



Validation of the Patient Health Questionnaire–9 and Patient Health Questionnaire–2 in the General Korean Population

Minah Kim^{1,2}, Sanghyup Jung³, Jee Eun Park^{1,2}, Jee Hoon Sohn^{1,4}, Su Jeong Seong⁵,
Byung-Soo Kim⁶, Sung Man Chang⁶, Jin Pyo Hong⁷, Bong-Jin Hahm^{1,3,8}, and Chan-Woo Yeom⁹✉

¹Department of Neuropsychiatry, Seoul National University Hospital, Seoul, Republic of Korea

²Department of Psychiatry, Seoul National University College of Medicine, Seoul, Republic of Korea

³Department of Psychiatry and Behavioral Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea

⁴Public Healthcare Center, Seoul National University Hospital, Seoul, Republic of Korea

⁵Department of Psychiatry, Kangdong Sacred Heart Hospital, Seoul, Republic of Korea

⁶Department of Psychiatry, School of Medicine, Kyungpook National University, Daegu, Republic of Korea

⁷Department of Psychiatry, Sungkyunkwan University School of Medicine, Samsung Medical Center, Seoul, Republic of Korea

⁸Institute of Human Behavioral Medicine, Medical Research Center, Seoul National University, Seoul, Republic of Korea

⁹Department of Psychiatry, Uijeongbu Eulji Medical Center, Eulji University School of Medicine, Uijeonbu, Republic of Korea

Objective The Patient Health Questionnaire–9 (PHQ-9) and PHQ-2 have not been validated in the general Korean population. This study aimed to validate and identify the optimal cutoff scores of the PHQ-9 and PHQ-2 in screening for major depression in the general Korean population.

Methods We used data from 6,022 participants of the Korean Epidemiological Catchment Area Study for Psychiatric Disorders in 2011. Major depression was diagnosed according to the Korean Composite International Diagnostic Interview. Validity, reliability, and receiver operating characteristic curve analyses were performed using the results of the PHQ-9 and Euro Quality of life-5 dimension (EQ-5d).

Results Of the 6,022 participants, 150 were diagnosed with major depression (2.5%). Both PHQ-9 and PHQ-2 demonstrated relatively high reliability and their scores were highly correlated with the “anxiety/depression” score of the EQ-5d. The optimal cutoff score of the PHQ-9 was 5, with a sensitivity of 89.9%, specificity of 84.1%, positive predictive value (PPV) of 12.6%, negative predictive value (NPV) of 99.7%, positive likelihood ratio (LR+) of 5.6, and negative likelihood ratio (LR-) of 0.12. The optimal cutoff score of the PHQ-2 was 2, with a sensitivity of 85.3%, specificity of 83.2%, PPV of 11.6%, NPV of 99.5%, LR+ of 5.1, and LR- of 0.18.

Conclusion The PHQ-9 and PHQ-2 are valid tools for screening major depression in the general Korean population, with suggested cutoff values of 5 and 2 points, respectively.

Psychiatry Investig 2023;20(9):853-860

Keywords Depression; PHQ-9; PHQ-2; Screening; Cutoff; Korea.

INTRODUCTION

Major depressive disorder (MDD) is a serious, yet common mental disorder that lowers patient quality of life and increases suicide rate.¹ In addition, MDD imposes an economic burden on patients and their families owing to poor patient functioning and medical expenses.² Therefore, accurate assessment

and treatment of MDD is one of the most important tasks in public mental health.

The prevalence of MDD in Korea is 6.7%,³ which is lower than that in the United States (10.4%)⁴ and Europe (11.32%)⁵; however, Korea's suicide rate is 24.6 per 100,000 people,⁶ which was the highest in the Organisation for Economic Co-operation and Development from 2003 to 2019. There is a gap between the prevalence of MDD and the suicide rate in Korea, considering that psychological autopsy studies of suicide victims reported comorbid MDD or other mood disorders in approximately 60% of cases.^{7,8} It has also been reported that early screening and management of depression in the community lowers the risk of suicide.⁹ Therefore, it is important to develop a concise and valid tool for screening depression in the general Korean population.

Received: April 1, 2023 **Revised:** May 26, 2023

Accepted: June 30, 2023

✉ **Correspondence:** Chan-Woo Yeom, MD

Department of Psychiatry, Uijeongbu Eulji Medical Center, Eulji University School of Medicine, 712 Dongil-ro, Uijeongbu 11759, Republic of Korea

Tel: +82-31-951-2379, **Fax:** +82-31-951-1093

E-mail: chwoye@gmail.com, chwoye@eulji.ac.kr

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Patient Health Questionnaire-9 (PHQ-9), developed by Spitzer et al.¹⁰ in 1999, is a self-report instrument designed to detect depression in primary care. The PHQ-9 has several advantages, which include being brief, multipurpose, free, and easy to score.¹¹ The PHQ-2, which consists of the first two items of the PHQ-9, was developed in response to the need for briefer measures.¹² Currently, the PHQ-9 and PHQ-2 are used worldwide, and validation studies and cutoff scores have been set for each country.^{13,14} In Korea, validation of the PHQ-9 or PHQ-2 has been carried out only for specific populations, such as the elderly,¹⁵ patients with migraine,¹⁶ and psychiatric patients.¹⁷ For the general public population, one study has been conducted to obtain the normative data of the PHQ-9 using the nationwide cross sectional survey data of Korea from 2014 to 2016 (Korea National Health and Nutrition Examination Survey)¹⁸; however, criterion validation and cutoff setting were not performed.

The present study therefore aimed to validate and set the cutoff points of PHQ-9 and PHQ-2 in major depression screening in the general Korean population. We calculated the sensitivity, specificity, and optimal cutoff of the PHQ-9 and PHQ-2 through receiver operating characteristic (ROC) analysis with a Korean national representative survey of mental health.

METHODS

Participants and procedures

The Korean Epidemiological Catchment Area Study for Psychiatric Disorders (KECA)-2011,¹⁹ a nationally representative survey of mental health in the general population aged ≥ 18 years, was conducted between March and December of 2011. Multistage stratified sampling was conducted across 12 catchment areas, where each included sample was independent of the others. Based on 2010 Population and Housing Census data,²⁰ 14,204 households were selected. The households selected for the survey were visited in advance; those suitable for the survey were identified and interviewees were selected. As a result of the previsit, we excluded households that were unsuitable; for example, when there was no actual household, redevelopment, no resident, no resident who met the criteria for the investigation, or when it was impossible to confirm who could be the target of the investigation. In total, 7,650 adults were contacted, and 6,022 completed the interview (response rate: 78.7%).^{3,19}

All participants were informed of the methods and purpose of the survey and provided written informed consent prior to participation. This study was approved by the Institutional Review Board of the Seoul National University Hospital (IRB No. C-1104-092-359).

Measures

In KECA-2011, structured interviews were conducted using the Korean version of the Composite International Diagnostic Interview 2.1 (K-CIDI 2.1) and various auxiliary tools, including the PHQ-9 and Euro Quality of life-5 dimension (EQ-5D). Prior to the interviews, 78 interviewers completed a 5-day training session based on the standard protocol and training materials developed by the World Health Organization (WHO).

K-CIDI

The CIDI is a fully structured interview used to identify mental illnesses.²¹ In the CIDI, an algorithm automatically makes a diagnosis based on the answers to each question. The K-CIDI was translated and validated by Cho et al.²² according to the WHO guidelines.²³ The K-CIDI can identify both the International Classification of Diseases 10th revision and Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) diagnoses. In this study, only DSM-IV MDD was diagnosed using Chapter E of the K-CIDI. Lifetime MDD (1 year, 1 month, and 2 weeks) can be diagnosed using the K-CIDI algorithm. For comparison with the PHQ-9, the 2-week K-CIDI MDD results were used.

PHQ-9 and PHQ-2

The PHQ-9 is a self-reported scale designed to identify depression in a primary care setting.²⁴ The PHQ-9 consists of nine symptoms that are the diagnostic criteria for MDD in the DSM-IV.²⁵ Each item is measured on a scale of 0 to 3 points (0=never, 1=a few days, 2=less than half of the day, and 3=most of the day) for the past two weeks, including the day the questionnaire was completed. The total score ranged from 0 to 27, and the higher the score, the higher the severity of depressive symptoms. Kroenke et al.²⁴ developed the PHQ-9 and conducted a validation study targeting primary care and obstetric gynecology clinics. As a result, when the cutoff score of PHQ-9 was ≥ 10 , major depression was detected with a sensitivity and specificity of 88%.

Kroenke et al.¹² also adapted the PHQ-9 to identify depression with only two items for use in busy clinical settings, naming this two-item questionnaire the PHQ-2. The PHQ-2 assesses depressive mood and loss of interest/pleasure, which are essential items for the diagnosis of DSM-IV MDD. As in the PHQ-9, each item of the PHQ-2 is scored from 0 to 3, and the total score ranges from 0 to 6. They confirmed that when the PHQ-2 score ≥ 3 , major depression could be detected with a sensitivity of 83% and a specificity of 92%.¹² We used the Korean version of the PHQ-9 translated by Park et al.²⁶ in 2010, who demonstrated the reliability and validity of the Korean version of the PHQ-9 in 86 psychiatric outpatients at a university hospital.

EQ-5D

The EQ-5D was developed by the Euro Quality of Life Group as a tool to measure health-related quality of life.²⁷ It consists of five areas that assess health status: mobility, self-care, usual activity, pain/discomfort, and anxiety/depression. Each item is rated on three levels: 1, 2, and 3 represent no problems, some problems, and extreme problems, respectively, and the EQ-5D index can be obtained using the formula.²⁸ A visual analog scale (EQ-VAS) is also included in the EQ-5D. Current

health is displayed on a vertical line drawn on a scale from 0 (worst imaginable health) to 100 (best imaginable health) on the EQ-VAS. The Korean version of the EQ-5D has been developed and its reliability and validity have been demonstrated in several clinical populations.²⁹

Statistical analysis

Sociodemographic characteristics were checked descriptively by dividing them into the total, those diagnosed with

Table 1. Demographic characteristics of the study sample

Characteristics	Total (N=6,022)	Major depression, diagnosed (N=150)	Major depression, undiagnosed (N=5,867)
Age (yr)	47.88±15.38	52.47±15.39	47.76±15.37
18–29 yr	838 (13.9)	19 (12.7)	819 (14.0)
30–39 yr	1,191 (19.8)	12 (8.0)	1,179 (20.1)
40–49 yr	1,175 (19.5)	25 (16.7)	1,150 (19.6)
50–59 yr	1,193 (19.8)	38 (25.3)	1,155 (19.7)
60–69 yr	1,025 (17.0)	31 (20.7)	994 (16.9)
≥70 yr	595 (9.9)	25 (16.7)	570 (9.7)
Sex			
Male	2,308 (38.3)	31 (20.7)	2,277 (38.8)
Female	3,714 (61.7)	119 (79.3)	3,590 (61.2)
Education			
No education	234 (3.9)	17 (11.3)	217 (3.7)
1–6 yr	901 (15.0)	44 (29.3)	857 (14.6)
7–9 yr	714 (11.9)	24 (16.0)	690 (11.8)
10–12 yr	2,009 (33.4)	34 (22.7)	1,975 (33.7)
≥12 yr	2,159 (35.9)	31 (20.7)	2,128 (36.3)
Employment			
Full time	3,924 (65.4)	86 (57.3)	3,838 (65.6)
Part time	386 (6.4)	8 (5.3)	378 (6.5)
Unemployed	1,693 (28.2)	56 (37.3)	1,637 (28.0)
Income			
<1,250 USD	2,136 (44.7)	79 (73.1)	2,057 (44.0)
1,250–2,500 USD	1,152 (24.1)	10 (9.3)	1,142 (24.4)
>2,500 USD	1,495 (31.3)	19 (17.6)	1,476 (31.6)
Marital status			
Married	3,914 (65.1)	80 (53.3)	3,834 (65.4)
Separated/divorced/widowed	1,024 (17.0)	49 (32.7)	975 (16.6)
Single	1,075 (17.9)	21 (14.0)	1,054 (18.0)
Habitation area			
Urban	4,685 (77.9)	110 (73.3)	4,575 (78.0)
Rural	1,332 (22.1)	40 (26.7)	1,292 (22.0)
PHQ-9 score	2.36±3.85	12.93±8.05	2.08±3.25
PHQ-2 score	0.63±1.14	3.41±2.01	0.56±1.01

Values are presented as mean±standard deviation or number (%). USD, United States dollar; PHQ-9, Patient Health Questionnaire-9; PHQ-2, Patient Health Questionnaire-2

major depression, and those undiagnosed with major depression. Cronbach's α coefficient was computed to ascertain internal consistency and was recalculated after each item of the PHQ-9 was deleted individually. Pearson's correlation coefficients were used to examine the relationships between the PHQ-9 and the EQ-5D for construct validity of PHQ-9. The sensitivity, specificity, positive predictive value (PPVs), negative predictive value (NPVs), positive likelihood ratio (LR+), and negative likelihood ratio (LR-) of the PHQ-9 and PHQ-2 were calculated using ROC analysis. Using ROC analysis, the optimal cutoff score was calculated by comparing the PHQ-9 and PHQ-2 with the 2-week major depression diagnosis of the K-CIDI. All statistical analyses were performed using Statistical Package for the Social Sciences (IBM SPSS version 26.0; IBM Corp., Armonk, NY, USA) for Windows. A $p < 0.05$ was considered statistically significant.

RESULTS

Sociodemographic characteristics and scores of PHQ-9 and PHQ-2

Sociodemographic and clinical characteristics are shown in Table 1, divided according to major depression as diagnosed using the K-CIDI. Among the participants included in the analysis, 2,308 (38.3%) were male and 3,714 (61.7%) were female. The mean patient age was 47.88 years. The number of participants diagnosed with major depression using the K-CIDI was 150 (2.5%) and the number of undiagnosed participants was 5,867. The mean scores of PHQ-9 and PHQ-2 in the major depression group were 12.93 (± 8.05) and 3.41 (± 2.01), respectively, and in the undiagnosed group, 2.08 (± 3.25) and 0.56 (± 1.01), respectively.

Reliability and item analysis

The Cronbach's α coefficient of PHQ-9 and PHQ-2 were 0.882 and 0.807, respectively. Table 2 shows that all items of the PHQ-9 were significantly and positively associated with the total PHQ-9 score. Cronbach's α coefficient did not decrease, even when each item was deleted individually from the PHQ-9.

Validity analysis

To confirm the construct validity of the PHQ-9 and PHQ-2, a correlation test with the EQ-5D was performed. The results are summarized in Table 3. All five items of the EQ-5D descriptive system, EQ-5D index, and EQ-VAS were significantly correlated with the PHQ-9 and PHQ-2. The EQ-5D anxiety/depression category had the highest correlation with the PHQ-9 and PHQ-2 among the items of the EQ-5D, with correlation coefficients of 0.471 ($p < 0.001$) and 0.451 ($p < 0.001$),

Table 2. Corrected item-total correlations and Cronbach's α when an item was deleted from the PHQ-9

Item of the PHQ-9	Corrected item-total correlations	Cronbach's α when an item was deleted
Depressed mood	0.710	0.862
Loss of interest/pleasure	0.720	0.861
Insomnia or hypersomnia	0.622	0.870
Poor appetite or overeating	0.622	0.872
Psychomotor agitation or retardation	0.603	0.874
Fatigue	0.613	0.875
Feeling worthless or excessive/inappropriate guilt	0.670	0.866
Decreased concentration	0.643	0.871
Thoughts of death/suicide	0.628	0.872

PHQ-9, Patient Health Questionnaire-9

Table 3. Correlations between depression (PHQ-9, PHQ-2) and health-related quality of life (EQ-5D)

Variable	Correlation coefficient	p
PHQ-9		
Mobility	0.191	<0.001
Self-care	0.089	<0.001
Usual activities	0.210	<0.001
Pain/discomfort	0.233	<0.001
Anxiety/depression	0.471	<0.001
EQ-5D index	0.311	<0.001
EQ-5D VAS	-0.317	<0.001
PHQ-2		
Mobility	0.171	<0.001
Self-care	0.088	<0.001
Usual activities	0.194	<0.001
Pain/discomfort	0.211	<0.001
Anxiety/depression	0.451	<0.001
EQ-5D index	0.288	<0.001
EQ-5D VAS	-0.272	<0.001

PHQ-9, Patient Health Questionnaire-9; PHQ-2, Patient Health Questionnaire-2; EQ-5D, Euro Quality of life-5 dimension; VAS, visual analogue scale

respectively. The correlation coefficients of the EQ-5D index with the PHQ-9 and PHQ-2 were 0.311 ($p < 0.001$) and 0.288 ($p < 0.001$), respectively, and the correlation coefficients of the EQ-VAS with the PHQ-9 and PHQ-2 were -0.317 ($p < 0.001$) and -0.272 ($p < 0.001$), respectively.

ROC analysis

The results of the PHQ-9 and PHQ-2 ROC analysis are

Table 4. Characteristics of cutoff scores for the PHQ-9 and PHQ-2

Depression score	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Positive likelihood ratio	Negative likelihood ratio
PHQ-9						
≥3	93.2	71.5	7.7	99.8	3.3	0.09
≥4	91.9	78.6	9.9	99.7	4.3	0.10
≥5	89.9	84.1	12.6	99.7	5.6	0.12
≥6	81.8	88.4	15.3	99.5	7.1	0.21
≥7	73.0	91.4	17.8	99.3	8.5	0.30
PHQ-2						
≥1	92.0	68.8	7.1	99.7	3.0	0.12
≥2	85.3	83.2	11.6	99.5	5.1	0.18
≥3	56.7	96.3	28.1	98.9	15.1	0.45

PHQ-9, Patient Health Questionnaire-9; PHQ-2, Patient Health Questionnaire-2

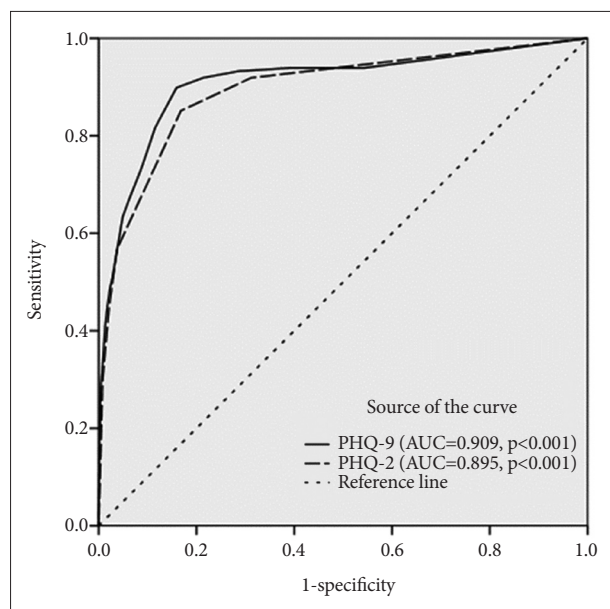


Figure 1. Receiver operating characteristic curves of the Patient Health Questionnaire-9 (PHQ-9) and Patient Health Questionnaire-2 (PHQ-2). AUC, area under the curve.

shown in Table 4, and the ROC curves of the PHQ-9 and PHQ-2 are shown in Figure 1. The area under the curve (AUC) of PHQ-9 was 0.909 (95% confidence interval [CI]=0.879–0.939; standard error [SE]=0.015; $p<0.001$). A cutoff score of 5 indicated the highest Youden's index. At this cutoff score, the sensitivity and specificity of the PHQ-9 were 89.9% and 84.1%, respectively, with a PPV of 12.6%, NPV of 99.7%, LR+ of 5.6, and LR- of 0.12. The AUC of PHQ-2 was 0.895 (95% CI=0.864–0.925; SE=0.015; $p<0.001$). A cutoff score of 2 indicated the highest Youden's index. At this cutoff score, the sensitivity and specificity of the PHQ-2 were 85.3% and 83.2%, respectively, with a PPV of 11.6%, NPV of 99.5%, LR+ of 5.1, and LR- of 0.18.

DISCUSSION

We analyzed KECA-2011 data collected from participants selected as representative of Korean society through a multi-stage and stratified cluster sampling method to confirm the validity and cutoff scores of the PHQ-9 and PHQ-2 in screening for MDD in the general population. We examined the reliability of the PHQ-9 and PHQ-2, conducted construct validity, calculated sensitivity and specificity, and obtained the AUC and optimal cutoff scores through ROC analysis.

The Korean versions of the PHQ-9 and PHQ-2 showed high reliability and validity. Cronbach's α coefficient of the PHQ-9 and PHQ-2 were 0.882 and 0.807, respectively, indicating relatively high internal consistency, according to Nunnally's guideline which suggests $\alpha \geq 0.70$ is an acceptable value for internal consistency.³⁰ These results are consistent with those of a previous study focusing on the general population, although few such studies have been conducted. The Cronbach's α of the PHQ-9 for the aforementioned study that confirmed the normative data of the PHQ-9 for the general population was 0.79 in Korea,¹⁸ 0.82 in Hong Kong,³¹ and 0.87 in Germany.³² The Cronbach's α of the PHQ-2 for the general population in Hong Kong was 0.76.³³ In addition, to confirm the construct validity of the PHQ-9 and PHQ-2, we checked their correlations with the EQ-5D, and the correlation coefficient between the EQ-5D index and the PHQ-9 and PHQ-2 were confirmed to be 0.311 and 0.288, respectively. The highest correlation was with the anxiety/depression category of the EQ-5D (PHQ-9: $r=0.471$, $p<0.001$; PHQ-2: $r=0.451$, $p<0.001$), which is similar to the results of a previous study that confirmed construct validity through the correlation between the PHQ-9 and EQ-5D.¹⁸

In our study, the appropriate cutoff scores for the PHQ-9 and PHQ-2 in the general population were 5 and 2, respec-

tively. In a previous systematic review, the cutoff scores of the PHQ-9 varied between 5 and 15 points depending on the study group and country, but were generally around 10 points.¹³ Our study showed lower cutoff values than other studies. These results are common in East Asian studies that analyzed the cutoff score of the PHQ-9, which usually shows relatively low cutoff scores of 5–7.^{34–36} In East Asian cultures, reporting mental health difficulties is sometimes viewed as acknowledgment of personal weakness,³⁷ and mental disorders are often viewed as stemming from personal or family conflicts.^{38,39} Therefore, East Asians may be reluctant to tick off depression-related symptoms on self-report questionnaires because of this cultural stigma.^{34,36} This tendency was also noted in the present study, which may have resulted in a low cutoff score. In addition, the low cutoff score might have been influenced by the present study being based on community-based data. The cutoff score for the PHQ-9 obtained from Korean psychiatric outpatient data was 10 points.¹⁷ A previous meta-analysis has shown that the optimal cutoff score for the PHQ-9 varies by population and setting.⁴⁰ Furthermore, a study that used the CIDI in a community-based sample, such as ours, suggested a cutoff score of 8,⁴¹ whereas a study that used a structured diagnostic interview in a psychiatric sample suggested a cutoff score of 13.⁴² The PHQ-2 has few items; therefore, there is no wide range of cutoffs across studies. However, a previous review of the PHQ-2 found that 19 out of 21 studies suggested a cutoff score of 3.¹⁴ Like the PHQ-9, the PHQ-2 appears to be influenced by culture and setting, resulting in lower scores than in other studies. A Taiwanese study suggested a PHQ-2 cutoff score of 2,⁴³ and a study that used the CIDI also indicated a cutoff score of 2.⁴⁴ The present study's results highlight that different cutoffs should be applied depending on culture, population, and setting.

The cutoff score of 5 for the PHQ-9 has appropriate sensitivity and specificity (89.9%, 84.1%, respectively); however, the low PPV (12.6%) may be an issue. The PHQ-2 exhibits a similar trend (cutoff score of 2, sensitivity of 85.3%, specificity of 83.2%, and PPV of 11.6%). PPV can express the probability of the correctness of a test result; however, it is known that PPV is highly affected by prevalence.⁴⁵ Therefore, the low PPV in this study was likely due to the fact that the prevalence of major depression was 2.5%. Our results showed an LR+ of 5.6 and LR- of 0.12 for the PHQ-9, and an LR+ of 5.1 and LR- of 0.18 for the PHQ-2. In general, when the LR+ is between 5 and 10, it is considered to provide moderate evidence in ruling out a diagnosis. Similarly, when the LR- is between 0.1 and 0.2, it is considered to provide moderate evidence in ruling out a diagnosis.⁴⁶ Therefore, the probabilities of PHQ-9 and PHQ-2 in this study were appropriate.

Our study has several limitations. First, major depression

was diagnosed using the K-CIDI, not by a medical professional. However, the results showed moderate to good concordance at the individual level between the CIDI and clinician-administered structured interviews.⁴⁷ Second, although a large number of subjects were investigated, the number of major depression cases was small, which affected not only the sensitivity and specificity but also the PPV. Third, as this was a cross-sectional study, test-retest reliability was not assessed, and further evaluation is required. However, the PHQ-9 and PHQ-2 were generally consistent.^{26,33}

In conclusion, the results of this study suggest that the Korean versions of the PHQ-9 and PHQ-2 are valid screening tools for major depression in Korea, with the optimal cutoff scores of PHQ-9 and PHQ-2 for screening major depression in a representative Korean general population using data from the Korean National Epidemiological Study (KECA-11) are ≥ 5 for the PHQ-9 (sensitivity 89.9%, specificity 84.1%) and ≥ 2 for PHQ-2 (sensitivity 85.3%, specificity 83.2%).

Availability of Data and Material

The datasets generated or analyzed during the study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Author Contributions

Conceptualization: Sanghyup Jung, Bong-Jin Hahm. Data curation: Jee Hoon Sohn, Su Jeong Seong, Jin Pyo Hong. Investigation: Bong-Jin Hahm, Chan-Woo Yeom. Methodology: Minah Kim, Sanghyup Jung, Bong-Jin Hahm. Software: Sanghyup Jung. Writing—original draft: Minah Kim. Writing—review & editing: Minah Kim, Sanghyup Jung, Jee Eun Park, Byung-Soo Kim, Sung Man Chang, Bong-Jin Hahm, Chan-Woo Yeom.

ORCID iDs

Minah Kim	https://orcid.org/0009-0006-9278-0279
Sanghyup Jung	https://orcid.org/0000-0002-2098-8410
Jee Eun Park	https://orcid.org/0000-0001-6097-3275
Jee Hoon Sohn	https://orcid.org/0000-0001-7206-460X
Su Jeong Seong	https://orcid.org/0000-0002-6789-8877
Byung-Soo Kim	https://orcid.org/0000-0001-6823-4813
Sung Man Chang	https://orcid.org/0000-0001-7077-7171
Jin Pyo Hong	https://orcid.org/0000-0001-5384-2605
Bong-Jin Hahm	https://orcid.org/0000-0002-2366-3275
Chan-Woo Yeom	https://orcid.org/0000-0003-2265-3291

Funding Statement

This study was supported by a research grant from the Korean Ministry of Health and Welfare.

REFERENCES

1. World Health Organization. Depression and other common mental disorders: global health estimates [Internet]. Available at: <https://apps.who.int/iris/handle/10665/254610>. Accessed October 14, 2022.
2. Lépine JP, Briley M. The increasing burden of depression. *Neuropsychiatr Dis Treat* 2011;7(Suppl 1):3-7.
3. Cho MJ, Seong SJ, Park JE, Chung IW, Lee YM, Bae A, et al. Prevalence and correlates of DSM-IV mental disorders in South Korean adults:

- the Korean epidemiologic catchment area study 2011. *Psychiatry Investig* 2015;12:164-170.
4. Hasin DS, Sarvet AL, Meyers JL, Saha TD, Ruan WJ, Stohl M, et al. Epidemiology of adult DSM-5 major depressive disorder and its specifiers in the United States. *JAMA Psychiatry* 2018;75:336-346.
 5. Gutiérrez-Rojas L, Porrás-Segovia A, Dunne H, Andrade-González N, Cervilla JA. Prevalence and correlates of major depressive disorder: a systematic review. *Braz J Psychiatry* 2020;42:657-672.
 6. Statistics Korea. Annual report on the cause of death statistics, 2020. [Internet]. Available at: https://kostat.go.kr/board.es?mid=a10301060200&bid=218&act=view&list_no=403046. Accessed October 14, 2022.
 7. Cavanagh JT, Carson AJ, Sharpe M, Lawrie SM. Psychological autopsy studies of suicide: a systematic review. *Psychol Med* 2003;33:395-405.
 8. Henriksson MM, Aro HM, Marttunen MJ, Heikkinen ME, Isometsä E, Kuoppasalmi KI, et al. Mental disorders and comorbidity in suicide. *Am J Psychiatry* 1993;150:935-940.
 9. Mann JJ, Apter A, Bertolote J, Beautrais A, Currier D, Haas A, et al. Suicide prevention strategies: a systematic review. *JAMA* 2005;294:2064-2074.
 10. Spitzer RL, Kroenke K, Williams JB; the Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *JAMA* 1999;282:1737-1744.
 11. Kroenke K. Enhancing the clinical utility of depression screening. *CMAJ* 2012;184:281-282.
 12. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care* 2003;41:1284-1292.
 13. Costantini L, Pasquarella C, Odone A, Colucci ME, Costanza A, Serafini G, et al. Screening for depression in primary care with Patient Health Questionnaire-9 (PHQ-9): a systematic review. *J Affect Disord* 2021;279:473-483.
 14. Manea L, Gilbody S, Hewitt C, North A, Plummer F, Richardson R, et al. Identifying depression with the PHQ-2: a diagnostic meta-analysis. *J Affect Disord* 2016;203:382-395.
 15. Han C, Jo SA, Kwak JH, Pae CU, Steffens D, Jo I, et al. Validation of the Patient Health Questionnaire-9 Korean version in the elderly population: the Ansan Geriatric study. *Compr Psychiatry* 2008;49:218-223.
 16. Seo JG, Park SP. Validation of the Patient Health Questionnaire-9 (PHQ-9) and PHQ-2 in patients with migraine. *J Headache Pain* 2015;16:65.
 17. Shin C, Lee SH, Han KM, Yoon HK, Han C. Comparison of the usefulness of the PHQ-8 and PHQ-9 for screening for major depressive disorder: analysis of psychiatric outpatient data. *Psychiatry Investig* 2019;16:300-305.
 18. Shin C, Ko YH, An H, Yoon HK, Han C. Normative data and psychometric properties of the Patient Health Questionnaire-9 in a nationally representative Korean population. *BMC Psychiatry* 2020;20:194.
 19. Ministry of Health and Welfare. The epidemiological survey of mental disorders in Korea. Sejong: Ministry of Health and Welfare; 2011.
 20. Statistics Korea. 2010 population and housing census [Internet]. Available at: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN1002&vw_cd=&list_id=&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=. Accessed March 25, 2011.
 21. Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, et al. The composite international diagnostic interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988;45:1069-1077.
 22. Cho MJ, Hahm BJ, Suh DW, Hong JP, Bae JN, Kim JK, et al. Development of a Korean version of the Composite International Diagnostic Interview (K-CIDI). *J Korean Neuropsychiatr Assoc* 2002;41:123-137.
 23. World Health Organization. Procedures for the development of new language versions of the WHO Composite International Diagnostic Interview (WHO-CIDI). Geneva: World Health Organization; 1998.
 24. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16:606-613.
 25. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, fourth edition, text revision (DSM-IV-TR). Washington, DC: American Psychiatric Publishing; 2010.
 26. Park SJ, Choi HR, Choi JH, Kim KW, Hong JP. Reliability and validity of the Korean version of the Patient Health Questionnaire-9 (PHQ-9). *Anxiety Mood* 2010;6:119-124.
 27. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001;33:337-343.
 28. Oh HS. Important significant factors of health-related quality of life (EQ-5D) by age group in Korea based on KNHANES (2014). *J Korean Data Inf Sci Soc* 2017;28:573-584.
 29. Kim MH, Cho YS, Uhm WS, Kim S, Bae SC. Cross-cultural adaptation and validation of the Korean version of the EQ-5D in patients with rheumatic diseases. *Qual Life Res* 2005;14:1401-1406.
 30. Nunnally JC. *Psychometric theory* (2nd ed). New York: McGraw-Hill; 1978.
 31. Yu X, Tam WW, Wong PT, Lam TH, Stewart SM. The Patient Health Questionnaire-9 for measuring depressive symptoms among the general population in Hong Kong. *Compr Psychiatry* 2012;53:95-102.
 32. Kocalevent RD, Hinz A, Brähler E. Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry* 2013;35:551-555.
 33. Yu X, Stewart SM, Wong PT, Lam TH. Screening for depression with the Patient Health Questionnaire-2 (PHQ-2) among the general population in Hong Kong. *J Affect Disord* 2011;134:444-447.
 34. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry* 2014;36:539-544.
 35. Inagaki M, Ohtsuki T, Yonemoto N, Kawashima Y, Saitoh A, Oikawa Y, et al. Validity of the Patient Health Questionnaire (PHQ)-9 and PHQ-2 in general internal medicine primary care at a Japanese rural hospital: a cross-sectional study. *Gen Hosp Psychiatry* 2013;35:592-597.
 36. Sung SC, Low CC, Fung DS, Chan YH. Screening for major and minor depression in a multiethnic sample of Asian primary care patients: a comparison of the nine-item Patient Health Questionnaire (PHQ-9) and the 16-item Quick Inventory of Depressive Symptomatology-Self-Report (QIDS-SR16). *Asia Pac Psychiatry* 2013;5:249-258.
 37. Jang Y, Chiriboga DA, Okazaki S. Attitudes toward mental health services: age-group differences in Korean American adults. *Aging Ment Health* 2009;13:127-134.
 38. Abdullah T, Brown TL. Mental illness stigma and ethnocultural beliefs, values, and norms: an integrative review. *Clin Psychol Rev* 2011;31:934-948.
 39. Park JE, Cho SJ, Lee JY, Sohn JH, Seong SJ, Suk HW, et al. Impact of stigma on use of mental health services by elderly Koreans. *Soc Psychiatry Psychiatr Epidemiol* 2015;50:757-766.
 40. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *CMAJ* 2012;184:E191-E196.
 41. Kiely KM, Butterworth P. Validation of four measures of mental health against depression and generalized anxiety in a community based sample. *Psychiatry Res* 2015;225:291-298.
 42. Beard C, Hsu KJ, Rifkin LS, Busch AB, Björgvinsson T. Validation of the PHQ-9 in a psychiatric sample. *J Affect Disord* 2016;193:267-273.
 43. Liu SI, Yeh ZT, Huang HC, Sun FJ, Tjung JJ, Hwang LC, et al. Validation of Patient Health Questionnaire for depression screening among primary care patients in Taiwan. *Compr Psychiatry* 2011;52:96-101.
 44. Arroll B, Goodyear-Smith F, Crengle S, Gunn J, Kerse N, Fishman T, et al. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *Ann Fam Med* 2010;8:348-353.
 45. Mischlinger J, Schernhammer E. A common trap of diagnostic tests: disease prevalence and positive predictive value. *Wien KlinWochen-*

schr 2017;129:583-584.

46. Kent P, Hancock MJ. Interpretation of dichotomous outcomes: sensitivity, specificity, likelihood ratios, and pre-test and post-test probability. *J Physiother* 2016;62:231-233.
47. Haro JM, Arbabzadeh-Bouchez S, Brugha TS, De Girolamo G, Guyer

ME, Jin R, et al. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys. *Int J Methods Psychiatr Res* 2006;15:167-180.