

# Validation of the Thai Version of the Patient Health Questionnaire for Adolescents (PHQ-A) in adolescent psychiatric patients

Wimonwan Panyawong<sup>1</sup>, Chosita Pavasuthipaisit<sup>2</sup>, Rattanasak Santitadakul<sup>3</sup>

<sup>1,2,3</sup> Child and Adolescent Mental Health Rajanagarindra Institute, the Department of Mental Health, Thailand.

\*Corresponding author: Wimonwan Panyawong

Email: w.panyawong@gmail.com; Phone number: +668 1697 5625

Received: 21 October 2019

First Revision: 06 February 2020

Accepted: 26 March 2020

## Abstract

A basic screening tool for depression for adolescents was insufficient in Thailand. Thus, Patient Health Questionnaire for Adolescents (PHQ-A) was translated into Thai. This study aimed to examine the psychometric properties of the Thai version of the PHQ-A. Samples were 11-20 year old patients who attended 12 child and adolescent psychiatric clinics across Thailand. Of 272 adolescent patients, 172 patients had depression and 100 patients did not have depression. To evaluate criterion validity, the PHQ-A scores were compared with the clinical diagnosis. Internal consistency and exploratory factor analysis (EFA) were also conducted for item and factor analysis Convergent validity was calculated from correlations between the PHQ-A and the Children's Depression Inventory (CDI) and the Centre for Epidemiologic Studies-Depression Scale (CES-D). The internal consistency coefficient of the Thai version of the PHQ-A was 0.92. EFA suggested that one-factor structure of depressive symptoms was suitable for the PHQ-A. The area under the curve of the PHQ-A total score against professional diagnosis was 0.88. The results indicated an optimal cut-off point with acceptable sensitivity and specificity of 8. Moreover, the PHQ-A score of 10 was useful for differentiating between patients without depression and those with moderate to severe major depressive disorder (MDD). Convergent validity was good with high correlations for both the CDI and the CES-D ( $r=0.83$  and  $0.87$ , respectively). The Thai version of the PHQ-A is a valid and reliable measurement. The tool is simple and easy to use for screening and monitoring the severity of depressive symptoms in children and adolescents in both clinical and community settings.

**Keywords:** Adolescents, Depression, PHQ-A, Screening

## Introduction

Depression is one of the major mental health problems in an adolescent. Numbers of young people who have depressive disorder tend to increase gradually (WHO, 2014; 2017b). Globally, the prevalence rate of depressive disorders, including major depressive disorder (MDD) and dysthymia, increased to 18.7% from

2005 to 2015 (Baranne & Falissard, 2018). Moreover, it is also the major contributor to suicidal behavior, self-harm, and self-medication which has led to serious health and mental health problems (Vos et al., 2016; WHO, 2014; 2017a). In Thailand, the prevalence rate of depressive disorder in adolescents was found to be 2-4% (Ekasawin, Phothisut, & Chomcheun,

2016; WHO, 2014). 15% of Thai young people reported having suicidal ideas and 13% of those had attempted suicide at least once (WHO, 2017b). However, there were only small numbers of young people who had access to mental health services (Jirapramukpitak, 2017). It was estimated that the number would be much lower in young people, especially at-risk adolescents, such as substance abusing, LGBTQI and rural teenagers (Brown, Rice, Rickwood, & Parker, 2015; Jirapramukpitak, 2017). According to various national guidelines for adolescents with depression, early screening for depression with standardized tools is an important starting point (McDermott, et al., 2010; Birmaher & Brent, 2007; National Collaborating Centre for Mental Health, 2005). It was recommended that screening for depression should be provided for all teenagers (Birmaher & Brent, 2007; National Collaborating Centre for Mental Health, 2005). In Thailand, the Strength and Difficulties Questionnaire (SDQ) has been widely used as a basic screening tool for mental health in schools (Wongpiromsarn, Wipulakorn, Nuanmanee, Woener, & Mongkol, 2011). However, the test contains only five items of emotional problems which are insufficient for detecting depressive disorders. Some screening tools, specifically designed for child and adolescent depression and available in Thailand, were the Children's Depression Inventory (CDI) for children (7–17 years old) (Trangkasombat & Liknapichitkul, 1996), and the Centre for Epidemiologic Studies–Depression Scale (CES–D) for teenagers (16–18 year-olds) (Trangkasombat, Larpoonsarb, & Havanont, 1997).

Both tools have been used by many child and adolescent mental health clinics but were rarely used in primary care settings due to the number of items and the complicated scoring method (Charemboon, 2011). A short and standardized screening tool for depression, available in Thailand, is the Thai version of the 9-item Patient Health Questionnaire (PHQ-9) (Kroenke, Spitzer, & Williams, 2001; Lotrakul, Sumrithe, & Saipanish, 2008). Nevertheless, this tool has only been studied using adult

samples (Lotrakul et al., 2008). In 2002, Johnson, Harris, Spitzer, & Williams modified the PHQ-9 to be suitable for use with teenagers and called it the Patient Health Questionnaire for Adolescents (PHQ-A), which had been recommended by national guidelines for adolescents with depression. The purpose of both the PHQ-9 and the PHQ-A was to screen for depression, assess depression severity, and monitor response to treatment (Cappelli et al., 1995; US Preventive Services Task Force, 2009; Siu, 2016; Williams, O'Connor, Eder, & Whitlock, 2009). Notably, the authors of these tests generously gave permission for free use and allowed them to be translated into various languages (Kroenke et al., 2001; Johnson, Harris, Spitzer, & Williams, 2002) such as Chinese (Zhang et al., 2013) and Hindi (Ganguly et al., 2013). Because of the above-mentioned reasons, the short and simple questionnaire which was validated for screening and monitoring depressive disorders in Thai adolescents, was deficient for the Thai mental health system. In 2017, the Child and Adolescent Mental Health Rajanagarindra Institute, the Department of Mental Health of Thailand, developed the Thai version of the PHQ-A to use with all Thai children and adolescents.

#### Development of the Thai version of PHQ-A

As mentioned earlier, the PHQ-A test was developed from the PHQ-9 (Johnson et al., 2002). Some minor changes were made to the original test to suit the unique symptoms of depressive disorder in adolescents. Firstly, the item of depressed mood was placed as the first item, instead of the item anhedonia. Next, the word “irritable” was added to the item of depressed mood. “Weight loss” was added to the item of appetite. The PHQ-A consists of 9 items in relation to diagnostic criteria of depressive disorder. It is a self-report questionnaire asking about each symptom over the past two weeks. There are four points on the rating scales: 0–not at all, 1–several days, 2–more than half the days, and 3–nearly every day. The interpretation of the PHQ-A total score follows suit of the original

PHQ-9 to indicate severities of depression: 0–4 for minimal, 5–9 for mild, 10–14 for moderate, 15–19 for moderately severe, and 20–27 for severe. In 2017, the PHQ-A was translated into Thai following the guidelines for cross-cultural adaptation of self-report measures (Beaton, Bombardier, Guillemin, & Ferraz, 2000). To begin with, two forward translations were done into Thai from the original PHQ-A by two independent researchers. Then, the back-translations were done by two bilinguals who were fluent in both English and Thai. Next, the first pre-final draft was evaluated by three experts of adolescent depression by reviewing and validating contents and wordings. Finally, fifteen adolescent patients attending the Child and Adolescent Mental Health Rajanagarindra Institute were invited to share their opinions on the test. As a result, in item 2 “little interest or pleasure in doing things”, the word “bored” was used to convey a more accurate meaning in Thai. The test-retest reliability method was used to evaluate the stability and reliability of the final version of the Thai PHQ-A within a two weeks’ time. The method was carried out in a school setting in Bangkok. Correlation between separate administrations of the test was 0.75, which indicated moderately good test-retest reliability.

## Objectives

This study aimed to examine the validation of a Thai version of the PHQ-A test in 11–20 year old patients from child and adolescent psychiatric services throughout Thailand.

## Methods

### **Samples**

Samples were 11–20 year old patients who attended child and adolescent psychiatric clinics from May to August 2018. Depression patients were selected from their recent diagnoses, including major depressive disorder (MDD) (F32.0–32.3), dysthymia (F34.1), recurrent depression (F33.0–33.2), and bipolar-current

episode depression (F31.3). Non-depression patients were other patients with no history of depressive disorder. Patients with any anxiety-related disorders (F40–43), which share some overlapping symptoms with depression, were excluded. Patients with intellectual disability, severe psychosis and substance use disorder were also excluded due to possible cognitive disturbance.

Participants were recruited from twelve child and adolescent psychiatric clinics throughout Thailand with four of the clinics being child and adolescent mental health institutes: Child and Adolescent Mental Health Rajanagarindra Institute in Bangkok, Northeastern and Southern branches, and the Rajanukul Institute. The remaining child and adolescent psychiatric clinics were from four psychiatric hospitals: Suan Prung Psychiatric Hospital, Srithanya Hospital, Somdet Chaopraya Institute of Psychiatry, and Prasrimahabodi Psychiatric Hospital and four tertiary care hospitals, namely, Prapokklao Hospital, Pranangkla Hospital, Nakhon Pathom Hospital, and Uttaradit Hospital. The total number of participants was 272 patients of whom 172 patients presented with depressive disorders and 100 patients presented without depression. The mean age of all participants was 15.1 years old. Among the depression group, 70% were female and 78% were patients with MDD. In the non-depression group, 71% were male and 88% had attention-deficit hyperactivity disorder (ADHD).

### **Validation measures**

#### ***Diagnostic criteria for depressive disorders***

To test criterion validity, the diagnoses from child and adolescent psychiatrists were used as a gold standard to compare with the PHQ-A score. There were twenty-three board-certified child and adolescent psychiatrists who participated in this project. The clinical diagnostic interview carefully followed the International Statistical Classification of Diseases and Related Health Problems 10<sup>th</sup> Revision (ICD-10) (WHO, 2016), which is routinely used in Thai hospitals (Udomratn, 2009). The Clinical Global Impressions

scale (CGI) (Busner & Targum, 2007) was also used as a guideline for identifying the severity of the depressive symptom.

#### **CDI**

The Thai version of the Children's Depression Inventory (CDI) is a 27-item self-report questionnaire which assesses depressive symptoms in children. Each item contains three choices which are scored from 0 to 2 points. Total scores of the CDI ranged from 0 to 54. The instrument had good internal consistency with an alpha coefficient of 0.83, a satisfactory validity of sensitivity at 78.7%, and specificity at 91.3% (Trangkasombat & Liknapichitkul, 1996). This study used the CDI to compare with the Thai PHQ-A in 11–15 year old participants.

#### **CES-D**

The Thai version of the Centre for Epidemiologic Studies–Depression Scale (CES-D) is a 20-item self-reporting rating scale. It is suitable for teenagers aged 16 and over. Scores for each item range from 0 (not at all) to 3 (all the time). Total scores ranged from 0 to 60. The instrument also contains 4 items of positive thoughts and moods along with 16 other items of depressive symptoms. Internal consistency of the test was good with an alpha coefficient of 0.86. The Thai version of the CES-D had satisfactory validity with a sensitivity of 72% and a specificity of 85% (Trangkasombat et al., 1997). The CES-D was used in comparison with the PHQ-A in the 16–20 year old age group of participants in this study.

#### **Procedures**

This study was approved by the institutional review board of the Department of Mental Health, Thailand (reference number DMH.IRB.02/2561). Field researchers were mental health professionals who were permanent staff at each clinic, including registered psychiatric nurses, certified clinical psychologists, and public health officers. Adolescent patients were invited to join this study during their regular psychiatric appointments. The 11–15 year old age group of participants completed the PHQ-A and the CDI, whereas the 16–20 year old age group of

participants completed the PHQ-A and the CES-D. Finally, participants were assessed by their regular psychiatrists, using the ICD-10 and the CGI, to identify diagnoses and severity of depression blindly from the PHQ-A.

#### **Data analysis**

The internal consistency of the Thai version of the PHQ-A was measured by Cronbach's alpha coefficient. Exploratory factor analysis was also used to identify and validate the factor structure of the PHQ-A for this clinical group. The criterion validity of the tool was examined by sensitivity, specificity, and likelihood ratio. The Receiver Operating Characteristic (ROC) curve was created and the Area Under Curve (AUC) was calculated to determine the best cut-off score of the PHQ-A against the presence of depressive disorders. This study also assessed the validity of the PHQ-A as a diagnosis tool by using the category algorithm of depressive disorder: if five or more of the nine items were rated as "more than half the days" (2 points) and at least one of those symptoms was either depressed mood (item 1) or anhedonia (item 2). Item 9 (suicidal idea) was counted if presented at all (Lotrakul et al., 2008). In order to classify the severity of depression, the total scores of the PHQ-A were also divided into the following levels: 0–4 (minimal), 5–9 (mild), 10–14 (moderate), 15–19 (moderately severe), and 20–27 (severe). To determine convergent validity, Pearson's correlation coefficient was used to assess the relationship between the PHQ-A and the CDI for 11–15 year old samples or the CES-S for 16–20 year old samples.

## **Results**

### **Reliability and item analysis of the Thai version of the PHQ-A**

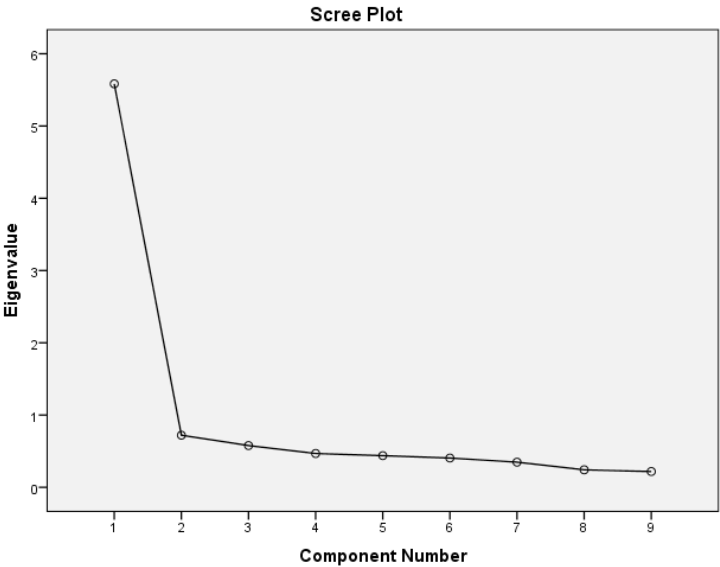
For all 272 participants, the internal reliability of the PHQ-A was excellent, with a Cronbach's alpha coefficient of 0.92. There was no item which would increase the total scale alpha if it was deleted. Item 7 (concentration problem) had the least item-total correlation (0.56) (Table 1).

**Table 1:** Item-total correlation and factor loading for PHQ-A. (N = 272)

PHQ-A item	Item-Total Correlation	Alpha if Item Deleted	Factor Loading
1. Feeling down, depressed, irritable, or hopeless?	0.83	0.91	0.88
2. Little interest or pleasure in doing things?	0.71	0.91	0.78
3. Trouble falling asleep, staying asleep, or sleeping too much?	0.68	0.92	0.75
4. Poor appetite, weight loss, or overeating?	0.72	0.91	0.78
5. Feeling tired, or having little energy?	0.76	0.91	0.82
6. Feeling bad about yourself – or feeling that you are a failure, or that you have let yourself or your family down?	0.78	0.91	0.84
7. Trouble concentrating on things like school work, reading, or watching TV?	0.56	0.92	0.63
8. Moving or speaking so slowly that other people could have noticed? Or the opposite – being so fidgety or restless that you were moving around a lot more than usual?	0.69	0.92	0.78
9. Thoughts that you would be better off dead, or of hurting yourself in some way?	0.77	0.91	0.83

Before running factor analysis, sampling adequacy was excellent (0.93) as per the Kaiser-Meyer-Olkin (KMO) criteria and Bartlett’s test of sphericity was statistically significant ( $p<.000$ ). The correlation matrix illustrated moderate-to-high correlation between individual item of the PHQ-A, ranged from 0.34 (item 7 and 9) to 0.74 (item 1 and 9). According to EFA

with Principal Component Analysis (PCA), only one factor was extracted with an eigenvalue  $>1$  which explained 62.04% of the variance with an eigenvalue of 5.58 as the scree plot showed in Figure 1. The one-factor structure was confirmed by a criterion of minimum factor loading of 0.63. (Table 1).



**Figure 1:** Scree plot showing the number of factors to retain

**Distribution of the PHQ-A scores according to the depression diagnosis criteria**

The mean of the PHQ-A score of the 100 patients without depression was 4.9 (SD 3.3) while that of the 172 patients with depression was 13.6 (SD 6.9). Within the depression group, the means of 141 patients with MDD and 31 patients with other depression were almost similar, with MDD at 13.5 (SD 7.0) and other depression at 13.9 (SD 6.7). Amongst MDD patients, the psychiatrists could also identify the severities of depressive disorder at 3 levels:

mild, moderate and severe MDD. The mean PHQ-A score was 10.5 (SD 5.8) in 63 patients with mild MDD, 15.7 (SD 6.5) in 46 patients with moderate MDD and 16.4 (SD 7.9) in 32 patients with severe MDD. Table 2 shows the distribution of the PHQ-A scores according to the depression diagnosis status. The PHQ-A scores of the majority of patients with no depressive disorder (90%) were less than 10, while those of most of the patients with moderate to severe depressive disorder (74% and 72%, respectively) were 10 and higher.

**Table 2:** Distribution of PHQ-A scores according to depression diagnosis status

Level of depression severity, PHQ-A score <sup>a</sup>	Major Depressive Disorder (N=141)			Other Depressive Disorder (N=31)	No Depressive Disorder (N=100)
	Mild (N=63)	Moderate (N=46)	Severe (N=32)		
	n (%)	n (%)	n (%)	n (%)	n (%)
Minimal, 0–4	7 (11.1)	1 (2.2)	2 (6.3)	1 (3.2)	47 (47)
Mild, 5–9	27 (42.9)	11 (23.9)	7 (21.9)	11 (35.5)	43 (43)
Moderate, 10–14	15 (23.8)	10 (21.7)	3 (9.4)	4 (12.9)	9 (9)
Moderately severe, 15–19	9 (14.3)	6 (13)	7 (21.9)	6 (19.4)	1 (1)
Severe, 20–27	5 (7.9)	18 (39.1)	13 (40.6)	9 (29)	0 (0)

**The validity of the Thai version of the PHQ-A**

Table 3 demonstrates the operating characteristic of the PHQ-A against the clinical diagnosis by psychiatrists which included sensitivity, specificity, and positive likelihood ratio (LR+). This study also conducted an ROC analysis, where the AUC was at 0.88 (SD=0.02, 95%CI 0.83 to 0.92). At a cut-off score of 8, the sensitivity was 76% and the specificity was 81%. The Likelihood ratio implied that a patient with depression was 4 times more likely to have a PHQ-A score of 8 or higher than a patient without depression. The validity of the PHQ-A as a diagnostic tool for depressive disorder, using category algorithm, had a sensitivity of 48% and a specificity of 100%.

The multilevel likelihood ratio was also calculated to determine further association between PHQ-A severity levels of depression and the likelihood of depressive disorder by

clinical diagnosis. The positive likelihood ratios of PHQ-A scores of 0–4, 5–9, 10–14, 15–19, and 20–27 for any depressive disorders were 0.1, 0.8, 2.1, 16.3, and 26.2, respectively. It meant that, for example, a PHQ-A score in the 0–4 range is only 1 in 10 times as likely in a patient with depression compared to a patient without depression, while a score of 10–14 is 2.1 times as likely and a score of 15–19 is 16.3 times as likely.



**Table 3:** Operating characteristics of PHQ-A cut-off scores

PHQ-A score	Sensitivity (%)	Specificity (%)	Likelihood Ratio (LR+)
≥ 5	93	47	1.8
≥ 6	90	65	2.6
≥ 7	84	76	3.5
≥ 8	76	81	4
≥ 9	68	87	5.2
≥ 10	61	90	6.1
≥ 11	58	92	7.2
≥ 12	55	95	11
≥ 13	55	97	18.2
≥ 14	48	98	24.2
≥ 15	42	99	42

To determine the convergent validity of the Thai version of the PHQ-A, the total score of the PHQ-A was correlated with the CDI for 11-15 year old participants and with the CES-D for 16-20 year old participants. Pearson’s correlation coefficient between the PHQ-A and

the CDI was 0.83, and was 0.87 between the PHQ-A and the CES-D. It was indicated that the PHQ-A had a high positive correlation with both existing Thai screening tools for depression in young people.

Discussion

The Thai version of the PHQ-A had been developed for assessing depression in adolescents who were 11-20 years old. This study aimed to evaluate its internal reliability and validity including the criterion and the convergent validity. The Thai version of the PHQ-A had excellent internal consistency with an alpha coefficient of 0.92. According to item analysis, the mood item (item 1) had the highest correlation with the total score illustrated that a word “irritable” was worth adding in the adolescent version (Richardson et al., 2010). On the contrary, the item of concentration (item 7) showed the lowest correlation with the total score. It could be explained by the nature of participants with no depressive disorder since the concentration problem is a common sign for both depression and ADHD (WHO., 2016). As a result, item 7 might not be a strong indicator to classify depressive disorder among adolescent

patients in a psychiatric clinic. The EFA confirmed one-factor structure of the PHQ-A as all items were clearly clustered into a single domain regarding diagnostic criteria of MDD. It was corresponded with other studies of the factor structure of the PHQ-A and the PHQ-9 which supported the one-factor model (Boothroyd, Dagnan, & Muncer, 2019; Naveed, Waqas, Memon, Jabeen, & Sheikh, 2019). Nonetheless, there were some evidences suggested two-factor model of the PHQ-9 in adult psychiatric population including cognitive /affective and somatic factor (Beard, Hsu, Rifkin, Busch, & Björqvinnsson, 2016; Guo et al., 2017). It was suggested a possible difference of the factor model relating to different age group which would need to be explored further in Thai population.

Means of the PHQ-A total score were gradually higher following the severity of depressive disorders classified by child and adolescent psychiatrists. Table 2 showed that

a large number of patients with a diagnosis of moderate and severe MDD also presented in moderate (score 10–14) to severe (score  $\geq 20$ ) categories of the PHQ-A. In contrast, 90% of patients without depression were on the minimal and mild grade of depression categorized by the PHQ-A. The results indicated a strength point of the PHQ-A in identifying the severity of depression (Bhatta, Champion, Young, & Loika, 2018). However, the PHQ-A severity levels were varied amongst the patients with mild MDD and other depressive disorders. Additionally, this study was a cross-sectional study, thus, the PHQ-A score could be varied depending on teenagers' emotional status at the moment of testing.

For consistency with previous studies, the criterion validity of the PHQ-A in this study was satisfactory (Johnson et al., 2002; Lotrakul et al., 2008). The cut-off point of 8 demonstrated acceptable sensitivity and specificity (76% and 81%, respectively). It was lower than the original recommended cut-off score of 10. However, it was in range with the satisfactory cut-off score of 8 to 11 from a meta-analysis study for an optimal score of the PHQ-9 (Manea, Gilbody, & McMillan, 2011). This slight difference can be explained by the settings and diagnostic method of this study. Manea et al. (2011) reported that the sensitivity of the PHQ-9 studies in a hospital setting tended to be lower than that of the primary setting, while specificity was rather stable. Moreover, this study was conducted using child and adolescent psychiatry as a gold standard, while previous studies often used structured clinical interviews (Johnson et al., 2002; Kroenke et al., 2001; Lotrakul et al., 2008; Zhang et al. 2013). The results of this study also supported the advantages of the severity interval of the PHQ-A (Johnson et al., 2002) (Table 3). For example, the cut-off point of 5 was great for discriminating between patients with depression and those without depression with high sensitivity (93%). Besides, the cut-off point of 10 was useful for differentiating between non-depression patients and depression patients with high specificity (90%). Finally, the

cut-off point of 15 was able to rule out almost all patients without depression with 99% specificity. Moreover, using the PHQ-A as a diagnostic tool following category algorithm resulted in exclusion of all patients with no depressive disorder. However, this method had low sensitivity (43%) which was similar to the study of the Thai version of the PHQ-9 (Lotrakul et al., 2008). Findings corresponded with the original PHQ-A which found moderate correlation with professional diagnosis, though high specificity for ruling out non-depression cases (Johnson et al., 2002).

The convergent validity of the PHQ-A was excellent compared with the CDI and the CES-D ( $r = 0.83$  and  $0.87$ , respectively). It was demonstrated that the 9-item PHQ-A can be used for assessing depression in children and adolescents, as well as the existing measurements. As mentioned earlier, both the CDI and the CES-D contain more numbers of items and more complicated scoring methods, than the PHQ-A. At this point, the PHQ-A is simple and easy to use. It can be used for monitoring both the severity of depressive symptoms and response to treatment (Siu, 2016; Williams et al., 2009). As a consequence, the tool has been used widely in primary care and school settings (Bhatta et al., 2018; Lewandowski et al., 2016).

## Limitations and Recommendations

This study was just a starting point for the Thai version of the PHQ-A. There were some limitations. Firstly, the results showed a high false positive rate among non-depressive patients who mostly had ADHD in that they showed similar symptoms to depressive disorder. To serve the purposes of a screening tool, the PHQ-A should be studied further using the general population, primary care settings, and some specific groups, such as students who are at risk of mental health problems, or adolescent patients with chronic illness. Additionally,



the high false negative rate of this study was generally derived from patients with mild MDD and other depressive disorders. It was possible that this study did not control the duration of depressive episodes. Specifically, amongst adolescent patients with mild MDD, it varied from critical patients through to chronic and almost recovered patients. As a result, the PHQ-A scores of this group were diverse. Consequently, a further clinical interview is essential for young people who are at risk of depression, even though their PHQ-A scores were low (Johnson et al., 2002).

In conclusion, the Thai version of the PHQ-A is an acceptable, simple, yet reliable and validated measurement for detecting depressive disorder in adolescents. Mental health professionals might apply the PHQ-A result as a part of their diagnosis and treatment planning in the future. Moreover, the test can be used as the first self-screening tool for children and teenagers in school and community settings. The tool is suitable for assisting Thai mental health and health professionals to improve the quality of services for adolescents with depression.

## Acknowledgements

This study was funded by UNICEF Thailand. We want to express our gratitude to all of the psychiatrists and mental health professionals from the 12 child and adolescent psychiatric clinics who assisted in data collection. We also wish to thank all those adolescent participants and their parents who agreed to join in this project.

## References

- Baranne, M. L., & Falissard, B. (2018). Global burden of mental disorders among children aged 5–14 years. *Child and adolescent psychiatry and mental health*, 12(1), 19.
- Beard, C., Hsu, K., Rifkin, L., Busch, A., & Björgvinsson, T. (2016). Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders*, 193, 267–273.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191.
- Bhatta, S., Champion, J. D., Young, C., & Loika, E. (2018). Outcomes of depression screening among adolescents accessing school-based pediatric primary care clinic services. *Journal of Pediatric Nursing*, 38, 8–14.
- Birmaher, B., & Brent, D. (2007). Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(11), 1503–1526.
- Boothroyd, L., Dagnan, D., & Muncer, S. (2019). PHQ-9: One factor or two?. *Psychiatry Research*, 271, 532–534.
- Brown, A., Rice, S. M., Rickwood, D. J., & Parker, A. G. (2015). Systematic review of barriers and facilitators to accessing and engaging with mental health care among at-risk young people. *Asia-Pacific Psychiatry*, 8(1), 3–22.
- Busner, J., & Targum, S. D. (2007). The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry (Edgmont)*, 4(7), 29–37.
- Cappelli, M., Clulow, M. K., Goodman, J. T., Davidson, S. I., Feder, S. H., Baron, P., ... & McGrath, P. J. (1995). Identifying depressed and suicidal adolescents in a teen health clinic. *Journal of Adolescent Health*, 16(1), 64–70.
- Charernboon, T. (2011). Measures and screening tests for depression in Thailand: A user's guide. *Thammasat Medical Journal*, 11(4),

- 667–676. (In Thai)
- Ekasawin, S., Phothisut, C., & Chomcheun, R. (2016). The prevalence of psychiatric disorders in Thai students aged 13–17 years. *Journal of Mental Health of Thailand*, 24(3), 141–153.
- Ganguly, S., Samanta, M., Roy, P., Chatterjee, S., Kaplan, D. W., & Basu, B. (2013). Patient health questionnaire-9 as an effective tool for screening of depression among Indian adolescents. *Journal of Adolescent Health*, 52(5), 546–551.
- Guo, B., Kaylor-Hughes, C., Garland, A., Nixon, N., Sweeney, T., Simpson, S., ... Morriss, R. (2017). Factor structure and longitudinal measurement invariance of PHQ-9 for specialist mental health care patients with persistent major depressive disorder: Exploratory structural equation modelling. *Journal of Affective Disorders*, 219, 1–8.
- Jirapramukpitak, T. (2017). *Mental health gap and related economic burden: a longitudinal community survey*. Retrieved from: <http://kb.hsri.or.th/dspace/handle/11228/4699?locale-attribute=th>. (In Thai)
- Johnson, J. G., Harris, E. S., Spitzer, R. L., & Williams, J. B. (2002). The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients. *Journal of Adolescent Health*, 30(3), 196–204.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9), 606–613.
- Lewandowski, R. E., O'Connor, B., Bertagnolli, A., Beck, A., Tinoco, A., Gardner, W. P., ... & Brace, N. E. (2016). Screening for and diagnosis of depression among adolescents in a large health maintenance organization. *Psychiatric Services*, 67(6), 636–641.
- Lotrakul, M., Sumrithe, S., & Saipanish, R. (2008). Reliability and validity of the Thai version of the PHQ-9. *BMC psychiatry*, 8(1), 46.
- Manea, L., Gilbody, S., & McMillan, D. (2011). Optimal cut-off score for diagnosing depression with the patient health questionnaire (PHQ-9): A meta-analysis. *Canadian Medical Association Journal*, 184(3), E191–E196.
- McDermott, B., Baigent, M., Chanen, A., Fraser, L., Graetz, B., Hayman, N., ... & Smalley, T. (2010). *Clinical practice guidelines: Depression in adolescents and young adults*. Melbourne: beyondblue: the national depression initiative.
- National Collaborating Centre for Mental Health. (2005). Depression in children and young people: identification and management in primary, community and secondary care. In *Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews [Internet]*. Centre for Reviews and Dissemination (UK).
- Naveed, S., Waqas, A., Memon, A.R., Jabeen, M., & Sheikh, M.H. (2019). Cross-cultural validation of the Urdu translation of the Patient Health Questionnaire for Adolescents among children and adolescents at a Pakistani school. *Public Health*, 168, 59–66.
- Richardson, L. P., McCauley, E., Grossman, D. C., McCarty, C. A., Richards, J., Russo, J. E., ... Katon, W. (2010). Evaluation of the patient health questionnaire-9 item for detecting major depression among adolescents. *Pediatrics*, 126(6), 1117–1123.
- Siu, A. L. (2016). Screening for depression in children and adolescents: US Preventive Services Task Force recommendation statement. *Pediatrics*, 137(3), e20154467
- Trangkasombat, U. & Liknapichitkul, D. (1996). Depression in children: study by using the Children's Depression Inventory. *Journal of the Psychiatric Association of Thailand*, 41, 221–230. (In Thai)
- Trangkasombat, U., Larpboosarb, V., & Havanont, P. (1997). Using CES-D for screening depression in adolescents. *Journal of the Psychiatric Association of Thailand*, 42. (In Thai)
- Udomratn, P. (2009). How Thailand has modified the section on mental disorders in the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10). *Asia-Pacific Psychiatry*, 1(2), 88–89.

- US Preventive Services Task Force. (2009). Screening and treatment for major depressive disorder in children and adolescents: US Preventive Services Task Force Recommendation Statement. *Pediatrics*, 123(4), 1223-1228.
- Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., ... & Coggeshall, M. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, 388(10053), 1545-1602.
- Williams, S. B., O'Connor, E. A., Eder, M., & Whitlock, E. P. (2009). Screening for child and adolescent depression in primary care settings: a systematic evidence review for the US Preventive Services Task Force. *Pediatrics*, 123(4), e716-e735.
- Wongpiromsarn, Y., Wipulakorn, P., Nuanmanee, S., Woener, W., & Mongkol, A. (2011). Strengths and Difficulties Questionnaire (SDQ) Thai improved version: change and administration. *Journal of Mental Health of Thailand*, 19(2), 128-134. (In Thai)
- World Health Organization (WHO) (2014). *Health for the world's adolescents: a second chance in the second decade*. Geneva: World Health Organization.
- World Health Organization (WHO) (2016). *International Statistical Classification of Diseases and Related Health Problems 10<sup>th</sup> Revision (ICD-10)*. Retrieved from <https://icd.who.int/browse10/2016/en>.
- World Health Organization (WHO) (2017a). *Depression and other common mental disorders: global health estimates*. Retrieved from <http://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf>.
- World Health Organization (WHO) (2017b). *Mental health status of adolescents in south-east Asia: evidence for action*. New Delhi: World Health Organization, Regional Office for South-East Asia. Retrieved from <http://apps.who.int/iris/handle/10665/254982>.
- Zhang, Y. L., Liang, W., Chen, Z. M., Zhang, H. M., Zhang, J. H., Weng, X. Q., ... & Zhang, Y. L. (2013). Validity and reliability of Patient Health Questionnaire-9 and Patient Health Questionnaire-2 to screen for depression among college students in China. *Asia-Pacific Psychiatry*, 5(4), 268-275.