The accuracy of Hopkins Symptom Checklist – 25 (HSCL-25) depression subscales (Indonesian version) on adolescents

Ayu Rahmawati Tirto¹, Sherly S. Turnip²

¹,²Faculty of Psychology, University of Indonesia, Indonesia
¹ayu.rahmawati51@ui.ac.id, ²sherly.saragih@gmail.com

**Introduction**

Adolescence is one of the lifetime periods when a person is susceptible to many mental health problems and disorders (Merikangas et al., 2010). One of the most widespread mental health problems is depression (Goyal, Srivastava, & Bansal, 2009; Jayanthi & Thirunavukarasu, 2015; Kementerian Kesehatan, 2013; The National Institute of Mental Health, 2017).

Studies showed that around 23% of adolescents had experienced symptoms of depression in their lifetime, with as many as .2% to 9.5% of community-based adolescents suffering from severe and extreme depression (Nair, Paul, & John, 2004; Sund, Larsson, & Wichstrom, 2011). Another study showed that the prevalence of both major and minor depression in adolescents is larger than 11% (Kessler & Walters, 1998; Merikangas et al., 2010).

Adolescents show a slightly different symptom of depression, particularly with regards to the prolonged sad mood. Sadness is often expressed in terms of unstable and irritable mood, long and persistent boredom, low energy, and awkwardness in a social
situation (Hammen & Watkins, 2008; Keena, 2005). A recent study found that suicidal thoughts, cognitive impairments, and having more negative thoughts about oneself also characterizes depressed adolescents (Orchard, Pass, Marshall, & Reynolds, 2017).

Depression in adolescents is mainly caused by hormonal imbalances from their developmental stage, impairment in social competences, and personal sense of failure in shifting from childhood to adulthood (Zgambo, Kalembo, Guoping, Honghong, & Honghong, 2012). Adolescents with low self-esteem and feelings of loneliness are also more prone to developing depression symptoms at a later stage in life compared to those with higher self-esteem (Brage & Meredith, 1994). Similarly, adolescents with negative relations to their social environment (e.g., peer, best friends, romantic partners) are also more susceptible to depression (La Greca & Harrison, 2005; Prinstein, Cheah, Borelli, Simon, & Aikins, 2005). Such symptoms often lead to poor social life, academic achievements, and relation with family members especially with siblings (Birmaher et al., 1996; Davis, 2000). It could also trigger other mental health issues in later life, such as substance abuse, bipolar disorder, and youth suicidal behavior (Birmaher et al., 1996; Ismail & Siste, 2010). Therefore, preventive measures for depression at an early stage becomes necessary to avoid the development of more severe symptoms and impairments.

One way to detect earlier symptoms of mental health disorders is by using a screening instrument as a guide for assessing early mental health status in professional health workers (Albert & Knoefel, 2011). The main advantages of using this tool are that it is easy-to-use and relatively time-efficient, avoiding over-consuming the professional health workers’ limited time (Taylor, 2002). One of the easy-to-use screening instrument forms is a self-report screening instrument. This is excellent tool to screen early mental health symptoms because it is cheap, relatively efficient in terms of time for test administration, and present more truthful results because answers are given by the participants themselves (Habere, Trabin, & Klinkman, 2013; McDonald, 2008).

HSCL-25 is one of the self-report instruments used to detect anxiety and depression symptoms. It is a brief version of HSCL, shortened from 58 items to 25 items. Initially, ten items were used to assess anxiety and 15 items for depression (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974). HSCL-25 is primarily used due to its simplicity, saving plenty of time and could be used by both professional and non-professional health workers (Ventevogel et al., 2007). Several studies have proven the use of this tool in assessing the psychological distress of adolescents in Oslo, Tibetan refugees, and Norway college students (Evans et al., 2008; Hjemdal, Friborg, & Stiles, 2012; Ystgaard, Tambs, & Dalgard, 1999).

This study uses the HSCL-25 Indonesian version adapted by Turnip and Hauff (2007) which has been applied in many adolescent studies in Indonesia. For example, this tool has been used in a study that assessed the relationship between psychological distress with social support, sense of community, and personality traits, as well as those that test the effectiveness of psychological therapies in decreasing psychological distress (Dewayani, Sukarlan, & Turnip, 2011; Megawanti, Sukarlan, & Turnip, 2011; Patria, 2012; Siregar, 2012; Wulanari, 2013). The previous examples measured psychological distress with the usual cut off score of 1.75.

Proper instruments which will be applied in specific population must show good accuracy through two main factors, namely sensitivity, and specificity. Accuracy is also shown through its positive and negative predictive values as well as likelihood ratios (LR) (Hulley, Cummings, Browner, Grady, & Newman, 2007; Spiers, Geller, & Kloss, 2013; Wiese, 2006). Meanwhile, accuracy testing is usually obtained by comparing the instrument with a gold standard as the criterion parameter (Bradley, 2013). In this study, we used a
diagnostic interview guide composing of SCID and few adolescents’ depression theories so it would be more accurate as the gold standard of adolescents’ depression.

The accuracy testing of HSCL-25 has been replicated in many studies across nations and cultures. These studies mainly showed good accuracy with moderate to high-level sensitivity and specificity in comparison to various gold standards (Fröjd, Håkansson, & Karlsson, 2004; Hinton et al., 1994; Lee, Kaaya, Mbwambo, Smith-Fawzi, & Leshabari, 2008; Ventevogel et al., 2007). The HSCL-25 Indonesian adapted version has undergone cultural validation (Turnip & Hauff, 2007) but the diagnostic testing, particularly accuracy testing of this version has never been done before.

This study aimed to test the accuracy (sensitivity and specificity) of HSCL-25 Indonesian adapted version for depression subscales. This study also aimed to test the appropriateness of the current cut-off score and later explore the optimum cut-off that could be used to differentiate between adolescents with depressive tendencies and those who are non-depressive. This study is part of a longitudinal study assessing the mental health status among adolescents in Indonesia. We hope that this study could increase the validity of the Indonesian version of HSCL-25 depression subscales, improving the preventive measures of depressive in adolescents.

Method

This research was a non-experimental study that compares two types of data, quantitative (HSCL-25 depression subscale score obtained from the preliminary study) and qualitative (depression diagnostic interview obtained from this study). These data collected to obtain the accuracy of HSCL-25 depression subscales, which are sensitivity, specificity, likelihood ratios, & predictive values.

This study was part of a longitudinal mental health study of Indonesian adolescents. Thus, the current participants were sampled from the preliminary research. The preliminary study was conducted on 623 participants from five randomly chosen samples of high schools that represent five Jakarta administrative regions (West, East, South, North, and Central Jakarta). Participants were freshmen students (grade X) aged between 14 to 18 years old. During the preliminary study, participants were asked to complete various mental health instruments, including HSCL-25 to assess their anxiety and depression.

Once the results were obtained, we used the HSCL-25 depression subscale scores to divide the participants into two categories based on the cut-off score (1.75). This HSCL-25 cut-off score is widely used across the cultures to determine whether someone is screened in the high or low-risk group in anxiety and depression (Dewayani et al., 2011; Hollander, Ekblad, Mukhamadiev, & Muminova, 2007; Sandanger et al., 1998; Veijola et al., 2003; Winokur, Winokur, Rickels, & Cox, 1984; Wulanari, 2013). The high depression group consists of participants who scored more than 1.75, while the low depression group comprised of participants whose depression subscale scored less than 1.75. These groups (excluding participants with comorbidity) were used as the population of the present study.

Next, we randomly selected 20 samples from each depression groups (of the preliminary study) to be the participants in this study. The researcher team (author was excluded) conducted the selection of the participants to avoid any expectation biases when collecting the data (double-blind study). There was a total amount of 40 random samples from three randomly chosen sample schools that participated in this study.

This study was a cross-sectional study conducted around November 2017, when the participants were in the middle grade of high school (Grade XI). We used HSCL-25
depression subscale as our primary instrument and a semi-structured interview guide as the gold standard. HSCL-25 is a brief self-report measure to detect the severity of anxiety and depression symptoms (Derogatis et al., 1974; Najarian & Davoodi, 2001). It is a short version of HSCL, containing 25 subscales that assess anxiety (10 items) and depression (15 items) (Derogatis et al., 1974). HSCL-25 measures the estimation of symptoms severity through a form of Likert scale, ranging from one (not at all) to four (extremely). The final score was obtained by dividing the total score with the number of items. The mean score used as the standard for high mental health disorder is 1.75, which was then used as the cut-off point for the preliminary study (Najarian & Davoodi, 2001). This entire instrument showed excellent internal consistency, while the depression subscale showed a Cronbach’s $\alpha$ of .85 (good) (Tay, Jayasuriya, Jayasuriya, & Silove, 2017). The Indonesian version of HSCL-25 was adapted by Turnip & Hauff (2007) and had been through cultural validity testing and resulted in good validity (Turnip & Hauff, 2007).

The gold standard we used in this study is a semi-structured diagnostic interview guide and clinical observation. This method was chosen because it is primarily used in health-related studies, specifically in assessing patients’ medical problems, becoming the basis for diagnosis and proper treatment plan (Rudolph, 2005). We constructed the interview guide based on many works of literature on adolescent depression (Hammen & Watkins, 2008; Keena, 2005) and the Structured Clinical Interview for DSM IV-TR (SCID) guide. We also used clinical observation to check the symptoms that they reported. SCID is a structured-interview guide for establishing the clinical diagnosis based on DSM IV-TR which is divided into several modules for assessing various mental health disorders. The depression symptoms belong to module A (First & Spitzer, 2002). SCID has good reliability in diagnosing Axis-I disorders in the DSM IV-TR (Lobbestael, Leurgans, & Arntz, 2011). The decision to add the adolescents’ depression theory on top of SCID, was done to increase the usefulness of the gold standard in capturing the depression phenomena among adolescents.

We measured the sensitivity and specificity of HSCL-25 through descriptive and crosstab analysis. Additionally, we also used the Receiver Operating Characteristic (ROC) curve analysis (conducted using IBM SPSS version 22.0) to examine whether the current cut-off score is appropriate for this population as well as explore the optimum cut-off point.

This study was obtained ethical clearance from the Ethical Committee of the Faculty of Psychology, University of Indonesia. Once permission was granted, we conducted individual interviews for all participants who have agreed to participate in the study in either the counseling or student council room. The interview was roughly 30-60 minutes per student, containing the introduction, explanation about the study, demographic and clinical data collection, and debriefing procedure. Several students rejected to participate after we asked the final permission. After we conducted another randomization, a total of 21 participants from high depression group and 19 from the low depression group were selected to participate in the interview process. The high and low depression groups were selected based on the HSCL-25 score in the preliminary study. Table 1 presents the demographic information of the study participants.
Table 1
The Distribution of Study Participants (n = 40)

<table>
<thead>
<tr>
<th>Gender</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>19</td>
<td>47.5%</td>
</tr>
<tr>
<td>Female</td>
<td>21</td>
<td>52.2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 years old</td>
<td>2</td>
<td>5%</td>
</tr>
<tr>
<td>16 years old</td>
<td>34</td>
<td>85%</td>
</tr>
<tr>
<td>17 years old</td>
<td>4</td>
<td>10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution of students per school</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Jakarta (“A” School)</td>
<td>9</td>
<td>22.5%</td>
</tr>
<tr>
<td>Central Jakarta (“B” School)</td>
<td>4</td>
<td>10%</td>
</tr>
<tr>
<td>East Jakarta (“D” School)</td>
<td>27</td>
<td>67.5%</td>
</tr>
</tbody>
</table>

Results

In the preliminary study, 520 out of 623 samples (83.5%) endorsed the high depression characteristics, based on the depression subscale of HSCL-25 (mean = 2.69) (see Table 2). This prevalence is very high compared to other mental health issues obtained in the preliminary study. Meanwhile, the diagnostic interview results in 14 out of 40 samples (35%) being endorsed into the high depression group (mean = 2.60, SD = .50), with the remaining rest 26 (65%) classified as either in no or low depression group (mean = 1.80, SD = .80).

Table 2
The Prevalence of Depression in the Preliminary Study

<table>
<thead>
<tr>
<th>Sample Participants (n = 623)</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>520</td>
<td>83.5%</td>
</tr>
<tr>
<td>Non-Depression</td>
<td>103</td>
<td>16.5%</td>
</tr>
</tbody>
</table>

Table 3 depicts the score comparison between HSCL-25 depression subscale and a diagnostic interview. The crosstab analysis shows that out of 14 students classified as having high depression symptoms by the diagnostic interview, 12 of them (85.7%) were also categorized as high depression by the HSCL-25 depression subscales (true positive group) (Spiers et al., 2013). The remaining two participants were categorized as low depression group by the HSCL-25 depression subscales, namely the false negative group (Carlson, 2012). Next, our analysis also resulted in 17 participants being classified into the true negative group (65.4%)(Spiers et al., 2013). Meanwhile, the remaining nine participants were categorized as high depression by HSCL-25 depression subscales. This group is considered as the false positive group (Carlson, 2012). The large false negative group could be caused by misinterpretation of the subscales as either participants’ current emotion or everyday life experiences, or just that they did not report their full consent in the interview. Based on these findings, then the sensitivity of HSCL-25 depression subscales was 85.7%, while the specificity was 65.4% (Spiers et al., 2013).
Table 3
Cross Tabulation Analysis of HSCL-25 Depression Subscales Compared to Gold Standard

<table>
<thead>
<tr>
<th>The results of diagnostic interview (gold standard)</th>
<th>Depression</th>
<th>Non-depression</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>12</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Percentage</td>
<td>85.7%</td>
<td>14.3%</td>
<td>100.0%</td>
</tr>
<tr>
<td>N</td>
<td>9</td>
<td>17</td>
<td>26</td>
</tr>
<tr>
<td>Percentage</td>
<td>34.6%</td>
<td>65.4%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Count</td>
<td>21</td>
<td>19</td>
<td>40</td>
</tr>
<tr>
<td>Percentage</td>
<td>52.5%</td>
<td>47.5%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Next, the result shows that 12 out of 21 participants were categorized as high depression group by both HSCL-25 depression subscales and the diagnostic interview, resulting in a positive predictive value of .57 (57%) for HSCL-25 (Spiers et al., 2013). Meanwhile, 17 out of 19 participants were categorized as low depression group by both HSCL-25 depression subscales and the diagnostic interview, indicating that the negative predictive value of HSCL-25 depression subscales is .90 (90%) (Hulley et al., 2007). The positive likelihood ratio of HSCL-25 depression subscales is 2.50, and the negative likelihood ratio is .20 (Wiese, 2006). Table 4 presents the predictive value and likelihood ratios.

Table 4
Predictive Values and Likelihood Ratios Results

<table>
<thead>
<tr>
<th>Predictive Values</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>57%</td>
<td>90%</td>
</tr>
<tr>
<td></td>
<td>(.57)</td>
<td>(.90)</td>
</tr>
<tr>
<td>Likelihood Ratios</td>
<td>2.5</td>
<td>.20</td>
</tr>
</tbody>
</table>

The area under the curve (AUC) of ROC (Figure 1) shows that the current sensitivity and specificity scores .75 (the confidence interval is between .59 and .91). The curve also formed in the middle to left of the entire area and the area under the ROC curve (AUC). The AUC and optimum cut-off points were presented in Table 5.

![ROC curve of HSCL-25 depression subscales](image-url)

Figure 1. ROC curve of HSCL-25 depression subscales
Next, by the distribution of symptoms analysis, the depression symptoms showed in the depression subscales of HSCL-25 is similar to the diagnostic interview. Table 6 shows the item distribution.

Table 6

The Distribution of Depression Symptoms in High Depression Group

<table>
<thead>
<tr>
<th>HSCL-25 Depression Subscale (n = 21)</th>
<th>Score 3</th>
<th>Score 4</th>
<th>Diagnostic Interview</th>
<th>Yes answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. 13 = 7</td>
<td>I. 13 = 5</td>
<td></td>
<td>Depressed, down or sad</td>
<td>9</td>
</tr>
<tr>
<td>I. 18 = 5</td>
<td>I. 18 = 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. 14 = 5</td>
<td>I. 14 = 5</td>
<td></td>
<td>Lost interest or pleasure in things that usually enjoyed</td>
<td>7</td>
</tr>
<tr>
<td>I. 23 = 14</td>
<td>I. 23 = 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. 15</td>
<td>7</td>
<td>10</td>
<td>Eating / appetite disturbance</td>
<td>6</td>
</tr>
<tr>
<td>I. 11 = 8</td>
<td>I. 11 = 4</td>
<td></td>
<td>Fidget, restless or slower than usual</td>
<td>9</td>
</tr>
<tr>
<td>I. 24 = 5</td>
<td>I. 24 = 2</td>
<td></td>
<td>Tired or low energy</td>
<td>8</td>
</tr>
<tr>
<td>I. 12 = 10</td>
<td>I. 12 = 2</td>
<td></td>
<td>Feeling worthless</td>
<td>7</td>
</tr>
<tr>
<td>I. 25 = 7</td>
<td>I. 25 = 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. 17 = 8</td>
<td>I. 17 = 9</td>
<td></td>
<td>Trouble in thinking, concentrating or making everyday decisions</td>
<td>11</td>
</tr>
<tr>
<td>I. 21 = 9</td>
<td>I. 21 = 6</td>
<td></td>
<td>The thought of things so bad until thinking about death, with/without specific attempts to self-hurt</td>
<td>11</td>
</tr>
<tr>
<td>I. 19</td>
<td>7</td>
<td>3</td>
<td>Mood changes between sadness to irritability</td>
<td>10</td>
</tr>
</tbody>
</table>

* I. = item, Score 3 = “quite a bit”, Score 4 = “extremely”
Discussion

Our findings show that the HSCL-25 has a sensitivity percentage score of 85.7%, while the specificity is 65.4%. These findings show that the Indonesian adapted version of HSCL-25 is considered to be an excellent instrument to screen people with symptoms of depression (Zhu, Zeng, & Wang, 2010). This is indicated by the higher numerical value of the sensitivity and specificity. This instrument is also suitable for adolescents. It is an excellent tool for preventive measures due to its’ high sensitivity (Wassertheil-Smoller & Smoller, 2015). These findings answer the research problem on whether the depression subscales of HSCL-25 have better sensitivity and specificity for detecting depression in adolescents, compared to the gold standard.

The above results are also supported by the fact that the +LR was considered as fair because it is not bigger than one. Meanwhile, the –LR was considered as good because it is lesser than one (Wiese, 2006). Thus, it show that HSCL-25 depressions subscales are considered as a fair to a good instrument. Similarly, the AUC analysis and ROC curve also indicate that the HSCL-25 depression subscale is a good instrument (Kumar & Indrayan, 2011).

Based on the optimum cut-off points offered by the AUC analysis (1.65 – 2.10), the current cut-off score (1.75) is considered to be an optimal score for HSCL-25 depression subscales. It lies between 1.65 (sensitivity: .929, 1 – specificity: .385) to 2.10 (sensitivity: .857, 1 – specificity: .345). This finding supports other studies across the globe that discover similar results regarding the optimum cut-off score (Dewayani et al., 2011; Hollander et al., 2007; Megawanti et al., 2011; Sandanger et al., 1998; Veijola et al., 2003; Wulanari, 2013). This also answers the research problem that questioned the optimum cut-off score of HSCL-25’s depression subscales in detecting depression among adolescents.

However, we suggest the cut-off score to be increased to 2.50 if the instrument is intended to be used as a diagnostic tool for people with high depression characteristics. This is because HSCL-25 can be used easily by many non-health professionals (Ventevogel et al., 2007). Thus, it would be more practical for them to use a single for all mental health assessments. This cut-off score would be good for HSCL-25 depression subscales as a diagnostic instrument because it still has a good compromise of sensitivity and specificity (sensitivity: .714, 1 – specificity: .269).

Next, based on the interview, mood change was a factor that many participants experienced (10 out of 14). This finding supports the existing literature that state that two of the most typical symptoms of depression found in adolescents were fluctuating mood and irritability (Keena, 2005). For that reason, we suggest adding items on mood changes to measure symptoms of depression in adolescents. By adding the mood change item to the HSCL-25 depression subscales, we hope that this instrument will better describe adolescents’ depression.

Our overall findings theoretically contribute to expanding the knowledge repertoire about the validity of HSCL-25, particularly the depression subscales. It also contributes to becoming the first study to describe the accuracy of HSCL-25 depression subscales on adolescents in Indonesia. Practically, these results could aid in diagnosing adolescents with mental health condition more accurately, allowing for a proper treatment plan to be constructed at an early stage.

However, several limitations exist in this study. First, the study participants were typically very homogenous in terms of age (16-18 years), familial, economic, residential, school, and cultural backgrounds. In other words, the result is limited in the matter of generalization. We suggest future studies to involve participants from a more heterogenic
demographic background to improve the generalizability and representativeness of the result.

Moreover, there were also issues regarding the relatively large number of false negatives in this study that were caused by misinterpretation of the items. We recommend prospective studies to use the diagnostic interview with longer duration or rapid sessions with a more in-depth approach. This suggestion is made to ensure that the researcher and participants have built the necessary rapport for participants to talk more openly about their mental health conditions. Additionally, it will also aid researchers in understanding the stories of each participant more deeply, clarifying the process of understanding the symptoms of depression.

Conclusion

According to the findings of this study, HSCL-25 depression subscales (Indonesian version) has the sensitivity percentage score of 85.7% and specificity of 64.3%, implying that it is a fair to good screening tool for depression in adolescents. These results are also supported by the predictive values, likelihood ratios, and ROC curve analysis results. The current cut-off for Indonesian adolescents (1.75) is an optimum cut-off score. Thus, it can still be used continuously for preventive measures of adolescents’ depression. However, the cut-off score should be changed to 2.50 if used for diagnosis purposes. From the symptom distribution analysis, we also suggest adding the mood changes item to complete this instrument for use on adolescents.

References


students born in Tibet compared to Tibetan students born in exile. \textit{Social Psychiatry and Psychiatric Epidemiology}. https://doi.org/10.1007/s00127-008-0346-9


Habrer, J. E., Trabin, T., & Klinkman, M. (2013). Furthering the reliable and valid measurement of mental health screening, diagnoses, treatment and outcomes through health information technology. \textit{General Hospital Psychiatry}. https://doi.org/10.1016/j.genhosppsych.2013.03.009


Ismail, R. I., & Siste, K. (2010). \textit{Depression disorder in Elvira, Silvia D., Hadisukanto, & Gitayanti (Eds.)} (Psychiatry). Jakarta: Medical Faculty of University of Indonesia.


Siregar, D. A. I. (2012). *Social skills training for university of Indonesia undergraduate students with high psychological distress*. University of Indonesia.


