



The Validation of Indonesian Version of Patient Health Questionnaire-9

Cut N. Dian, Almeida Effendy*, Mustafa M. Amin

Department of Psychiatry, Faculty of Medicine, Universitas Sumatera Utara, Medan, Indonesia

Abstract

Edited by: Branislav Filipović
Citation: Dian CN, Effendy E, Amin MM. The Validation of Indonesian Version of Patient Health Questionnaire-9. *Open-Access Maced J Med Sci.* 2022 Mar 28; 10(T7):193-198. <https://doi.org/10.3889/oamjms.2022.9293>

Keywords: Indonesian version of Patient Health Questionnaire; Depression level; Medical students; Validation Patient Health Questionnaire-9; Mini ICD 10
***Correspondence:** Almeida Effendy, Department of Psychiatry, Faculty of Medicine, Universitas Sumatera Utara, Indonesia. E-mail: elmeida.effendi@usu.ac.id

Received: 07-Mar-2022

Revised: 15-Mar-2022

Accepted: 18-Mar-2022

Copyright: © 2022 Cut N. Dian, Almeida Effendy, Mustafa M. Amin

Funding: This research did not receive any financial support.

Competing Interest: The authors have declared that no competing interest exists

Open Access: This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

BACKGROUND: The level of depression among college students differs to each other based on their activities, particularly in terms of their educational backgrounds. However, the national depression level among these students needs to be investigated further due to the differences of social and cultural environments.

AIM: This research aims to investigate the level of depression among medical students in the range of ages 17–22.

METHODS: The study performed the Indonesian version of Patient Health Questionnaire (PHQ-9) which was assessed to 500 medical students. The Indonesian version of PHQ-9 was prepared through forward-backward translation. The statistical analysis such as the receiver operating characteristic (ROC), the internal concurrent reliability, and concurrent validity was performed.

RESULTS: Our findings have confirmed the correlation of Indonesian version of PHQ-9 to the original version. The ROC analysis has suggested that the area under the curve (AUC) was accounted for 92.0, whereas the cutoff value was 5.5. With sensitivity and specificity, respectively, were 0.907 and 0.865, the results could confirm the depression level that equals 5.5 or above.

CONCLUSION: The PHQ-9 translated version could be used as an evaluation to determine the depression level to the sample of medical students with acceptable validation results.

Introduction

Depression, one of the common mental disorders in low-to-middle income nations, has been prioritized since 2008 by the World Health Organization (WHO) [1]. The priority is due to the prevalence that has been associated including genders, ages, educational background, marital status, and financial income [2]. As the society in the developing countries has been struggling in terms of educational background and financial income in which both of them are social issue, in Indonesia, the death rate due to depression has been reported in between 15 and 29 for 3.6%; the most common cause for this rate was solitariness (7%), anxiety (5%), and unsocialized (3%) [1]. Therefore, the investigation of depression symptoms for Indonesians is essential to be performed.

Physically speaking, depression can be observed through several symptoms. There are three common physical and behavioral appearances, and a little desire of doing something which could be followed by sleeping disorder and losing energy is the real example of depression [3]. Educational and cultural problems such as family and colleagues may have caused lack of motivation to do some works [2],

and the ages that are commonly found to experience depression are mostly in younger ages from 15 to 29 years of age [4], [5], [6]. Furthermore, a survey has reported that depression is one of the main causes for disability which accounted for 11% [7]. Hence, it is important to understand the prevalence of depression.

Mental disorder, particularly in medical students, is important for the country to prevent its domino effects. As the developing countries focus on how to improve the health-care facilities as well as human resources, deep understanding in providing health-care services and health policies must be considered. However, since 2018, no survey which presented the national depression level in Indonesia has not been reported until it was found that 15% of adult females and 24.4% of college students in Jakarta to be suffering depression [7], [8]. Consequently, the depressed behaviors are not considered as special consideration due to medical treatment, as the drugs that have been prescribed took only 9%, indicating a little use of medications [9].

Clinical depression assessment focuses on early screening and diagnosis through questionnaire and interviews. Certain instruments of assessment have been in Indonesia, including the Mini International Interview Version ICD 10 (Mini ICD 10) to detect the

Beck's Depression Inventory, and the Patient Health Questionnaire (PHQ-9) [10], [11]. Although the PHQ-9 is favorable due to its benefits such as short-term assessment (depression level in recent 2 weeks) [12], its ability to detect the depression level in the range of ages 18–66 and above, and brief questionnaire statements [13], to the best of our knowledge, the validation of this instrument in Indonesia, particularly for medical students, need to be investigated further. Moreover, this study aims to design PHQ-9 Indonesian version by adopting the original version.

Research Methodologies

Participants

The study recruited samples from the medical students. The recruitment was performed in Faculty of Medicine, Universitas Sumatera Utara, Medan, Indonesia, whereas the samples who were the students were in between 17 and 22 years of age. Research subjects were collected through non-probability sampling which was the purposive sampling.

Due to its purposiveness, the sample subjects were recruited based on certain criteria. The inclusion criteria are medical students who have been studying in the Faculty of Medicine, Universitas Sumatera Utara, with ages in between 17 and 22 years of old, with both males and females. On the other hand, the exclusion criteria are subjects who refused to participate in this study as well as not completing the assessment that has been given.

As this study involved with human subjects, this study has been approved by the Health Research Ethical Committee of the Faculty of Medicine, Universitas Sumatera Utara, with the given approval number is No: 67/UN5.2.1.1/PPM/2019. The ethical clearance was also received which is based on the Nuremberg Code and 1964 Helsinki Declaration. Thus, the written informed consent was provided to the participants, and they were requested to sign as the indicator of agreement to be subject samples in this study.

PHQ-9 translation framework

The questionnaire was prepared by following the method proposed by Greco, Wallpo, and Eastridge [14]. In brief, there were four steps that are required to be done. First, the original PHQ-9 issued in English was translated in forward method. The process of translation employed two translators who have been chosen accordingly. These bilingual translators were chosen due to their ability in using English and translating the questionnaire into Bahasa Indonesia [29], [30], [31]. Then, the results of translation

were discussed to combine and determine the similarity of the two translations altogether by the authors and the translators. Thus, forward translation was obtained [14].

Next, the backward translation was carried out from the forward translation. In this step, the previous translators were requested to translate the questionnaire into their origin language (mother tongue) [14], [15], [16]. To avoid bias understanding, the backward translator should have understood the concept of the PHQ-9 [17]. Finally, these backward translations were discussed again in seeking the similarity. Thus, the questionnaire that has been obtained was considered as the Indonesian version of PHQ-9. The following Figure 1 displays the flowchart of forward-backward translation.

Statistical measurements of Indonesian version of PHQ-9

In validating the translated version and finding the reliability the final translation, the statistical analysis was performed in several parameters. First, the concurrent validity was measured to find the correlation between the original version and the translated version through Pearson's correlation [15]. Next, the internal consistency reliability for the translated version was statistically analyzed to verify the reliability of the translated version of PHQ-9, and the analysis conducted to seek the reliability was done by Cronbach's

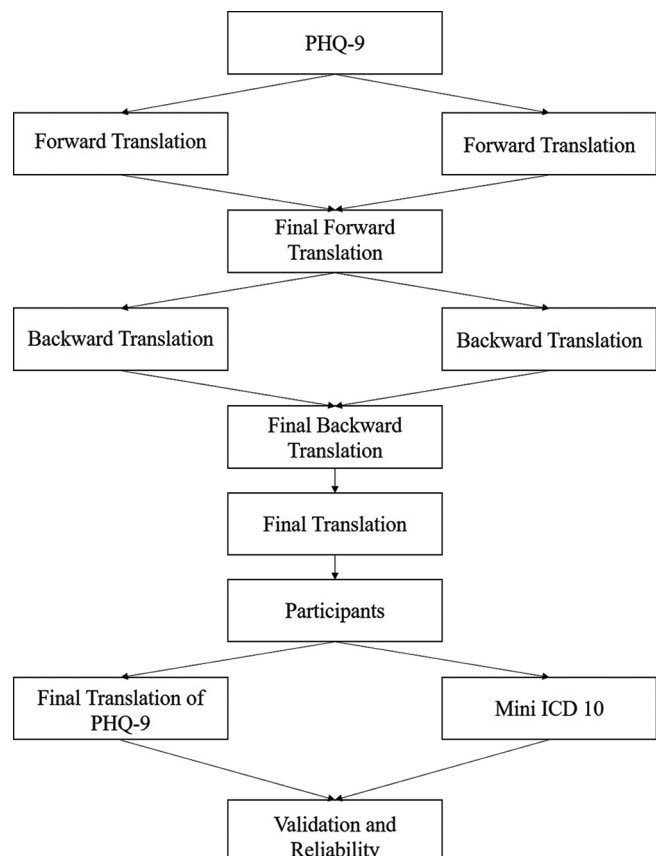


Figure 1: The flowchart of PHQ-9 Assessment to the Depression Level of The Medical Students

alpha correlation coefficient. Finally, the translated version was analyzed to determine its diagnostic ability with receiving operating characteristics (ROC). This analysis was performed under the measurement of area under the curve (AUC) of the ROC within the graph.

The assessment of PHQ-9

The final translation of the PHQ-9 was performed in self-questioner. The subjects were also invited to do the Mini ICD 10 to compare the results. The following Figure 1 displays the schematic diagram of the assessment.

Results

Translation version of PHQ-9

The translated version of PHQ-9 was carried out by two translators through forward-backward translation. The following Table 1 shows the comparison between the English and translation version of PHQ-9 based on the questions.

Table 1: The comparison of the English and translated version of PHQ-9

Items	English version	Translation version
1.	Kurang tertarik atau kurang suka dalam melakukan apapun	A little desire or pleasure in doing something
2.	Merasa murung, pilu, atau putus asa	Depressed, difficult, or hopeless
3.	Sulit untuk tidur atau bertahan tidur, atau terlalu banyak tidur	Difficulty in sleeping, or sleeping too much
4.	Merasa Lelah atau kurang bertenaga	Tired or weak
5.	Kurang nafsu makan atau makan terlalu banyak	Lack of appetite or over-eating
6.	Merasa buruk akan diri sendiri – atau merasa bahwa anda adalah orang yang gagal atau telah mengecewakan diri sendiri atau keluarga Anda	Discouraged or feel like a failure or disappoint yourselves or your family
7.	Sulit berkonsentrasi pada sesuatu, misalnya membaca koran atau menonton televisi	Cannot concentrate on something such as reading a newspaper or watching TV
8.	Bergerak atau berbicara sangat lambat sehingga orang lain memperhatikannya. Atau sebaliknya sedemikian resah, dan gelisah sehingga Anda bergerak jauh lebih banyak dari biasanya	To slow to talk or move until people realize it or vice versa
9.	Memikirkan bahwa Anda lebih baik mati atau melukai diri Anda sendiri dengan sesuatu cara	The desire to hurt yourself or kill yourself

These questions were used to assess the depression of research subjects. To determine the validity and reliability, the subjects were chosen accordingly through purposive sampling technique. As the subjects were medical students in Universitas Sumatera Utara aged from 17 to 22 years, the assessment was carried out after class dismissed in which the informed consent was agreed by them. Thus, the participants were also invited to fill the MINI ICD 10.

Demographical characteristics of participants

The participants were medical students at Universitas Sumatera Utara. In Indonesia, the medical students are students who have not been completed

their undergraduate studies, in which the completion of the study is appreciated with the title “Bachelor of Medicine.” Thus, to be a general practitioner, they must follow 2 years of medical training which is known as coassistant.

In this study, 500 medical students were recruited. As the curriculum for the students requires 4 years of study, then students who were 17–22 years of age were invited to involve in this study. The following Table 2 displays the characteristics of the participants.

Table 2: The demographical characteristics of participants

Variables	n (%)
Sex	
Male	159 (31.8)
Female	341 (68.2)
Ages	
17 years of old	14 (2.8)
18 years of old	141 (28.2)
19 years of old	107 (21.4)
20 years of old	91 (18.2)
21 years of old	46 (9.2)
22 years of old	101 (20.2)
Education	
Undergraduate students	500 (100.0)
Occupation	
Medical students	500 (100.0)
Marital status	
Single	500 (100.0)
University entrance year	
2016	86 (17.2)
2017	128 (25.6)
2018	177 (35.4)
2019	109 (21.8)
PHQ-9 score	
Median (maximum-minimum)	8 (0–26)
MINI ICD 10	
Yes	389 (77.8)
No	111 (22.2)

From Table 2, 500 participants were recruited based on the inclusion criteria. It was obtained that the female participants had the higher population with 341 (68.2%), whereas the male students were just above 30%. As the participants were medical students from the interval of age 17–22, the oldest participants were 22 years of age with 20% of the population, while 14 youngest subjects with were 17 years of old. However, in terms of the highest population, the most populated participants were students who were 18 years of-old, and both top two and three were dominated by 19 and 20 years of age students. These results also confirmed that most of the participants are in the 2nd, 3rd, and 4th year of study, which imply to the long duration of study. Interestingly, the MINI ICD assessment showed majority of participants experience depression.

Concurrent validity

To confirm the validity of the Indonesian version of PHQ-9, two statistical tests were performed which were normality test and data transformation. The following Tables 3 and 4 display the normality and transformation results test, respectively.

In the normality tests, the result which was $p < 0.05$ has confirmed the abnormality of the data. By transforming the data into log 10 function, the data have confirmed the abnormal distribution (Table 4). These two tests resulted in non-conclusive situation which

leads to the tests through Eta correlation. The following Table 5 shows the linear relationship that was obtained through the Eta correlation.

Table 3: Test of normality of the PHQ-9 Indonesian Version

Tests	Kolmogorov–Smirnov ^a			Shapiro–Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
PHQ-9 Score	0.117	500	0.000	0.952	500	0.000
Mini ICD 10 Score	481	500	0.000	0.513	493	0.000

Lilliefors significance correction.

Table 4: Data transformation of PHQ-9 Indonesian version

Tests	Kolmogorov–Smirnov ^a			Shapiro–Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
PHQ-9 Score	0.117	500	0.000	0.952	500	0.000
Mini ICD 10 Score	481	500	0.000	0.513	493	0.000

Lilliefors significance correction.

Table 5: The correlation between PHQ-9 and MINI ICD 10 Indonesian version

Eta	Value	
	Score PHQ-9 dependent	Mini ICD 10
	0.527	0.889

Based on Table 5, provided that the PHQ-9 score is considered as the base for the dependent variable, the correlation coefficient was 0.527. In contrast, when the Mini ICD 10 was considered to be dependent variable, then the correlation coefficient accounted for 0.889.

Internal consistency of reliability

According to Table 6, the reliability has been confirmed by the Cronbach's alpha value which was accounted for 0.885. This value is acceptable as long as all statements were valid. As it is displayed in the corrected item-total item correlation, the value was accounted for above 0.3.

Table 6: The internal consistency of reliability of Indonesian version PHQ-9

Testing variables	Corrected item-total item correlation	Cronbach's alpha if item deleted	Cronbach's alpha
P1	0.816	0.876	0.885
P2	0.823	0.866	
P3	0.790	0.879	
P4	0.787	0.872	
P5	0.795	0.879	
P6	0.807	0.866	
P7	0.812	0.870	
P8	0.841	0.872	
P9	0.863	0.872	

The ROC

The ROC and the AUC were performed to determine the assessment. As it is found in Table 1, the Mini ICD 10 has determined that 389 participants were classified as experiencing depression, while the other 111 subjects were not. Table 7 shows the average score of PHQ-9.

Table 7: The ROC analysis of PHQ-9

	Average	N	AUC	IK 95%	Cutoff	Sensitivity	Specificity
Score PHQ-9	8.77	500	92.0	88.1 – 96.0	5.5	0.907	0.865

Based on statistical analysis, the optimal cutoff

line was ≥ 5.5 with sensitivity of 90.7% and 86.5% of specificity. This implies that the depression has been experienced by subjects with above 5.5 of depression score.

Discussion

This study is conducted to determine the Indonesian version of PHQ-9 due to its original instrument of English. As the process of assessment might have been resulted an invalid score, the concept of PHQ-9 assessment should have been transformed into the language that the participants are involved. On the other hand, the process of translation must be performed in acceptable ways through exact assessment. Thus, an acceptable method to translate the assessment is using the forward-backward translation, and this method has been approved by the WHO as worldwide guidance.

In addition to obtain valid results, the translation of the forward-backward requires at least two translators [18]. These bilingual translators must translate the instruments into their mother tongue, and one of the translators must have been aware about the concept of instruments, while the other is unaware. Hence, in this study, one of the translators has an educational background as a psychiatrist, while the other one has no medical educational background. Subsequently speaking, due to the forward-backward translation, the forward and backward translators must do the translation into the English version from their mother tongue too, respectively.

Based on Table 1, the demographical characteristics of the respondents have shown widely. The respondents were medical students studying at Universitas Sumatera Utara, Medan, Indonesia, with various ages from 17 to 22. From 500 subjects, up to 70% of the respondents were female students, while the rest of a third was male students (31.8%). In a study conducted in 2018, a precaution regarding to the depression has been reported commonly in people who are in between teenagers and young adults, and this level of depression tends to decrease as the aging [5]. Meanwhile, in a survey, it has been reported that 21.8% of people experienced medium to severe depression, and from this survey, mostly the women who were suffering depression were slightly higher than that in men, even though the percentage was no significantly difference, male (21.4%) and female (22.3%) [8]. Thus, the highest prevalence of severe depression only for the women itself is commonly found in the age of 15–19 years of old [5].

In this study, most of the respondents are those who are in the age of 18, 19, and 22. In 2017, a study has reported that the major depression in the USA

was around 7% differently to each aged group. Thus, individual prevalence is different in the level of ages, in which the age of 18–29 could have 3 times higher than elder people (above 60 years of old). As far as gender is concerned, females experience depression higher than that in males (around 1.5–3 times higher), which commonly started in the beginning of teenaged ages [6].

The validity of the PHQ-9 of Indonesian version has been tested by comparing the PHQ-9 results with Mini ICD 10, and the analysis through Eta correlation has confirmed. According to Table 5, the Eta correlation analysis has shown $r = 0.527$, and this shows positive correlation in between PHQ-9 of Indonesian version. However, the positive correlation had medium strength, so as long as the score of PHQ-9 is high, then, the Mini ICD score is also high.

The value of Cronbach's alpha is 0.873 for the PHQ-9 of Indonesian version. George and Mallery in 2003 have provided practical guidance, in which the value of 0.873 is considered to be acceptable due to its interval in between 0.7 and 0.95 [16]. A lower value of alpha appears due to the number of questions and their connection among them. On the other hand, as the alpha's score is high, it implies to the depth analysis of one item in different point of views [19]. In the meantime, the value of corrected item-total item correlation was determined to be above the minimal coefficient correlation, in which the valid number is 0.3. Therefore, the nine statements that have been translated forwardly-backwardly are valid statements.

As the valid statements of the Indonesian version of PHQ-9 have been obtained, the assessment could be performed. Subsequently speaking, our findings that were confirmed through AUC value through ROC method were determined for 92% (Table 7; IK 95%, 88.1–96%) with $p < 0.05$. Statistical views have suggested that the above number of 92% is classified as very good, in which it can be considered as the indicator of diagnosis. Hence, our PHQ-9 Indonesian version was used in diagnosing whether the presence of depression in 100 subjects could be determined, which is 92 subjects. On the other hand, the optimal cutoff which is in ≥ 5.5 with sensitivity of 90.7% and specificity 86.5% has confirmed the minimum score for someone to experience depression. This implies to every sample that has score to be equal or above 5.5 which is diagnosed to be experiencing depression, however, this diagnosis only focused on the medical students at Universitas Sumatera Utara.

Conclusion

The assessment of depressions level which is based on the PHQ-9 may produce different results due different culture. In addition, to the best of our

knowledge, the Indonesian version of PHQ-9 is considered to be acceptable and valid which can be seen from the statistical results in terms of the Eta correlation with values of $r = 0.527$. As the objective of this study is to determine the validity of PHQ-9 level for assessing the depression level of Indonesian medical students in particular for the 1st and 3rd year due to its educational and social needs, demographical factors limited our study. Nevertheless, this Indonesian version of PHQ-9 can be used as one of the assessments levels for students, which certainly in the future could have been used for descriptive study.

References

1. World Health Organisation. Depression and Other Common Mental Disorders. Geneva: World Health Organisation; 2017.
2. Shin C, Kim Y, Park S, Yoon S, Ko YH, Kim YK, *et al*. Prevalence and associated factors of depression in general population of Korea: Results from the Korea national health and nutrition examination survey, 2014. *JKoreanMedSci*. 2017;32(11):1861-9. <https://doi.org/10.3346/jkms.2017.32.11.1861> PMID:28960042
3. Tasman A, Kay J, Lieberman JA, First MB, Riba MB. *Psychiatry*. 4th ed. New York: John Wiley and Sons, Ltd.; 2015.
4. American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorder DSM-IV. Washington, DC: American Psychiatric Publishing; 2005.
5. Peltzer K, Pengpid S. High prevalence of depressive symptoms in a national sample of adults in Indonesia: Childhood adversity, sociodemographic factors and health risk behaviour. *Asian J Psychiatr*. 2018;33:52-9. <https://doi.org/10.1016/j.ajp.2018.03.017>
6. Chen Y, Bennett D, Clarke R, Guo Y, Yu C, Bian Z, *et al*. Patterns and correlates of major depression in Chinese adults: A cross-sectional study of 0.5 million men and women. *Psychol Med*. 2017;47(5):958-70. <https://doi.org/10.1017/S0033291716002889> PMID:27919307
7. Christiani Y, Byles J, Tavener M, Dugdale P. Socioeconomic related inequality in depression among young and middle-adult women in Indonesia's major cities. *J Affect Disord*. 2015;182:76-81. <https://doi.org/10.1016/j.jad.2015.04.042> PMID:25978717
8. Peltzer K, Pengpid S. Depressive symptoms and social demographic, stress and health risk behaviour among university students in 26 low-, middle- and high-income countries. *Int J Psychiatry Clin Pract*. 2015;19(4):259-65. <https://doi.org/10.3109/13651501.2015.1082598> PMID:26265420
9. Riskesdas. Prevalensi Depresi pada Penduduk Umur 15 Tahun ke Atas Menurut Provinsi; 2018.
10. Lecrubier JP, Weiller E, Hergueta T, Amorim P, Bonora LI. MINI Mini International Neuropsychiatric Interview French current DSM-IV; 2015. <https://doi.org/10.13140/RG.2.1.2792.9440>
11. Shahid A, Wilkinson K, Marcu S, Shapiro CM. Beck Depression Inventory. New York: Springer; 2011.
12. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-13. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>

- PMid:11556941
13. Muñoz-Navarro R, Cano-Vindel A, Medrano LA, Schmitz F, Ruiz-Rodríguez P, Abellán-Maeso C, *et al.* Utility of the PHQ-9 to identify major depressive disorder in adult patients in Spanish primary care centres. *BMC Psychiatry*. 2017;17(1):291. <https://doi.org/10.1186/s12888-017-1450-8>
PMid:28793892
14. Sirait PJ. Validitas dan Reliabilitas Overt Agitation Severity Scale (OASS) versi Bahasa Indonesia. Universitas Sumatera Utara; 2013.
15. Osborne JW, Costello AB. Practical Assessment Research and Evaluation, *Psychology*. 2004;9:11.
16. Dahlan MS. Statistik untuk Kedokteran dan Kesehatan: Deskriptif, Bivariat dan Multivariat; 2015.
17. American Academy of Psychiatrics. *Mental Health Screening and Assessment Tools for Primary Care*. Vol. 125. American Academy of Psychiatrics; 2010.
18. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth*. 2017;11(Suppl 1):S80-9. https://doi.org/10.4103/sja.SJA_203_17
19. Sinclair S, Blais M, Blacker D. Psychiatric rating scale. In: *Comprehensive Textbook of Psychiatry*. 10th ed. Philadelphia: Wolters Kluwer Medknow Publications; 2017. p. 2722-48.
16. Dahlan MS. Statistik untuk Kedokteran dan Kesehatan: