최근 한 달 사이에 체중이 변화하여 평소에 입던 옷이 맞지 않았다.
4	Insomnia or hypersomnia	-	지난 2주간 밤중에 잠에서 깨는 경우가 많았고, 한 번 깨면 다시 잠들기 쉽지 않았다. 지난 2주간 나는 거의 온종일을 잤다.
5	Psychomotor agitation or retardation	-	지난 2주간 안절부절못하여 진정하기 어려웠다. 지난 2주간 다른 사람이 알아챌 정도로 생각이나 말이 느려졌다.
6	Fatigue or loss of energy	-	지난 2주간 너무 피로하고 지친 나머지, 이전에 할 수 있던 것 중 많은 일을 하지 못하였다. 지난 2주간 하는 일마다 힘들다고 느꼈다.
7	Worthlessness or excessive or guilt	-	지난 2주간 나는 벌을 받고 있다고 생각했다. 지난 2주간 나는 사회에 필요하지 않은 사람이라고 생각했다.
8	Diminished ability to think or concentrate, or indecisive	-	지난 2주간 매사에 결단력이 부족하다고 생각했다. 지난 2주간 어떤 일에도 집중할 수 없었다.
9	Recurrent thoughts of death	-	지난 2주간 자살하려고 유서를 쓴 적이 있었다. 지난 2주간 현실적으로 실현 가능한 자살 방법을 생각했다.
10	Impairment of functioning	-	지난 2주간 규칙적인 생활을 하는 것이 힘들어 직장에 갈 수 없었다. 지난 2주간 일상생활을 유지할 만한 능력이 떨어졌다.
11	-	Somatic symptoms	지난 2주간 심장이 두근거리고 식은땀이 나거나, 두통 등의 증상을 경험했다. 지난 2주간 신체적 건강 때문에 주변 사람들에게 불평하거나 도움을 청하였다.
12	-	Anxiety, worry or fear	지난 2주간 걱정할 게 없음에도 불구하고 불안했다. 지난 2주간 별다른 이유 없이 겁을 먹거나 공포에 휩싸였다.
13	-	Manic symptom	에너지가 넘쳤던 시기에 주변 사람들을 짜증 나고 진 빠지게 했다. 기분이 고양된 상태에서 폭음하거나 무분별한 성행위를 했던 시기가 있었다.
14	-	Hopelessness	지난 2주간 내 장래는 희망적이지 않다고 생각했다. 지난 2주간 내가 진정으로 원하는 것을 얻지 못할 것이라는 생각이 들었다.

(K-DEP)를 개발하고자 한다. 대다수의 우울증 평가도구가 주요우울장애의 핵심 증상을 측정하지만, 그 진단기준을 모두 포함하는 경우는 드물며, Park 등<sup>24</sup>에 따르면 DSM-5의 주요우울장애의 핵심 진단기준 9개를 모두 충족하는 경우는 주요우울장애 환자의 3%에 불과하고, 대부분 진단 기준의 세부적인 조합이 다양하게 분포되어있다. 주요우울장애의 진단이 관련 증상의 조합으로 이루어진다는 점을 고려하면, 우울증 환자의 집단 내 이질성을 극복하기 위해서 주요우울장애의 핵심 진단기준을 고루 질문할 수 있는 평가도구가 필요하다. K-DEP는 주요우울장애의 핵심 진단기준에 더해 조증, 불안 증상, 신체증상, 무망감 영역을 보완하여 양극성 장애, 불안 장애와 감별이 용이하도록 개선하고자 한다. 또한 문항반응 이론(item response theory)을 활용한 최적 및 최소 문항의 선별을 통해 좋은 선별도구의 요건인 정확성(민감도 및 특이도)과 비용 효율성, 실용성을 도모하고자 한다. 더불어 검사

의 타당화를 위한 연구 설계에 K-DEP의 실시 및 채점자와 진단 면접을 실시하는 전문가를 다르게 설정하는 이중 맹검법을 도입하여 기존의 우울증 선별 및 평가도구보다 더욱 합리적이고 유용한 검사임을 입증하는 방향으로 타당화 하고자 한다.

이렇게 개발된 K-DEP는 지역사회, 정신의학과 현장, 일차 의료 현장 등에서 주요우울장애를 신뢰롭게 선별하여 잠재적인 우울증 환자에 대한 초기 개입을 통해 질환의 만성화를 예방하며, 잠재적인 우울증 환자의 오진으로 인한 사회적 비용을 절감할 수 있을 것으로 기대된다. 선별도구가 효과적으로 사용된다면 지역사회, 일차 의료현장 등 잠재적 환자들이 접하기 쉬운 장면에서부터 활용되어야 하는데, 이를 위해서는 비용에 대한 부담이 적어야 한다. K-DEP는 보건복지부의 정신건강사업의 일환으로 개발되는 것으로 무료 배포가 가능하여 국가 연구 과제에서 비용에 대한 부담 없이 활용될 수 있

을 것이다. 한국인의 특수성을 고려하여 제작되고 측정 도구로서의 신뢰성과 타당성, 그리고 진단적 타당성을 갖춘 한국형 우울증 선별도구가 개발된다면 잠재적 수검자와 의료 현장에 도움이 되는 성과가 될 것으로 예측한다.

**중심 단어** : 주요우울장애 · 선별도구 · 진단적 타당성 · 평가도구 개발 · 근거기반 심리 측정.

## REFERENCES

- Kroenke K, Spitzer RL, Williams JB, Monahan, PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine* 2007;146:317-325.
- Von Korff M, Shapiro S, Burke JD, Teitlebaum M, Skinner EA, Turner RW, et al. Anxiety and depression in a primary care clinic: comparison of Diagnostic Interview Schedule, General Health Questionnaire, and practitioner assessments. *Archives of General Psychiatry* 1987;44:152-156.
- Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD et al. Screening for depression in adults: a summary of the evidence for the US Preventive Services Task Force. *Annals of Internal Medicine* 2002;136:765-776.
- Jung SW, Lee EJ, Choi YK, Sung HM, Kim DH, Choi JS, et al. A Guide of Diagnostic Evaluation for Depression: focused on Assessment Instrument for Depression. *The Korean Journal of Psychopathology* 2009;18:21-38.
- Yoon SY, Lim JH, Han C. Rating scales for measurement-based clinical practice of depression. *Korean Journal of Psychopharmacology* 2012;23:136-146.
- Sung HM, Kim JB, Park YN, Bai DS, Lee SH, Ahn HN. A study on the reliability and the validity of Korean version of the Beck Depression Inventory-II (BDI-II). *Journal of the Korean Society of Biological Therapies in Psychiatry* 2008;14:201-212.
- Han K, Kim M, Park JM. The Edinburgh postnatal depression scale, Korean version: Reliability and validity. *Journal of Korean Society of Biological Therapies in Psychiatry* 2004;10:201-207.
- Shin SC, Kim MK, Yun KS, Kim JH, Lee MS, Moon SJ, et al. The center for epidemiologic studies-depression scale: its use in Korea. *Journal of Korean Neuropsychiatric Association* 1991;30:752-767.
- Chon KK, Rhee MK. Preliminary development of Korean version of CES-D. *Korean Journal of Clinical Psychology* 1992;11:65-76.
- Cho MJ, Kim KH. Diagnostic validity of the CES-D (Korean version) in the assessment of DSM-III-R major depression. *Journal of Korean Neuropsychiatric Association* 1993;32:381-399.
- Choi HS, Choi JH, Park KH, Joo KJ, Ga H, Ko HJ, et al. Standardization of the Korean version of Patient Health Questionnaire-9 as a screening instrument for major depressive disorder. *Journal of the Korean Academy of Family Medicine* 2007;28:114-119.
- Lim KH, Park YN, Kim DH, Shin IH, Lee WS, Kim JB. A preliminary study of the standardization of the Korean version of the patient health questionnaire-9. *Korean Journal of Health Promotion* 2009; 9:275-281.
- Lee MS, Rhee MK. A development of Korea depression scale. *Journal of Korean Neuropsychiatric Association* 2003;42:492-506.
- Hwang SH, Rhee MK, Kang RH, Lee HY, Ham BJ, Lee YS, et al. Development and validation of a screening scale for depression in Korea: the Lee and Rhee Depression Scale. *Psychiatry Investigation* 2012;9:36-44.
- Kim YK, Hur JW, Kim KH, Oh KS, Shin YC. Clinical application of Korean version of Edinburgh postnatal depression scale. *Journal of Korean Neuropsychiatric Association* 2008;47:36-44.
- Lee JB, Kim CS, Park HB, Park BT, Lee JH, Cheung SD. A Study on Anxiety-Depression in College Freshmen. *Yeungnam University Journal of Medicine* 1987;4:29-38.
- Lim SY, Lee EJ, Jeong SW, Kim HC, Jeong CH, Jeon TY, et al. The validation study of Beck Depression Scale 2 in Korean version. *Anxiety and Mood* 2011;7:48-53.
- Oh SM, Min KJ, Park DB. A comparison of normal, depressed and anxious groups: a study on the standardization of the hospital anxiety and depressed scale for Koreans. *Journal of Korean Neuropsychiatric Association* 1999;38:289-296.
- Jung IK, Kwak DI, Shin DK, Lee MS, Lee HS, Kim JY. A reliability and validity study of geriatric depression scale. *Journal of Korean Neuropsychiatric Association* 1997;36:103-112.
- Cho MJ, Bae JN, Suh GH, Hahm BJ, Kim JK, Lee DW, et al. Validation of geriatric depression scale, Korean version GDS in the assessment of DSM-III-R major depression. *Journal of Korean Neuropsychiatric Association* 1999;38:48-63.
- Bae JN, Cho MJ. Development of the Korean version of the Geriatric Depression Scale and its short form among elderly psychiatric patients. *Journal of Psychosomatic Research* 2004;57:297-305.
- Kim JY, Park JH, Lee JJ, Huh Y, Lee SB, Han SK, et al. Standardization of the Korean version of the Geriatric Depression Scale: reliability, validity, and factor structure. *Psychiatry Investigation* 2008;5: 232-238.
- Jung IK, Kwak DI, Joe SH, Lee HS. A study of standardization of Korean form of Geriatric Depression Scale (KGDS). *Journal of Korean Geriatric Psychiatry* 1997;1:61-72.
- Park SC, Kim JM, Jun TY, Lee MS, Kim JB, Yim HW, et al. How Many Different Symptom Combinations Fulfill the Diagnostic Criteria for Major Depressive Disorder? Results from the CRESCEND Study. *Nordic Journal of Psychiatry* 2017; 71:217-222.
- An JY, Seo ER, Lim KH, Shin JH, Kim, JB. Standardization of the Korean version of screening tool for depression (Patient Health Questionnaire-9, PHQ-9). *Journal of the Korean Society of Biological Therapies in Psychiatry* 2013;19:47-56.
- Chon KK, Choi SC, Yang BC. Integrated adaptation of CES-D in Korea. *Korean Journal of Health Psychology* 2001;6:59-76.
- Kim JI. A validation study on the translated Korean version of the Edinburgh postnatal depression scale. *Korean Journal of Women Health Nursing* 2006; 12:204-209.
- Kim JS. A Study on Anxiety-Depression and Pyscnoticism in Hospitalized Patients. *Yeungnam University Journal of Medicine* 1992;9: 54-67.
- Kim MS, Lee IS, Lee CS. The Validation Study 1 of Korean BDI-2: In Female University Students Sample. *The Korean Journal of Clinical Psychology* 2007; 26:997-1014
- Lee H, Kim KH. Validation of the Korean Version of the Mood and Anxiety Symptom Questionnaire (K-MASQ). *Korean Journal of Clinical Psychology* 2014; 33:395-411
- Lee JH. Development of the Korean form of Zung's self-rating depression scale. *Yeungnam University Journal of Medicine* 1995;12: 292-305.
- Lee SA, Kim KH, Cho, SM. Validation of the Mood and Anxiety Symptom Questionnaire in Korean Adolescents. *Psychiatry Investigation* 2015;12:218-226.
- Lee SA, Kim KH, Cho SM. Validation of the Mood and Anxiety Symptom Questionnaire in Korean Adolescents. *Psychiatry Investigation* 2015;12:218-226.
- Lee YH, Song JY. A study of the reliability and the validity of the BDI, SDS, and MMPI-D scales. *Korean Journal of Clinical Psychology* 1991;10:98-113.
- Rhee MK, Lee YH, Jung HY, Choi JH, Kim SH, Kim YK, et al. A Standardization Study of Beck Depression Inventory 2-Korean Version (K-BDI): Validity-II. *Korean Journal of Psychopathology* 1995; 4:96-104.
- Shin HC, Kim CH, Park YW, Cho BL, Song SW, Yun YH, et al. Validity of Zung's self-rating depression scale: detection of depression in primary care. *Journal of the Korean Academy of Family Medicine* 2000;21:1317-1330.