Research paper

Screening for depressive and anxiety disorders among adolescents in Indonesia: Formal validation of the centre for epidemiologic studies depression scale – revised and the Kessler psychological distress scale

Thach Duc Tran, Fransiska Kaligis, Tjhin Wiguna, Lisa Willenberg, Peter Azzopardi, Jane Fisher

1 Introduction

It is estimated that 10–20% of adolescents in the world are experiencing at least one mental health problem (Kieling et al., 2011; Patton et al., 2016). Mental disorders including depressive disorders and anxiety disorders account for up to 17% of the disability-adjusted life-years (DALYs) lost among 15–19 year-old adolescents (Mokdad et al., 2016). Despite the significance of adolescent mental health, it is yet to be prioritised in research and public health policy in low- and middle-income countries (LAMICs) which causes a huge gap between the needs and the resources available to meet them (Belfer, 2008; Erskine et al., 2017; Fisher et al., 2011).

Valid and reliable screening tests for adolescent mental health problems play a vital role in addressing these public health problems. Screening tests, which are affordable and easily completed, are necessary first for large-scale epidemiological studies that can provide knowledge about the magnitude of and risk and protective factors for the problem. Second, screening tests are needed for community and primary health care settings in order to increase recognition and early detection of the problems. Depressive and anxiety disorders are the most common mental health problems among adolescents. Without effective treatments, these mental health problems can lead to more severe conditions including psychoses (Fusar-Poli et al., 2012; Pelizza et al., 2018), self-harm and suicide (Hawton et al., 2012). Early
detection and treatment of these mental health problems is essential to improving prognosis and reducing disability (WHO, 2001). However, screening tools that have been locally translated, culturally verified and formally validated for use among adolescents are not available in most LAMICs.

The Centre for Epidemiologic Studies Depression Scale (CES-D) was created by Lenore Radloff an epidemiologist statistician at the National Institute of Mental Health (NIMH) in 1977 (Radloff, 1977), and revised (CESD-R) in 2004 by William Eaton and colleagues (Eaton et al., 2004). The CES-D is a self-report tool amalgamated from Zung's Depression Scale (Zung, 1965), the Beck Depression Inventory (Beck et al., 1996), and the Minnesota Multiphasic Personality Inventory (MMPI) (Hathaway and McKinley, 1951). The CES-D was revised to reflect the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) definition of Major Depressive Disorder. The CES-D/CESD-R are currently used worldwide to screen for depression among adults. The original CES-D has been validated for use among adolescents in high-income settings and found to be reliable and valid (Heo et al., 2018; Logsdon and Myers, 2010; Radloff, 1991; Stockings et al., 2015). The construct validity of the CESD-R short version (10 items) was established in large samples of adolescents in the USA (Harox et al., 2014). However, the updated CESD-R (full version) has not been validated for use among adolescents.

The Kessler Psychological Distress Scale – 10 items (K10) is one of the few screening instruments that is relatively short (10 questions and takes 2–3 min to complete). It is designed to screen for psychological distress that includes not only depressive disorders but also anxiety, another more common mental disorder. Kessler et al. (2002) developed the K10 using item response theory methods in two large-scale population-based surveys. The K10 and K6 have been demonstrated through national surveys to be useful screening instruments among general adult population samples in USA, Australia, New Zealand, and Canada (Furukawa et al., 2003; Kessler et al., 2003; Oakley Browne et al., 2010) and the World Health Organization World Mental Health (WMH) Survey Initiative in 28 countries (Furukawa et al., 2008; Kessler et al., 2010; Kessler and Ustün, 2004). The K10 and K6 have been used in studies of adolescent mental health (Green et al., 2010; Huang et al., 2009; Jaisoorya et al., 2017; Sampasa-Kanyinga and Hamilton, 2016). Several studies have shown that K10 and K6 have good psychometric properties for use among adolescents in diverse settings including China, India, and Australia (Chen and Feng, 2014; Huang et al., 2009; Jaisoorya et al., 2017; Mewton et al., 2016). Green et al. (2010) found that the K6 provides a fairly good prediction of serious emotional disturbance in American adolescents.

Indonesia is a lower-middle-income country in South East Asia. With a population of 261 million people, Indonesia is one of the world’s most populous countries and is home to 48 million adolescents (people aged 10–19 years) (UNICEF, 2018). As in other LAMICs, the burden of mental health problems is unclear in Indonesia. There is no locally validated tool available for use to screen for common mental disorders among adolescents in Indonesia. The aim of this study was to culturally verify and establish the empirical psychometric properties of the Indonesian versions of the CES-D-R and the K10/K6 and their overall performance in detecting depressive and anxiety disorders among older adolescents in Indonesia.

2. Methods

2.1. Study design

This study included two inter-linked components: (1) translation and cultural verification and (2) a formal validation study.

2.2. Setting

This study was conducted in Jakarta, the largest city and capital of Indonesia. Jakarta is the most populous city in Southeast Asia with a population of >10 million people. Most people are Muslim, with small Christian & Buddhist populations.

2.3. Participants and procedure

2.3.1. Component 1

The CESD-R and the K10 were translated and culturally verified using a four-step process. In the first step, two bilingual research and clinical psychiatrists independently forward translated the scales from English to Indonesian. All difficulties and differences in translations were discussed among the two translators and the Indonesian investigators to obtain consensus on the content of Version 1. In the second step, 6 health workers and psychiatric researchers formed a group to discuss the meaning and comprehensibility of each item and response option in Version 1. The Indonesian research team (FK and TW) reviewed all suggestions raised in the small group discussion and made the necessary adaptations of language to obtain Version 2. In the third step, Version 2 was pilot-tested among a small group of eight adolescents to evaluate comprehensibility and acceptability. An open-ended question was included after each item of the scales asking about if it is comprehensible and acceptable and seeking any suggestions for changes. The Indonesian team made any further adaptions based on participant feedback to generate Version 3 of the translation. In the last step, Version 3 was independently back-translated into English by two other bilingual health workers. A clinical psychologist and researcher (JF) who is an English native speaker checked the translations against the original version of the scales and identified any items that required specific verification. The Indonesian team made the final adaptation based on the results of the back translation to finalise the Indonesian language versions of the CESD-R and K10.

2.3.2. Component 2

The formal validation study was nested within a survey of non-communicable disease and associated risks among Indonesian adolescents (main survey).

In the main survey, a total of 627 adolescents (16–18 years old) attending senior high schools in Jakarta, Indonesia were recruited to participate using a multiple stage sampling method. First, a total of 11 schools were randomly selected from a list of 580 public, private and religious senior high schools across Jakarta. Second, in each of the selected schools, one class per grade 10–12 was randomly selected. Third, all students in the selected classes were invited to participate. They were recruited if they and their parents each provided written informed consent to participate in the study.

Senior high schools include Grades / Years 10–12. There are no major differences in the socio-economic-demographic characteristics and academic results of students in public, private and religious schools. Among 580 senior high schools in Jakarta, 116 are public, 370 private and 94 religious.

The main survey assessed social and demographic characteristics and included the translated, culturally verified versions of the CESD-R, and the K10. Students self-completed the survey in class time using individual handheld tablets with REDCap (Research Electronic Data Capture) software. All data collection for the main survey was conducted in classes under the supervision of research assistants with the agreement of schools and teachers.

The formal validation study is a sub-study of the main survey. A subsample of adolescents were selected from the students who had participated in the main survey and had provided informed consent to participate in sub-studies and whose parents had also consented. In each class which participated in the main survey, 6 students were randomly selected from a list of participating students who had completed the main survey and had parental consent to being selected for the interview. The selections were made by research assistants using the ‘RAND’ function in Microsoft Excel. They were interviewed using the
The K6 has 1 item selected for each of the psychological dimensions mental health problems, defined as severe depressive disorders and current suicidal thoughts or recent suicidal behaviours, were referred to the Child/Adolescent Psychiatric Unit in the main hospital in Jakarta. Parents/guardians were notified about the young person’s needs via phone and a letter containing the contact details of the two study investigators who facilitated the referral process. Participants diagnosed with minor mental health problems during the interview, including mild-moderate depressive disorders, dysthymia, or an anxiety disorder were given a letter for their parents/guardians containing recommendations to seek further counselling at the nearest psychiatric service within their area. All letters for parents/guardians were written in general terms and did not contain details of the participants’ mental health problems. All personal contact information provided by study participants (for this purpose) was immediately destroyed once the referral was complete. All data collected from participants for the study were identified only by a code number and not their name. Data were entered and stored in password-protected computer files on a secure server at Universitas Indonesia. The database was transferred to Monash University on a secure cloud storage facility and stored on a server at Universitas Indonesia. This diagnostic tool has been used widely in clinics in Indonesia.

2.4. Measures

2.4.1. The Centre for Epidemiologic Studies Depression Scale – revised (CESD-R) Indonesian version

The CESD-R includes 20 items measuring depressive symptoms in nine different groups including sadness, loss of interest, appetite, sleep, thinking/concentration, guilt, tired, movement, and suicidal ideation. The version that was used was the final version of the Component 1 translation and cultural verification process (Appendix 1). It has 20 items, each scored on a 5-point ordinal scale for frequency (0 = not at all or less than 1 day in the last week; 1 = 1 to 2 days in the last week; 2 = 3–4 days in the last week; 3 = 5–7 days in the last week; 4 = nearly every day for 2 weeks). The range of possible total scores is from 0 to 80, with higher scores indicating more depressive symptoms. The CESD-R can be scored to have the same range as the original CES-D version. This method is called the ‘CES-D style score’ in which the values for the top two responses (‘5–7 days in the last week’ and ‘nearly every day for 2 weeks’) are given the same value (3) and the total scores range from 0 to 60. In this study, the results will be reported for scoring using both styles (hereby called CESD-R 80 and CESD-R 60).

In adults, having a total CESD-R 60 score less than 16 indicates no clinically significant depressive symptoms (Eaton et al., 2004).

2.4.2. The Kessler psychological distress scale – 10 items (K10)

The 10 items include feelings of depression (2 items), nervousness (2 items), restlessness (2 items), worthlessness (1 item), being fatigued (1 item), hopelessness (1 item), and finding everything an effort (1 item). A 6-item subset of K10 was developed by Kessler and colleagues; the K6 has 1 item selected for each of the psychological dimensions (Kessler et al., 2002). The version that was used is the final version of the Component 1 translation and cultural verification process (Appendix 2). Each item is scored on a 5-point ordinal scale for frequency (0 = ‘none of the time’; 1 = ‘a little of the time’; 2 = ‘some of the time’; 3 = ‘most of the time’; 4 = ‘all of the time’). The total K10 scores range from 0 to 40. The K6’s items are scored in the same method as K10’s items. The total K6 scores range from 0 to 24.

In adults, K10 total scores are categorised as 0–5 ‘Low distress’; 6–11 ‘Moderate distress’; 12–19 ‘high distress’; and 20–40 ‘very high distress’ (Andrews and Slade, 2001). For K6 total scores, adults with scores of 13–24 are classified as having a probable serious mental health problem and those with scores of 0–12 as probably not having one (Kessler et al., 2010).

2.4.3. The Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID)

The MINI Kid is a structured diagnostic interview for the major DSM-IV and ICD-10 psychiatric disorders experienced by children and adolescents (Sheehan et al., 2010). In this study, the modules for Major Depressive Episode, Dysthymia, Panic Disorder, Separation Anxiety Disorder, and Generalized Anxiety Disorder of the Indonesian version of the standard MINI Kid were administered. Depressive disorder was diagnosed if a young person had symptoms meeting diagnostic criteria for a major depressive episode and/or dysthymia. Anxiety disorder was diagnosed if symptoms met diagnostic criteria for current panic disorder, separation anxiety disorder, and/or generalized anxiety disorder. The MINI Kid has been translated by Dr. Ika Widyawati et al., Division of Child and Adolescent Psychiatry, Department of Psychiatry, Universitas Indonesia. This diagnostic tool has been used widely in clinics in Indonesia.

2.5. Data management and statistical analysis

The MINI-Kid data were recorded on paper-based questionnaires and then double-entered using REDCap. The MINI-Kid data were linked to the main survey data using participants’ unique identification numbers.

The internal consistency of the scales was evaluated using Cronbach’s alpha coefficients with a cut-off of 0.8 used to indicate high internal reliability.

To identify the overall performance of the scales to detect a mental health problem, Receiver Operating Characteristic (ROC) Curves analyses were performed. A ROC curve is the plot of True Positive Fraction (sensitivity) and False Positive Fraction (1 – specificity) across varying cut-offs of the scale. The Area Under the ROC Curve (AUC) summarising the entire location of the ROC curve is a combined measure of sensitivity and specificity that describes the inherent validity of the scale to detect a problem. The closer the AUC is to 1, the higher the overall accuracy of the scale. AUCs between 0.5 and 0.7 are categorised as low accuracy; between 0.7 and 0.9 as moderate accuracy; and between 0.9 and 1.0 as high accuracy (Fischer et al., 2003).

The ROC curve provides detailed values for the sensitivity and specificity for each cut-off point. Based on this information, an optimal cut-off can be identified.

In this study, an optimal cut-off point to detect a mental health problem was identified using Youden’s index (Youden’s index = sensitivity + specificity – 1) (Youden, 1950). This index at a cut-off point is the height of the vertical line between the ROC curve at that point and the diagonal, chance line and it is also equivalent to the area under the curve subtended by that operating point (Hajian-Tilaki, 2018). The optimal cut-off point is the cut-off value that maximizes the Youden index. The sensitivity, specificity, positive likelihood ratio (LR+), and negative likelihood ratio (LR-) were reported for each optimal cut-off point.

The CESD-R was validated against MINI depressive disorder. The K10 and K6 were validated against MINI depressive disorder, anxiety disorder, and any depressive or anxiety disorder.

Sensitivity analyses were performed including weighting adjustment to check if the main results are robust. For any main demographic characteristic of the sample that was different to that of the corresponding population, post-stratification weights were used to adjust for the difference.
For each of K10 and CESD-R, only cases with missing data less than or equal to 30% number of the scale items were included in all analyses. Missing data were imputed using regression models that included as predictors all other items of the two scales and the social-demographic characteristics. The regression models were constructed under the Structural Equation Modelling framework with the full information maximum likelihood method (FIML) for missing data. All analyses were performed using STATA version 14.2.

2.6. Ethics approvals

This study was approved by the Alfred Ethics Committee (No. 141/17 on 14th December 2017) and the Ethics Committee of the Faculty of Medicine, Universitas Indonesia (No. 527/UN2.F1/ETIK/2017 on 5th June 2017).

3. Results

3.1. Cultural verification

In step 2 (small group discussion of 6 health workers and psychiatric researchers) and step 3 of Component 1 (a pilot test among 8 adolescents), all 20 items of CESD-R and 10 items of K10 were found to be culturally appropriate, comprehensible, and meaningful. There were some suggestions of language to be used for ease of understanding that have been incorporated in the final Indonesian versions.

3.2. Sample

Of the 11 schools randomly selected for this study, 6 are public (4 of which were standard education schools and 2 vocational schools), 3 private (2 of which were standard and 1 of which was vocational), and 2 religious (all were standard). Overall, 196 adolescents provided data for this study. The socio-demographic characteristics of the participants are presented in Table 1.

Among 196 participants, 41 had any missing CESD-R data and 10 missing K10 data. There were 7 participants with more than 30% of CESD-R data (>6 items) missing. No participant had more than 30% (3 items) of K10 data missing. Therefore, data from 189 participants were included for analyses related to CESD-R and all 196 participants were included for analyses related to K10/K6.

The distributions of the CESD-R 60, CESD-R 80, K10 and K6 total scores are presented in Table 2.

3.3. Internal reliability

The internal reliability Cronbach's alpha coefficient of the CESD-R 60 and CESD-R 80 was 0.90 (95%CI: 0.88–0.92); the K10 was 0.89 and the K6 was 0.83 (95%CI: 0.80–0.86). All of the scales have Cronbach's alpha >0.8.

### Table 1

| Statistics | 
|------------|------------|
| Age in years – Mean [SD] | 16.5 [6.07] |
| Grade level – n (%) |  
| 10 | 64 (32.6%) |
| 11 | 84 (42.9%) |
| 12 | 48 (24.5%) |
| Gender |  
| Boy | 86 (43.9%) |
| Girl | 110 (56.1%) |
| Religion |  
| Islam | 171 (87.2%) |
| Christianity | 16 (8.2%) |
| Others | 9 (4.6%) |

SD: standard deviation.

4. Discussion

This is the first study to describe the performance and identify the optimal cut-off points of the CESD-R, K10, and K6 to detect depressive and anxiety disorders among adolescents in Indonesia. The data reveal that these scales have good internal consistency and discriminant ability to detect any depressive disorder or any depressive /anxiety disorder among 16–18 year-old Indonesian students in Jakarta. They indicate that the CESD-R Indonesian version is a comprehensible and sensitive tool for screening for depressive disorder and both the K10 and K6 Indonesian versions are for screening for any depressive or anxiety disorder among 16–18 year-old Indonesian students living in major cities in Indonesia.

Similar to the original CES-D (Stockings et al., 2015), the revised version (CESD-R) has high Cronbach’s alpha and high sensitivity/specificity to detect any depressive disorder among adolescents in this study. The results of this study suggest a cut-off value of ≥20 (in CES-D style score) to detect any depressive disorder among adolescents in Indonesia that is higher than the common cut-off of ≥16 that is
suggested for use among adults. This is the first study validating the CESD-R against a clinician-administered diagnostic tool among adolescents so that there is no prior evidence to compare it with. Haroz et al. (2014) demonstrated that a short 10-item version of CESD-R had good a factorial construct among two large samples of 13–18 years old American. Stockings et al. (2015) review of validation studies showed that the optimal cut-off values identified for the original CES-D version were ≥12 (students, aged 12–14 years in the USA), 16 (students, aged 15–18 years in the USA), 22 (girl students, aged 12–14 years in the USA and community sample, aged 14–16 years in the Netherlands), 24 (students, aged 14–18 years in the USA), and 30 (community sample, aged 10–17 years in Rwanda). In Vilagut et al.’s review (2016), the optimal cut-offs of ≥16 (were reported in 22 studies), 20 (in 12 studies), and 22 (in 7 studies) among general adult populations in different settings including the USA, England, the Netherlands, Germany and Columbia. This heterogeneity strongly suggests that the cut-off value can be different across different populations. The cut-off score identified in this study is lower than the one among adults in America and might reflect the differences in the performance of the scale between adolescents and adults, or between Indonesians and Americans, or both of these possibilities.

In line with the results of previous studies (Chan and Fung, 2014; Huang et al., 2009; Jaisooya et al., 2017; Newton et al., 2016), the results of this study confirmed that K10/K6 have high comprehensibility and internal reliability for use among adolescents. However, there is to date no study validating the K10 against a gold standard diagnostic instrument for depressive and anxiety disorders. This study is the first in the world reporting the evidence of the optimal cut-off values for K10/K6 to detect depressive and anxiety disorders. Green et al. (2010) found that the K6 performed well to predict depressive and anxiety disorders with AUCs ranged from 0.75 to 0.83 in American students that are similar to what we found in this study in Indonesian students. Green et al. did not however report optimal cut-off scores. This study indicates that the K10’s optimal cut-off value is ≥18 to detect depressive disorders, anxiety disorders, or any depressive or anxiety disorder among Indonesian adolescents, which is lower than that for adults in America. Similarly, the corresponding cut-off of K6 is ≥12 is slightly lower than the cut-off of ≥13 for American adults.

Depression and anxiety disorders commonly co-occur (Goldberg and Huxley, 1992). Most risk factors for these mental health problems are shared. They should be addressed at the same time, especially in community or school settings, or in primary health care. The K10/K6 were designed to capture psychological distress, including both depression and anxiety. The results of this study suggest both K10 and K6 can be used to screen for depressive or anxiety disorders, each with high sensitivity and specificity. The K6 is preferable for use in school settings or primary health care because it is shorter and has similar good psychometric properties. The CESD-R was constructed to screen for depressive disorders only. It is longer than the K10/K6. We would not suggest CESD-R for use in primary health care but for programs or epidemiological studies that focus on depression among Indonesian adolescents.

The optimal cut-off points we suggested in this study were identified using the widely accepted method to maximise the total of sensitivity and specificity (or to minimize the sum of false positive and false negative errors). However, based on the detailed data on sensitivity and specificity that is provided in the Supplementary file, more sophisticated methods to identify the optimal cut-offs for specific purposes can be calculated. For instance, the information about the financial cost of false positive and false negative estimates and the costs of a further workup for diagnosis of the disease can be combined with ROC curve data using utility-based decision theory to determine optimal cut-offs (Hajian-Tilaki, 2018).

The strengths of this study are that the scales were validated against the MINI Kid, in psychiatrist-administered structured diagnostic interviews. It permits confident identification of the performance and the optimal cut-off values to detect the common mental health problems among older Indonesian adolescents. We acknowledge that the school-based sample in this study limits the generalisability of the findings to this group. It was beyond the resources of this study to identify and assess a randomly selected sample of out-of-school adolescents, and we acknowledge that the properties of these scales might not be identical among them. We used the version of MINI-Kid that has been translated and validated by Dr. Ika Widyawati at al., Division of Child and Adolescent Psychiatry, Department of Psychiatry, Universitas Indonesia. However, the data of that validation study have not been published. We acknowledge that as a limitation of our study. We also acknowledge that the characteristics of the sample of this study were not properly balanced by the representation in terms of grade and gender distributions. However, we have conducted the sensitivity analyses using post-stratification weights to adjust for the imbalances and the main results did not change.

### 5. Conclusions

This study suggests that the CESD-R is a useful tool for screening for depressive disorders and both the K10 and K6 are useful for screening for any depressive or anxiety disorder among Indonesian adolescents. Future studies can verify the findings of this study among younger adolescents and out-of-school adolescents in Indonesia.

### Contributors

TT and JF conceptualized and designed the study. FK, TW, and LW collected the data and trained and supervised the health researchers who conducted the interviews. TT and HN conducted data management and statistical analyses. TT wrote the first draft of the manuscript. All authors reviewed and edited the draft. JF and TT reviewed and finalised the manuscript.
Role of funding

This study is funded by The Australia-Indonesia Centre under the Health Cluster Tactical Research Funding. TDT and PA are supported by Early Career Fellowships from the Australian National Health and Medical Research Council. JF is supported by a Monash Professonal Fellowship and the Jean Hailes Professorial Fellowship, which is funded by a grant to the Jean Hailes Foundation from the H and I Hect Trust managed by Perpetual Trustees. The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Acknowledgements

We are grateful to the Australia-Indonesia Centre for funding and to the adolescents who provided the data. We thank Prof. DR. Dr. R. Irawati Ismail, Dr. Noorhana SW, Dr. AAAA Kusumawardhani, DR. Dr. Nurmiati Amir, for participating the cultural verification; Karin Nadia Utami, Muhammad Insan, and Jessica Marsigit, who participated in the data collection and supported this study.

Conflict of interest

The authors declare no conflicts of interest.

Supplementary materials


References