

Validitas dan Reliabilitas *Patient Health Questionnaire-4* (PHQ-4) Versi Bahasa Indonesia sebagai Kuesioner Deteksi Dini Gejala Cemas dan Depresi pada Anak dan Remaja

Tjhin Wiguna,¹ Kusuma Minayati,¹ Osi Kusuma Sari,² Yunita Arihandayani,² Imran Pambudi,² Wafa Sofia Fitri³

¹Departemen Psikiatri Rumah Sakit dr. Cipto Mangunkusumo-Fakultas Kedokteran Universitas Indonesia, Jakarta

²Direktorat Pelayanan Kesehatan Kelompok Rentan, Kementerian Kesehatan Republik Indonesia

³Program Studi Subspesialis Kedokteran Jiwa Fakultas Kedokteran Universitas Indonesia, Jakarta

Latar belakang. Gangguan kecemasan dan depresi pada anak dan remaja dilaporkan meningkat terutama dalam beberapa tahun terakhir ini, sementara instrumen deteksi dini yang singkat dan tervalidasi dalam Bahasa Indonesia masih terbatas. PHQ-4 merupakan suatu kuesioner deteksi dini singkat sehingga berpotensi untuk digunakan luas di masyarakat.

Tujuan. Mengevaluasi validitas dan reliabilitas PHQ-4 versi Bahasa Indonesia pada anak dan remaja usia 7–18 tahun.

Metode. Studi potong lintang ini melibatkan 1846 anak dan remaja di Jakarta, Tangerang Selatan, dan Serang; yang terdiri dari 440 anak usia 7–10 tahun dan 1406 remaja usia 11–18 tahun. Validitas konten dinilai dengan CVI, reliabilitas dengan Cronbach's alpha, *construct validity* dengan *Exploratory Factor Analysis* (EFA), dan validitas kriteria dengan ROC menggunakan *The Mini International Neuropsychiatric Interview for Children and Adolescents* (MINI-KID). Semua analisa dilakukan dengan menggunakan program SPSS versi 29.

Hasil. Seluruh butir PHQ-4 versi Bahasa Indonesia untuk anak usia 7 – 10 tahun menunjukkan I-CVI $\geq 0,83$ setelah dilakukan modifikasi menjadi kalimat tanya, dan I-CVI=1,00 untuk PHQ-4 bagi anak usia 11–18 tahun. Reliabilitas PHQ-4 untuk anak usia 7–10 tahun adalah Cronbach's $\alpha=0,644$, sedangkan untuk anak usia 11–18 tahun Cronbach's $\alpha=0,693$. Analisis EFA untuk PHQ-4 untuk anak usia 7–10 tahun maupun 11–18 tahun menunjukkan KMO 0,712 dan 0,664, dengan *loading factor* seluruh butir $>0,5$ dan varians dijelaskan 52,3% dan 51,4%. Analisis ROC menunjukkan AUC untuk anak usia 7–10 tahun 54% dengan nilai ambang abatas adalah 5. Untuk anak usia 11–18 tahun, AUC ditemukan 69.9% nilai ambang batas 6.

Kesimpulan. PHQ-4 versi Bahasa Indonesia adalah kuesioner unidimensiaonal yang valid dan reliabel sebagai kuesioner deteksi dini gejala cemas dan depresi pada populasi anak dan remaja, meskipun studi selanjutnya masih diperlukan. **Sari Pediatri** 2026;28(1):8-18

Kata kunci: PHQ-4, validitas, reliabilitas, kecemasan, depresi, anak, remaja

Validity and Reliability of the Indonesian Version of the Patient Health Questionnaire-4 (PHQ-4) as a Screening Tool for Anxiety and Depressive Symptoms in Children and Adolescents

Tjhin Wiguna,¹ Kusuma Minayati,¹ Osi Kusuma Sari,² Yunita Arihandayani,² Imran Pambudi,² Wafa Sofia Fitri³

Background. Anxiety and depressive disorders among children and adolescents have reportedly increased, particularly in recent years, while brief and validated early screening instruments in the Indonesian language remain limited. The PHQ-4 is a brief early screening questionnaire with the potential for widespread use in the community.

Objective. To evaluate the validity and reliability of the Indonesian version of the PHQ-4 among children and adolescents aged 7–18 years.

Methods. This cross-sectional study involved 1846 participants from Jakarta, South Tangerang, and Serang; consisted of 440 children aged 7–10 years old and 1406 adolescents aged 11–18 years old. Content validity was assessed using the Content Validity Index (CVI), reliability using Cronbach's alpha, construct validity using Exploratory Factor Analysis (EFA), and criterion validity using ROC analysis with *The Mini International Neuropsychiatric Interview for Children and Adolescents* (MINI-KID). All analyses were conducted using SPSS version 29.

Results. All items of the Indonesian PHQ-4 for children aged 7–10 years showed an I-CVI ≥ 0.83 after modification into interrogative sentences, while the I-CVI for children aged 11–18 years was 1.00. Reliability testing of the PHQ-4 for children aged 7–10 years showed Cronbach's $\alpha = 0.644$, whereas for adolescents aged 11–18 years Cronbach's $\alpha = 0.693$. EFA results for both age groups demonstrated KMO values of 0.712 and 0.664, with all factor loadings > 0.5 and explained variances of 52.3% and 51.4%, respectively. ROC analysis showed that, among children aged 7–10 years, the AUC values were 54% with a cut-off score of 5. Among adolescents aged 11–18 years, the AUC values were 69.9% with a cut-off score of 6.

Conclusion. The Indonesian version of the PHQ-4 is a valid and reliable unidimensional questionnaire for the early detection of anxiety and depressive symptoms in children and adolescents, although further studies are still needed. **Sari Pediatri** 2026;28(1):8-18

Keywords: PHQ-4, validity, reliability, anxiety, depression, children, adolescents

Alamat korespondensi: Tjhin Wiguna. Departemen Ilmu Kesehatan Jiwa, Rumah Sakit Cipto Mangunkusumo-Fakultas Kedokteran Universitas Indonesia, Jl. Pangeran Diponegoro No.8, RW.5, Kenari, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10430. Email: twiga00@yahoo.com

Kesehatan jiwa merupakan komponen esensial pada anak dan remaja yang berperan penting dalam menentukan kualitas hidup, fungsi sosial, serta pencapaian akademik. Gangguan kecemasan dan depresi adalah masalah kesehatan jiwa paling sering ditemukan pada populasi usia muda, dengan *onset* umumnya dimulai sejak masa kanak-kanak hingga remaja.¹ Berbagai studi menunjukkan bahwa gangguan ini tidak hanya berdampak pada kesejahteraan psikologis, tetapi juga meningkatkan risiko gangguan perilaku, penyalahgunaan zat, penurunan prestasi akademik, bahkan bunuh diri pada usia dewasa.^{2,3}

Secara global, prevalensi distres psikologis seperti gangguan kecemasan dan depresi pada anak dan remaja terus meningkat. Kondisi ini diperburuk oleh berbagai faktor risiko, seperti tekanan akademik, perubahan lingkungan sosial, paparan media digital, serta faktor keluarga dan ekonomi.⁴ Di Indonesia, data Riset Kesehatan Dasar (*Riskesdas*) menunjukkan peningkatan prevalensi gangguan emosional dari 6% (2013) menjadi 9% (2018),⁵ yang mencerminkan meningkatnya beban masalah kesehatan jiwa di masyarakat.

Namun, deteksi dini gangguan kesehatan jiwa pada anak dan remaja masih menjadi tantangan. Salah satu hambatan utama ialah keterbatasan kuesioner deteksi dini yang praktis, valid, serta mudah digunakan di layanan kesehatan primer maupun sekolah. Alat ukur yang tersedia umumnya memiliki jumlah butir yang banyak, seperti *Strengths and Difficulties Questionnaire* (SDQ-25 butir), *Patient Health Questionnaire-9* (PHQ-9 butir), dan sebagainya, sehingga kurang efisien untuk deteksi dini pada populasi anak dan remaja maupun dalam praktik klinis dengan waktu terbatas.

Patient Health Questionnaire-4 (PHQ-4) merupakan kuesioner deteksi dini singkat yang dikembangkan dengan mengombinasikan dua butir dari PHQ-9 (depresi) dan dua butir dari GAD-7 (kecemasan).⁶ Dengan hanya empat pertanyaan, PHQ-4 memungkinkan identifikasi cepat distres psikologis pada dua subskala (kecemasan dan depresi) tanpa mengurangi akurasi secara signifikan. Instrumen PHQ-4 awalnya dikembangkan untuk dewasa, tetapi penelitian terdahulu menunjukkan bahwa instrumen ini juga dapat dimanfaatkan untuk anak dan remaja. Studi di Australia melaporkan bahwa PHQ-4 memiliki validitas dan reliabilitas yang baik, serta kemampuan diskriminasi yang memadai untuk mendeteksi distres psikologis pada populasi anak dan remaja.⁷

Penggunaan kuesioner dalam konteks lintas budaya memerlukan proses adaptasi dan validasi yang ketat. Perbedaan bahasa, budaya, serta tahap perkembangan kognitif anak dapat memengaruhi pemahaman terhadap butir kuesioner dan interpretasi respons. Hal ini terutama relevan pada anak usia sekolah dasar (SD) yang masih berada pada tahap operasional konkret sehingga berpotensi mengalami kesulitan memahami konsep abstrak.⁸

Di Indonesia, ketersediaan kuesioner deteksi dini distres psikologis yang spesifik untuk gejala depresi dan kecemasan, serta telah tervalidasi pada populasi anak dan remaja, masih terbatas. Selain itu, penelitian yang mengevaluasi performa psikometrik kuesioner pada berbagai kelompok usia dalam konteks budaya lokal juga masih jarang. Oleh karena itu, diperlukan penelitian yang tidak hanya melakukan adaptasi bahasa, tetapi juga mengevaluasi validitas konten, validitas konstruk, reliabilitas, serta kemampuan diskriminasi instrumen dalam mendeteksi gejala kecemasan dan depresi.

Penelitian ini bertujuan mengevaluasi validitas dan reliabilitas PHQ-4 versi Bahasa Indonesia sebagai kuesioner deteksi dini gejala kecemasan dan depresi pada anak dan remaja usia 7–18 tahun. Hasil penelitian ini diharapkan memberikan dasar ilmiah bagi penggunaan PHQ-4 sebagai kuesioner deteksi dini yang efisien, valid, dan aplikatif, terutama di layanan kesehatan primer maupun lingkungan sekolah di Indonesia.

Metode

Penelitian ini dirancang sebagai studi observasional analitik dengan desain potong lintang untuk mengevaluasi validitas dan reliabilitas kuesioner *Patient Health Questionnaire-4* (PHQ-4) versi Bahasa Indonesia. Penelitian dilaksanakan pada siswa sekolah dasar (SD), sekolah menengah pertama (SMP), dan sekolah menengah atas (SMA) di wilayah Jakarta, Tangerang Selatan, dan Kabupaten Serang selama periode Agustus hingga September 2025.

Sebanyak 1846 subjek penelitian dilibatkan, terdiri atas 440 anak usia 7–10 tahun dan 1406 remaja usia 11–18 tahun. Pemilihan subjek dilakukan dengan metode *consecutive sampling* di sekolah-sekolah yang berpartisipasi. Kriteria inklusi meliputi: (1) anak dan remaja yang bersedia mengikuti penelitian, dibuktikan

dengan persetujuan tertulis dari orang tua/wali (*informed consent*) dan persetujuan subjek (*assent*); (2) mampu mengisi kuesioner secara mandiri atau dengan pendampingan orang tua, terutama pada kelompok usia 7–10 tahun. Subjek dengan gangguan perkembangan berat, gangguan jiwa akut, atau gangguan komunikasi verbal yang dapat menghambat pemahaman saat mengisi kuesioner dikeluarkan dari penelitian (kriteria eksklusi).

Prosedur adaptasi instrumen kuesioner PHQ-4 versi Bahasa Indonesia dikembangkan melalui proses adaptasi lintas budaya yang sistematis dengan tahapan sebagai berikut.

- Penerjemahan dari bahasa Inggris ke dalam Bahasa Indonesia dilakukan oleh dua penerjemah independen yang kompeten.
- Diskusi panel ahli dilakukan untuk menilai kesesuaian terminologi, kejelasan makna, dan relevansi budaya, dengan panel yang terdiri atas 4 psikiater anak dan remaja, 2 psikiater, 4 dokter anak subspecialis tumbuh kembang dan pediatri sosial, serta 2 psikolog.
- Penerjemahan kembali (*back-translation*) ke dalam bahasa Inggris dilakukan oleh dua penerjemah independen lainnya untuk memastikan kesetaraan makna dengan versi asli.
- Diskusi panel ahli ulang dilaksanakan dengan komposisi ahli yang sama untuk menyelesaikan perbedaan interpretasi.
- Uji coba dilakukan pada sejumlah kecil subjek untuk mengevaluasi keterpahaman bahasa dan aspek budaya.
- Finalisasi kuesioner dilakukan berdasarkan hasil evaluasi seluruh tahapan sebelumnya.

Proses adaptasi ini bertujuan memastikan persamaan linguistik dan konseptual antara instrumen asli dan versi Bahasa Indonesia.

Kuesioner utama yang digunakan adalah PHQ-4, yang terdiri atas empat item dengan skala *Likert* 0–3; dengan demikian, total skor berkisar antara 0 hingga 12. Distres psikologis yang diukur oleh kuesioner ini mencakup dua subskala utama, yaitu gejala kecemasan dan gejala depresi. Sebagai standar baku emas (*gold standard*), *Mini International Neuropsychiatric Interview for Children and Adolescents* (MINI-KID) digunakan untuk menilai validitas kriteria. Domain yang digunakan dari MINI-KID adalah domain gangguan depresi dan domain gangguan cemas.

Untuk prosedur pengumpulan data dilakukan di lingkungan sekolah melalui pengisian kuesioner secara terstruktur di dalam kelas. Berdasarkan tahap perkembangan kognitif,⁹ subjek dibagi menjadi dua kelompok usia, yaitu 7–10 tahun dan 11–18 tahun. Pada kelompok usia 7–10 tahun, pengisian kuesioner dilakukan dengan pendampingan orang tua atau guru guna meningkatkan akurasi respons. Adapun pada kelompok usia 11–18 tahun, pengisian dilakukan secara mandiri. Seluruh subjek yang telah mengisi kuesioner ditanyakan kesediaannya untuk mengikuti wawancara klinis menggunakan MINI-KID, yang dilakukan oleh psikiater yang telah dilatih khusus untuk penelitian ini. Sebanyak 114 anak usia 7–10 tahun dan 217 remaja usia 11–18 tahun menyatakan bersedia mengikuti wawancara klinis tersebut. Data demografi yang dikumpulkan meliputi usia, jenis kelamin, tingkat pendidikan, karakteristik orang tua, serta riwayat kesehatan dan perkembangan.

Analisis data dilakukan dengan menggunakan *Statistical Package for the Social Sciences* (SPSS) versi 29.

- Evaluasi psikometrik instrumen dilakukan melalui beberapa tahapan berikut.
- Validitas konten dinilai dengan *Content Validity Index* (I-CVI dan S-CVI) berdasarkan penilaian panel ahli.
- Reliabilitas internal dievaluasi menggunakan koefisien *Cronbach's alpha* pada 440 anak usia 7–10 tahun dan 1406 remaja usia 11–18 tahun, dengan nilai $\geq 0,6$ dianggap dapat diterima untuk kuesioner deteksi dini singkat.
- Validitas konstruk dianalisis dengan *Exploratory Factor Analysis* (EFA) yang dilakukan pada 220 anak usia 7–10 tahun dan 703 remaja usia 11–18 tahun; kedua jumlah subjek tersebut dipilih secara acak dan telah memenuhi perhitungan kebutuhan subjek untuk EFA. Kriteria kelayakan dalam EFA ditetapkan, yaitu nilai *Kaiser-Meyer-Olkin* (KMO) $> 0,5$ dan uji *Bartlett's test* yang signifikan ($p < 0,05$).
- Validitas kriteria dinilai melalui analisis *Receiver Operating Characteristic* (ROC) dengan menghitung *Area Under the Curve* (AUC), sensitivitas, spesifisitas, serta menentukan titik potong optimal menggunakan indeks *Youden*.

Seluruh analisis dilakukan secara terpisah berdasarkan kelompok usia untuk mengevaluasi performa instrumen pada tahap perkembangan yang berbeda.

Penelitian ini telah mendapatkan persetujuan etik dari Komite Etik Penelitian Kesehatan Fakultas Kedokteran Universitas Indonesia–Rumah Sakit Cipto Mangunkusumo (FKUI-RSCM) dengan nomor KET-746/UN2.F1/ETIK/PPM.00.02/2025. Seluruh prosedur penelitian dilaksanakan sesuai dengan prinsip etik penelitian yang berlaku, termasuk penghormatan terhadap kerahasiaan data dan hak partisipan.

Hasil

Sebanyak 1846 subjek penelitian dilibatkan dalam penelitian ini, yang terdiri atas siswa SD, SMP, dan SMA (Tabel 1). Distribusi usia menunjukkan bahwa 440 subjek (24,38%) termasuk dalam kelompok usia 7–10 tahun, sedangkan 1406 subjek (75,62%) termasuk dalam kelompok usia 11–18 tahun. Berdasarkan jenis kelamin, 1.043 subjek (56,5%) berjenis kelamin perempuan dan 803 subjek (43,5%) berjenis kelamin laki-laki.

Karakteristik demografi orang tua menunjukkan bahwa sebagian besar orang tua (77,1%) memiliki tingkat pendidikan terakhir SMP dan SMA. Latar belakang sosial ekonomi orang tua sebagian besar (69,9%) berada pada kategori pendapatan sama dengan atau lebih dari upah minimum regional (UMR).

Hasil validitas konten menunjukkan bahwa sebagian besar butir pada PHQ-4 versi Bahasa Indonesia memiliki tingkat kesepakatan yang baik di antara panel ahli. Pada kelompok usia 7–10 tahun, putaran pertama panel ahli menghasilkan nilai I-CVI dan S-CVI yang kurang optimal dan tidak memenuhi nilai ambang yang ditetapkan. Konsultasi dengan ahli Bahasa Indonesia kemudian dilakukan, dan disarankan untuk memodifikasi kalimat pernyataan menjadi kalimat tanya. Setelah penilaian ahli putaran kedua pada kuesioner bentuk kalimat tanya, diperoleh nilai I-CVI 1,00 untuk butir 1, 0,67 untuk butir 2, serta 0,83 untuk butir 3 dan 4. Nilai S-CVI/UA sebesar 0,25 mengindikasikan adanya variasi penilaian antar ahli;

Tabel 1. Karakteristik Demografi Responden (n=1846)

Karakteristik	Kelompok usia 7–10 Tahun (n=440)	Kelompok usia 11–18 Tahun (n=1406)
Usia responden, rerata (SB)	8,74 (1,31)	15,08 (1,52)
Usia responden, median	9	15
Jenis kelamin, n (%)		
Laki-laki	206 (46,8%)	597 (42,5%)
Perempuan	234 (53,2%)	809 (57,5%)
Usia ayah, rerata (SB)	41,71 (7,14)	47,70 (6,12)
Usia ayah, median	41	47
Usia ibu, rerata (SB)	37,53 (5,80)	44,21 (5,72)
Usia ibu, median	37,5	44
Pendidikan terakhir orang tua, n (%)		
SD	311 (70,7%)	95 (6,8%)
SMP	81 (18,4%)	722 (51,4%)
SMA	38 (8,6%)	586 (41,7%)
D3	5 (1,1%)	1 (0,1%)
D4/S1	5 (1,1%)	2 (0,1%)
Wilayah sekolah, n (%)		
Jakarta	239 (54,3%)	990 (70,4%)
Tangerang Selatan	152 (34,5%)	151 (10,7%)
Serang	49 (11,1%)	265 (18,9%)
Pendapatan keluarga, n (%)		
≥ Upah minimum regional	237 (53,9%)	1055 (75,0%)
< Upah minimum regional	203 (46,1%)	351 (25,0%)

SB = Simpang Baku

oleh karena itu, revisi redaksional dilakukan kembali, khususnya pada butir pertanyaan nomor 2. Setelah revisi, panel ahli putaran ketiga dilakukan dan seluruh butir menunjukkan nilai I-CVI $\geq 0,83$ dengan nilai Kappa $\geq 0,816$, yang menandakan validitas konten yang memadai. Pada kelompok usia 11–18 tahun, seluruh butir memiliki nilai I-CVI 1,00 dengan S-CVI/Ave 1,00 dan nilai Kappa 1,00, yang menunjukkan kesepakatan penuh antar panel ahli dan validitas konten yang sangat baik.

Uji reliabilitas menunjukkan bahwa PHQ-4 versi Bahasa Indonesia memiliki konsistensi internal dalam kategori moderat. Pada kelompok usia 7–10 tahun, nilai Cronbach's α untuk PHQ-4 versi Bahasa Indonesia adalah 0,644. Untuk subskala gejala kecemasan, diperoleh Cronbach's α sebesar 0,548 dengan rerata korelasi antar-butir 0,379; sedangkan untuk subskala gejala depresi, Cronbach's α sebesar 0,597 dengan rerata korelasi antar-butir 0,428.

Pada kelompok usia 11–18 tahun, PHQ-4 versi Bahasa Indonesia memiliki nilai Cronbach's α 0,693. Adapun Cronbach's α untuk subskala gejala kecemasan sebesar 0,698 dengan rerata korelasi antar-butir 0,536, sedangkan untuk subskala gejala depresi diperoleh Cronbach's α sebesar 0,605 dengan rerata korelasi antar-butir 0,434. Temuan ini mengindikasikan bahwa reliabilitas instrumen lebih tinggi pada kelompok usia remaja dibandingkan dengan kelompok anak usia sekolah dasar.

Analisis faktor eksploratori (*Exploratory Factor Analysis*, EFA) pada kelompok anak usia 7–10 tahun ($n=220$) dan remaja usia 11–18 tahun ($n=703$) menunjukkan bahwa PHQ-4 versi Bahasa Indonesia merupakan kuesioner satu dimensi dengan validitas konstruk yang baik. Pada kelompok usia 7–10 tahun, nilai *Kaiser-Meyer-Olkin* (KMO) sebesar 0,712 dan uji Bartlett menunjukkan signifikansi ($p<0,001$), yang mengindikasikan kecukupan jumlah subjek untuk

PHQ-4 versi Bahasa Indonesia untuk anak usia 7–10 tahun

Komponen	Parameter	Nilai
KMO & Bartlett's Test	KMO	0,712
<i>Bartlett's Chi-square</i> (df=6)	499,195	
Nilai p	<0,001	
<i>Anti-image Correlation (MSA)</i>	PHQ-4_1	0,704
PHQ-4_2	0,683	
PHQ-4_3	0,768	
PHQ-4_4	0,739	
<i>Communalities (Extraction)</i>	PHQ-4_1	0,602
PHQ-4_2	0,669	
PHQ-4_3	0,258	
PHQ-4_4	0,564	
<i>Total Variance Explained</i>	<i>Component 1 Eigenvalue</i>	2,093
Varians dijelaskan (%)	52,325%	
Kumulatif (%)	52,325%	
<i>Factor Loadings (Component 1)</i>	PHQ-4_1	0,776
PHQ-4_2	0,818	
PHQ-4_3	0,508	
PHQ-4_4	0,751	

Tabel 3. Hasil analisis faktor eksploratori (EFA) PHQ-4 versi Bahasa Indonesia untuk remaja usia 11–18 tahun

Komponen	Parameter	Nilai
<i>KMO & Bartlett's Test</i>	KMO	0,664
<i>Bartlett's Chi-square (df=6)</i>	498,825	
Nilai p	<0,001	
<i>Anti-image Correlation (MSA)</i>	PHQ-4_1	0,653
PHQ-4_2	0,655	
PHQ-4_3	0,658	
PHQ-4_4	0,687	
<i>Communalities (Extraction)</i>	PHQ-4_1	0,581
PHQ-4_2	0,565	
PHQ-4_3	0,329	
PHQ-4_4	0,579	
<i>Total Variance Explained</i>	<i>Component 1 Eigenvalue</i>	2,055
Varians dijelaskan (%)	51,381%	
Kumulatif (%)	51,381%	
<i>Factor Loadings (Component 1)</i>	PHQ-4_1	0,762
PHQ-4_2	0,752	
PHQ-4_3	0,574	
PHQ-4_4	0,761	

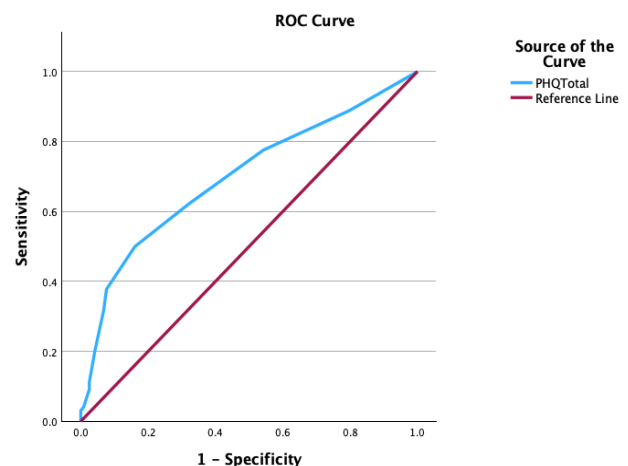
Validitas kriteria (analisis ROC)

analisis faktor. Faktor utama menjelaskan 52,3% varians total, dengan *loading factor* masing-masing butir berkisar antara 0,508 hingga 0,818. Pada kelompok usia 11–18 tahun, nilai KMO sebesar 0,664 dengan uji Bartlett juga signifikan ($p < 0,001$), dan varians yang dijelaskan sebesar 51,4%. *Loading factor* pada kelompok ini berkisar antara 0,574 hingga 0,762. Seluruh butir memiliki *loading factor* di atas 0,5, yang mengindikasikan bahwa PHQ-4 versi Bahasa Indonesia memiliki konstruk unidimensional yang stabil (Tabel 2 dan Tabel 3).

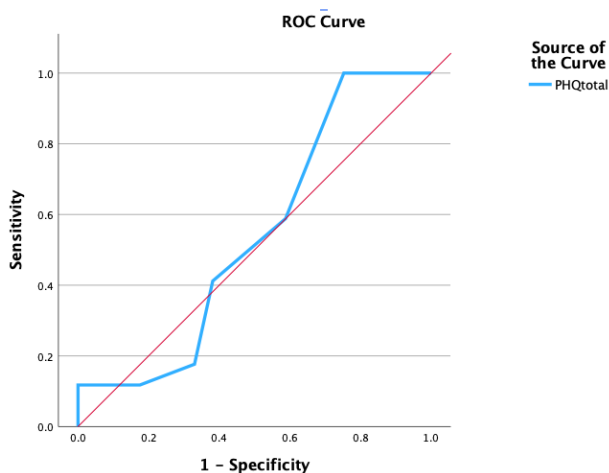
Tabel 2. Hasil analisis faktor eksploratori (EFA) validitas kriteria yang dinilai melalui analisis *Receiver Operating Characteristic* (ROC) menunjukkan bahwa PHQ-4 versi Bahasa Indonesia memiliki kemampuan diskriminasi yang memadai untuk deteksi dini gejala kecemasan dan depresi pada anak dan remaja. Pada kelompok usia 7–10 tahun ($n=114$), nilai *Area Under the Curve* (AUC) untuk deteksi dini distres psikologis (gejala kecemasan dan depresi) adalah 54%, dengan nilai batas ambang optimal 5, sensitivitas 1, spesifisitas 0,25, dan nilai indeks Youden 0,247. Adapun AUC untuk subskala gejala kecemasan adalah 74,4% dengan nilai batas ambang optimal 4, sensitivitas 0,67, dan indeks Youden 0,333; sedangkan untuk subskala gejala depresi, AUC sebesar 67,6% dengan nilai batas ambang

optimal 3, sensitivitas 0,58, dan indeks Youden 0,243 (Gambar 1).

Pada kelompok remaja usia 11–18 tahun ($n=217$), nilai AUC untuk distres psikologis adalah 69,9% dengan nilai batas ambang optimal 6, sensitivitas 0,62, spesifisitas 0,68, dan indeks Youden 0,300. Adapun AUC untuk subskala gejala kecemasan adalah 63,7%



Gambar 1. Hasil analisis ROC PHQ-4 untuk anak usia 7–10 tahun



Gambar 2. Hasil analisis ROC PHQ-4 untuk remaja usia 11–18 tahun

dengan nilai batas ambang optimal 4, sensitivitas 0,57, dan indeks Youden 0,259; sedangkan untuk subskala gejala depresi, AUC sebesar 69,3% dengan nilai batas ambang optimal 3, sensitivitas 0,89, dan indeks Youden 0,285 (Gambar 2).

Secara keseluruhan, nilai indeks Youden berkisar antara 0,243 hingga 0,337, yang menunjukkan performa diskriminasi dalam kategori cukup baik.

Pembahasan

Penelitian ini menunjukkan bahwa PHQ-4 versi Bahasa Indonesia memiliki validitas dan reliabilitas yang cukup memadai sebagai alat deteksi dini distress psikologis pada anak dan remaja usia 7–18 tahun, yang mencakup dua subskala utama, yaitu gejala kecemasan dan depresi. Secara umum, temuan ini konsisten dengan studi sebelumnya yang menempatkan PHQ-4 sebagai kuesioner deteksi dini distress psikologis yang singkat dengan performa psikometrik yang baik,^{7,10} meskipun terdapat beberapa perbedaan dalam nilai reliabilitas dan kemampuan diskriminasi.

Dari aspek validitas konten, seluruh butir PHQ-4 menunjukkan nilai I-CVI $\geq 0,83$ setelah proses adaptasi, bahkan mencapai 1,00 pada kelompok usia 11–18 tahun. Hal ini mengindikasikan kesesuaian konseptual yang tinggi antara versi Bahasa Indonesia dan kuesioner aslinya. Temuan ini sejalan dengan penelitian

Kocalevent dkk¹¹ dan Christodoulaki dkk¹² yang melaporkan bahwa PHQ-4 memiliki struktur analisis butir yang stabil dan representatif terhadap konstruk distress psikologis, termasuk untuk subskala gejala kecemasan dan depresi.⁶ Namun, pada kelompok usia 7–10 tahun, ditemukan kebutuhan revisi redaksional pada butir yang bersifat abstrak, khususnya terkait konsep “mengendalikan kekhawatiran”. Fenomena ini dapat dijelaskan melalui teori perkembangan kognitif Piaget,⁹ yang menyatakan bahwa anak usia sekolah dasar berada pada tahap operasional konkret sehingga memiliki keterbatasan dalam memahami konsep abstrak dan introspektif. Dengan demikian, adaptasi bahasa menjadi faktor krusial dalam mempertahankan validitas konten pada populasi anak.

Dari sisi reliabilitas, nilai *Cronbach's alpha* dalam penelitian ini berkisar antara 0,644 hingga 0,698, yang tergolong moderat namun masih dapat diterima untuk kuesioner deteksi dini singkat. Nilai ini sedikit lebih rendah dibandingkan studi Lushington dkk⁷ pada populasi yang sama (usia 7–18 tahun), yang melaporkan nilai *Cronbach's alpha* PHQ-4 sebesar 0,73 pada anak usia praremaja, 0,83 pada remaja pertengahan, dan 0,86 pada remaja akhir.⁷ Perbedaan ini kemungkinan disebabkan oleh beberapa faktor. Pertama, jumlah butir yang sangat sedikit (hanya dua butir per subskala) secara matematis membatasi nilai *Cronbach's alpha*, karena koefisien ini sangat dipengaruhi oleh jumlah butir dalam setiap skala.¹³ Kedua, variabilitas pemahaman emosi pada anak dan remaja, terutama pada usia yang lebih muda, dapat menyebabkan inkonsistensi respons. Ketiga, faktor lingkungan saat pengisian kuesioner di sekolah, seperti diskusi antarsiswa, berpotensi meningkatkan bias pengukuran. Oleh karena itu, nilai reliabilitas moderat dalam penelitian ini masih dapat diterima dalam konteks kuesioner deteksi dini di komunitas dan layanan primer.^{13,14}

Hasil analisis faktor eksploratori menunjukkan bahwa seluruh butir memiliki *loading factor* $> 0,5$ dengan varians yang dijelaskan sekitar 51–52%, yang mengindikasikan kemungkinan bahwa PHQ-4 versi Bahasa Indonesia berfungsi sebagai kuesioner *unidimensional* dengan kestabilan yang baik. Temuan ini konsisten dengan studi Lushington dkk⁷ yang menunjukkan *loading factor* serupa, sehingga memberikan gambaran bahwa PHQ-4 merupakan kuesioner *unidimensional* yang dapat digunakan untuk deteksi dini distress psikologis, meskipun terdiri atas dua subskala (gejala kecemasan dan depresi). Secara mekanistik, hal ini dapat dijelaskan

oleh adanya tumpang tindih neurobiologis antara gejala kecemasan dan depresi, termasuk disfungsi sistem limbik, disregulasi neurotransmitter seperti serotonin dan norepinefrin, serta aktivasi berlebihan pada aksis hipotalamus–hipofisis–adrenal (aksis HHA).¹⁵ Kondisi ini menyebabkan gejala kecemasan dan depresi sering muncul secara bersamaan dan sulit dipisahkan secara klinis, sehingga mendukung struktur faktor tunggal pada kuesioner deteksi dini singkat seperti PHQ-4.

Analisis validitas kriteria dalam penelitian ini menunjukkan nilai AUC berkisar antara 63,7% hingga 74,4%, yang mengindikasikan kemampuan diskriminatif sedang hingga baik untuk deteksi dini distres psikologis maupun deteksi dini subskala gejala kecemasan dan depresi pada populasi anak usia 7–18 tahun. Nilai ini lebih rendah dibandingkan dengan studi pada populasi dewasa, yang melaporkan AUC PHQ-4 dapat mencapai 78,7–83,5%.¹⁵ Perbedaan ini dapat dijelaskan oleh beberapa faktor. Pertama, pada populasi anak dan remaja, ekspresi gejala emosional sering kali lebih heterogen dan tidak selalu sesuai dengan kriteria diagnostik klasik. Misalnya, anak dan remaja dengan depresi lebih sering menunjukkan *iritabilitas* dibandingkan perasaan sedih.¹⁶ Kedua, keterbatasan kemampuan introspeksi pada anak dapat memengaruhi akurasi *self-report*. Ketiga, penggunaan MINI-KID sebagai standar baku yang berbasis wawancara klinis dapat menghasilkan diskrepansi dengan kuesioner deteksi dini berbasis *self-report*.^{17,18}

Menariknya, pada kelompok usia 7–10 tahun, sensitivitas untuk deteksi dini gejala kecemasan mencapai 74,7% pada nilai batas ambang 4, yang menunjukkan bahwa PHQ-4 sensitif dalam mendeteksi kasus pada populasi ini. Akan tetapi, spesifisitasnya relatif lebih rendah, sehingga terdapat potensi *positif palsu* yang lebih tinggi. Secara klinis, hal ini masih dapat diterima dalam konteks deteksi dini, di mana prioritas utama ialah tidak melewatkan kasus (*high sensitivity*).¹⁹ Sebaliknya, pada kelompok remaja, sensitivitas PHQ-4 tampak lebih rendah, yang mengindikasikan bahwa PHQ-4 mungkin kurang sensitif dalam mendeteksi kasus ringan atau subklinis pada kelompok usia ini.

Dari perspektif implementasi klinis, temuan ini mendukung penggunaan PHQ-4 sebagai alat deteksi dini distres psikologis, termasuk gejala kecemasan dan depresi, yang efisien terutama di layanan kesehatan primer dan lingkungan sekolah. Dengan waktu pengisian kurang dari lima menit, PHQ-4 merupakan kuesioner deteksi dini distres psikologis yang lebih

mudah dikerjakan oleh populasi anak dan remaja. Namun demikian, mengingat nilai reliabilitas yang moderat dan kemampuan diskriminasi yang masih dalam taraf cukup baik, hasil deteksi dini yang positif sebaiknya selalu diikuti dengan evaluasi lanjutan menggunakan wawancara psikiatri klinis atau alat ukur diagnostik yang lebih komprehensif.

Selain itu, dalam konteks pelayanan kesehatan di Indonesia, hasil deteksi dini yang positif menggunakan PHQ-4 dapat ditindaklanjuti melalui sistem rujukan berjenjang yang terintegrasi dengan layanan kesehatan primer, seperti puskesmas, klinik, maupun layanan kesehatan sekolah. Individu dengan indikasi distres psikologis dapat dirujuk ke psikolog klinis atau psikiater di fasilitas kesehatan rujukan tingkat lanjut sesuai kebutuhan klinis dan mekanisme yang berlaku. Akan tetapi, implementasi di lapangan masih menghadapi berbagai kendala, seperti keterbatasan tenaga kesehatan jiwa, distribusi layanan psikiatri yang belum merata terutama di daerah pedesaan, rendahnya pemanfaatan sistem rujukan akibat kurangnya pelatihan tenaga kesehatan, serta stigma masyarakat terhadap gangguan mental yang dapat menghambat akses layanan lanjutan. Upaya mengatasi stigma dapat dilakukan melalui edukasi kesehatan mental berbasis sekolah dan komunitas, peningkatan literasi kesehatan mental pada orang tua dan guru, serta penyediaan layanan konseling yang bersifat ramah remaja dan menjaga kerahasiaan individu. Pendekatan ini penting untuk meningkatkan penerimaan terhadap layanan kesehatan jiwa, sehingga individu yang teridentifikasi mengalami distres psikologis lebih bersedia mencari bantuan profesional sejak dini. Oleh karena itu, optimalisasi deteksi dini perlu disertai penguatan sistem rujukan dan edukasi kesehatan mental agar proses identifikasi hingga intervensi dapat berjalan lebih efektif dan berkesinambungan.²⁰

Penelitian ini memiliki beberapa keterbatasan. Pertama, desain potong lintang tidak memungkinkan evaluasi reliabilitas *test-retest* atau sensitivitas terhadap perubahan klinis. Kedua, terdapat potensi bias pengisian kuesioner di lingkungan sekolah. Ketiga, distribusi sosial ekonomi responden yang didominasi oleh kelompok tertentu dapat membatasi generalisasi hasil. Meskipun demikian, ukuran subjek penelitian yang besar dan cakupan yang baik di sekolah negeri maupun swasta menjadi kekuatan utama penelitian ini.

Secara keseluruhan, hasil penelitian ini memperkuat bukti bahwa PHQ-4 versi Bahasa Indonesia merupakan kuesioner deteksi dini distres psikologis yang valid,

reliabel, praktis, dan relevan secara klinis untuk populasi anak dan remaja di Indonesia, dengan tetap mempertimbangkan keterbatasannya sebagai kuesioner deteksi dini yang singkat.

Kesimpulan

Penelitian ini menunjukkan bahwa instrumen PHQ-4 versi Bahasa Indonesia memiliki validitas dan reliabilitas yang cukup memadai sebagai kuesioner deteksi dini distress psikologis, termasuk subskala gejala kecemasan dan depresi, pada anak dan remaja usia 7–18 tahun. Validitas konten yang tinggi ($I-CVI \geq 0,83$ dan mencapai 1,00 pada kelompok remaja) menegaskan kesesuaian adaptasi dalam Bahasa Indonesia yang baik. Reliabilitas internal berada pada kategori moderat (*Cronbach's alpha* 0,644–0,698), dengan nilai lebih tinggi pada remaja dibandingkan anak usia praremaja. Analisis faktor menunjukkan bahwa seluruh butir memiliki loading factor $>0,5$ dengan varians yang dapat dijelaskan sekitar 51–52%, yang mendukung struktur unidimensional. Analisis ROC menunjukkan kemampuan diskriminasi yang cukup hingga baik, dengan AUC berkisar antara 54% hingga 69,9%.

Namun, perlu dipertimbangkan bahwa pada anak usia 7–10 tahun, AUC PHQ-4 untuk deteksi distress psikologis hanya sebesar 54%, sehingga penggunaan nilai masing-masing subskala secara terpisah lebih direkomendasikan untuk kelompok usia ini. Sebaliknya, pada remaja usia 11–18 tahun, nilai AUC untuk distress psikologis mencapai 69,9% dengan nilai batas ambang optimal 6, dan nilai AUC untuk masing-masing subskala juga cukup memadai. Secara keseluruhan, PHQ-4 dapat direkomendasikan sebagai kuesioner deteksi dini yang cukup efisien di layanan kesehatan primer dan sekolah. Namun, karena penggunaannya bersifat sebagai alat deteksi dini, hasil positif tetap memerlukan evaluasi lanjutan melalui wawancara klinis psikiatri yang memadai sebagai alat diagnostik.

Daftar pustaka

- Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J Child Psychol Psychiatry* 2015;56:345-65. Doi:10.1111/jcpp.12381
- Pine DS, Cohen P, Gurley D, Brook J, Ma Y. The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Arch Gen Psychiatry* 1998;55:56-64. Doi:10.1001/archpsyc.55.1.56
- Clayborne ZM, Varin M, Colman I. Systematic review and meta-analysis: adolescent depression and long-term psychosocial outcomes. *J Am Acad Child Adolesc Psychiatry* 2019;58:72-9. Doi:10.1016/j.jaac.2018.07.896
- Bunting L, Nolan E, McCartan C, dkk. Prevalence and risk factors of mood and anxiety disorders in children and young people: findings from the Northern Ireland youth wellbeing survey. *Clin Child Psychol Psychiatry* 2022;27:686-700. Doi:10.1177/13591045221089841
- Kementerian Kesehatan RI. Laporan nasional riskesdas 2018. Published online 2018. Didapat dari: <https://repository.kemkes.go.id/book/1323>
- Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics* 2009;50:613-21. Doi:10.1176/appi.psy.50.6.613
- Lushington K, Washer S, Wicking P, Wicking A, Dorrian J. Validation of the PHQ-4 as a unidimensional screener in 191,519 Australian children and adolescents. *J Affect Disord* 2026;400:121216. Doi:10.1016/j.jad.2025.121216
- Borgers N, de Leeuw E, Hox J. Children as respondents in survey research: cognitive development and response quality. *Bull Sociol Methodol* 2000;66:60-75. Doi:10.1177/075910630006600106
- Piaget J. The stages of the intellectual development of the child. *Bull Menninger Clin* 1962;26:120-8.
- Löwe B, Wahl I, Rose M, dkk. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affect Disord* 2010;122:86-95. Doi:10.1016/j.jad.2009.06.019
- Kocalevent RD, Finck C, Jimenez-Leal W, Sautier L, Hinz A. Standardization of the Colombian version of the PHQ-4 in the general population. *BMC Psychiatry* 2014;14:205. Doi:10.1186/1471-244X-14-205
- Christodoulaki A, Baralou V, Konstantakopoulos G, Touloumi G. Validation of the Patient Health Questionnaire-4 (PHQ-4) to screen for depression and anxiety in the Greek general population. *J Psychosom Res* 2022;160:110970. Doi:10.1016/j.jpsychores.2022.110970
- Kim HW, Shin C, Lee SH, Han C. Standardization of the Korean version of the Patient Health Questionnaire-4 (PHQ-4). *Clin Psychopharmacol Neurosci* 2021;19:104-11. Doi:10.9758/cpn.2021.19.1.104
- Eisinga R, te Grotenhuis M, Pelzer B. The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *Int J Public Health* 2013;58:637-42. Doi:10.1007/s00038-012-0416-3
- Bautista J, Hidalgo-Tinoco C, Di Capua Delgado M, Viteri-Recalde J, Guerra-Guerrero A, López-Cortés A. The gut-brain-circadian axis in anxiety and depression: a critical review. *Front Psychiatry* 2025;16:1697200. Doi:10.3389/fpsy.2025.1697200

16. Khubchandani J, Brey R, Kotecki J, Kleinfelder J, Anderson J. The psychometric properties of PHQ-4 depression and anxiety screening scale among college students. *Arch Psychiatr Nurs* 2016;30:457-62. Doi:10.1016/j.apnu.2016.01.014
17. Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. *Lancet* 2012;379:1056-67. Doi:10.1016/S0140-6736(11)60871-4
18. Sheehan DV, Sheehan KH, Shytle RD, dkk. Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). *J Clin Psychiatry* 2010;71:313-26. Doi:10.4088/JCP.09m05305whi
19. Grimes DA, Schulz KF. Uses and abuses of screening tests. *Lancet* 2002;359:881-4. Doi:10.1016/S0140-6736(02)07948-5
20. Kementerian Kesehatan Republik Indonesia. Pedoman penyelenggaraan pelayanan kesehatan jiwa di fasilitas pelayanan kesehatan primer. Jakarta: Kemenkes RI; 2022.

Lampiran

Kuesioner PHQ-4 untuk anak usia 7-10 tahun

Instruksi	Tidak sama sekali (0)	Beberapa hari (1)	Lebih dari separuh waktu/hari (2)	Hampir setiap hari (3)
Selama <u>2 minggu terakhir</u> , seberapa sering kamu terganggu oleh masalah-masalah di bawah ini?				
Apakah kamu sering merasa takut, jantung berdebar, atau tidak tenang?				
Apakah kamu sering merasa khawatir terus-menerus meski sudah mencoba berhenti?				
Apakah kamu sudah tidak terlalu senang melakukan hal-hal yang biasanya kamu suka (misalnya main, menggambar, bermain dengan teman)?				
Apakah kamu sering merasa sedih atau tidak bersemangat?				

Kuesioner PHQ-4 untuk remaja usia 11-18 tahun

Instruksi	Tidak sama sekali (0)	Beberapa hari (1)	Lebih dari separuh waktu/hari (2)	Hampir setiap hari (3)
Selama <u>2 minggu terakhir</u> , seberapa sering Anda terganggu oleh masalah-masalah berikut?				
Merasa gugup, cemas, atau gelisah				
Tidak dapat menghentikan atau mengendalikan rasa cemas				
Tidak terlalu berminat atau senang melakukan sesuatu				
Merasa sedih, depresi, atau putus asa				