Canadian Psychiatric Association
Association des psychiatres du Canada

Original Research

The Canadian Journal of Psychiatry / La Revue Canadienne de Psychiatrie
2019, Vol. 64(9) 647-657
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0706743718818414
TheCJP.ca | LaRCP.ca
SAGE

The Psychometric Properties of the Kessler Psychological Distress Scale (K6) in an Epidemiological Sample of Canadian Youth

Les propriétés psychométriques de l'échelle de détresse psychologique de Kessler (K6) dans un échantillon épidémiologique d'adolescents canadiens

Mark A. Ferro, PhD[1]

Abstract

Objective: This study (1) describes the distribution of Kessler Psychological Distress Scale (K6) scores in an epidemiological sample of youth, (2) reports its item correlations and internal consistency reliability, (3) tests for measurement invariance by age (youth vs. adults) and sex, and (4) examines its predictive power for past-year psychiatric disorders.

Method: Youth aged 15 to 19 years ($n = 2010$) and adults aged 20 to 64 years ($n = 2010$) from the Canadian Community Health Survey–Mental Health who completed the K6 were included. Past-year psychiatric disorders were measured using the World Health Organization Composite International Diagnostic Interview 3.0 (WHO-CIDI). Polychoric correlation matrices and ordinal reliability coefficients were calculated to assess internal consistency of the K6, and confirmatory factor analysis was used to test for measurement invariance. Area under the curves (AUCs) were computed to determine the extent to which the K6 predicted a positive screen on the WHO-CIDI.

Results: K6 scores showed a J-shaped distribution, with >50% of youth having scores ≤ 3. Item and total scores were higher for females versus males. Item correlations were robust (0.31-0.78) and internal consistency was high ($\alpha = 0.86$). Full measurement invariance was demonstrated between youth and adults, as well as between male and female youth. The K6 was a strong predictor of major depressive episode (AUC = 0.848), generalized anxiety disorder (AUC = 0.847), and bipolar disorder (AUC = 0.853).

Conclusions: The K6 is a valid and reliable measure of psychological distress among youth. Its brevity and robust predictive power for psychiatric disorder confirm its utility in clinical and community settings to identify youth needing comprehensive psychiatric assessment.

Abrégé

Objectif : Cette étude: (1) décrit la distribution des scores à l'échelle de détresse psychologique de Kessler (K6) dans un échantillon épidémiologique d'adolescents; (2) rend compte des corrélations de ses items et de la fiabilité de la cohérence interne; (3) vérifie l'invariance de la mesure selon l'âge (adolescents c. adultes) et le sexe; et (4) examine son pouvoir de prédiction des troubles psychiatriques de l'année précédente.

Méthode : Des adolescents de 15 à 19 ans ($n = 2010$) et des adultes de 20 à 64 ans ($n = 2010$) de l'Enquête sur la santé dans les collectivités canadiennes–Santé mentale qui ont répondu à la K6 ont été inclus. Les troubles psychiatriques de l'année précédente ont été mesurés à l'aide de l'Entrevue composite diagnostique internationale de l'Organisation mondiale de la

[1] School of Public Health and Health Systems, University of Waterloo, Ontario

Corresponding Author:
Mark Ferro, University of Waterloo, School of Public Health and Health Systems, 200 University Avenue West, Waterloo, Ontario N2 L 3G1, Canada.
Email: mark.ferro@uwaterloo.ca

santé 3.0 (WHO-CIDI). Les matrices des corrélations polychoriques et les coefficients de fiabilité ordinaux ont été calculés pour évaluer la cohérence interne de la K6, et l'analyse factorielle confirmatoire a servi à vérifier l'invariance de la mesure. Les zones sous la courbe (ZSC) ont été calculées pour déterminer à quel degré la K6 prédisait un dépistage positif à la WHO-CIDI.

Résultats : Les scores à la K6 présentaient une distribution en forme de J, et >50% des adolescents avaient des scores $\leq$3. Les scores des items et totaux étaient plus élevés chez les femmes que chez les hommes. Les corrélations des items étaient fortes (0,31-0,78) et la cohérence interne était élevée ($\alpha = 0,86$). L'invariance complète de la mesure était démontrée entre les adolescents et les adultes, ainsi qu'entre les adolescents et les adolescentes. La K6 prédisait fortement l'épisode dépressif majeur (ZSC = 0,848), le trouble d'anxiété généralisée (ZSC = 0,847), et le trouble bipolaire (ZSC = 0,853).

Conclusions : La K6 est une mesure valide et fiable de la détresse psychologique chez les adolescents. Sa concision et son fort pouvoir de prédiction des troubles psychiatriques en confirment l'utilité en milieu clinique et communautaire pour repérer les adolescents nécessitant une évaluation psychiatrique complète.

Keywords
adolescent, measurement, mental disorder, mental health, validity

The Kessler Psychological Distress Scale (K6) is a psychometrically robust measure of psychological distress for adult populations.[1] It was developed specifically to measure the prevalence of severe mental illness according to the Alcohol, Drug Abuse, and Mental Health Administration Reorganization Act in the United States.[1] Since then, the K6 has become the most commonly used screening measure in the general population. It outperforms other screening measures (e.g., General Health Questionnaire) in predicting severe mental illness and consequently is used in epidemiological surveys globally, including the World Health Organization (WHO) World Mental Health Survey.[2] Recently, there has been a research interest in examining the psychometric properties of the K6 in adolescent and youth populations.

Studies have reported the distribution of K6 scores among youth to be similar to that observed in adults—the majority of youth have no or low levels of psychological distress.[3-7] Internal consistency reliability of the K6 has been reported to be adequate ($\alpha = 0.78$ to 0.90),[4-7] suggesting that its items load onto a single psychological distress construct—a finding confirmed in studies examining the factor structure of the K6 in youth samples.[3,5-7] Research into the measurement invariance of the K6 across subgroups of youth has produced mixed results. One study found the K6 to be fully invariant across sex, age, and race in a school-based sample of American youth.[8] In contrast, a school sample of Australian youth demonstrated a lack of invariance between sexes, suggesting that male and female youth may interpret K6 items and the psychological distress construct differently.[5] Thus, observed differences between male and female youth may not reflect true differences but instead be an artifact of differential interpretation.

Determining the extent to which the K6 is invariant across subgroups of youth is critical to establishing its validity in this population; however, notably absent in the literature is research examining whether the K6 is invariant between youth and adult samples. Because the K6 was initially designed to measure psychological distress among adults, research is needed to confirm its intended use among youth. Doing so will allow researchers to examine differences in developmental changes from adolescence to adulthood across populations. Whilst preliminary evidence (similar distribution, internal consistency, and factor structure) suggests that the K6 is performing as intended in youth population, direct evidence—via measurement invariance testing—is needed.

Similar to findings from adult populations,[2] the K6 is a robust predictor of severe mental illness among youth,[7] particularly for mood and anxiety disorders[3] and less so for behaviour disorders.[5] Using data from the National Comorbidity Survey Replication Adolescent Supplement (NCS-A),[4] the K6 was moderately predictive of past-year *Diagnostic and Statistical Manual of Mental Disorders IV* (*DSM-IV*) major depressive episode (area under the curve [AUC] = 0.77), bipolar disorder (AUC = 0.74), and generalized anxiety disorder (AUC = 0.77) as measured using the Composite International Diagnostic Interview. That same study showed that the K6 only weakly predicted behaviour disorders, including attention-deficit hyperactivity disorder (AUC = 0.58), oppositional defiant disorder (AUC = 0.68), and conduct disorder (AUC = 0.71).

Use of the K6 as a valid and reliable measure of psychological distress among youth populations is promising; however, there remains an absence of evidence of its psychometric properties among Canadian youth. Research investigating the appropriateness of the K6 among Canadian youth is important. Contemporary estimates from epidemiological studies indicate that the prevalence of mental disorder among youth is substantial, with approximately 1 in 5 youth affected.[9-11] Disability related to psychiatric disorder peaks[12] and more than half of all psychiatric disorders in adulthood have their onset during adolescence.[13] Thus, the ability to accurately, reliably, and efficiently identify youth who require more comprehensive assessment of their mental health is paramount to reducing (1) the time needed for diagnosis, (2) delays in accessing appropriate psychiatric services, and (3) individual, family, and systems burdens of psychiatric disorder.

In a national epidemiological study of Canadian youth aged 15 to 19 years, the objectives of this study were to (1) describe the distribution of K6 scores, (2) report item correlations and internal consistency reliability of the K6, (3) test for measurement invariance by age (youth vs. adults) and sex (males vs. females [youth sample only]), and (4) examine the predictive power of the K6 with past-year *DSM-IV* psychiatric disorders.

Method

Data Source

Data were obtained from the public use file of 2012 Canadian Community Health Survey–Mental Health (CCHS-MH), an epidemiological study conducted by Statistics Canada to collect information about the prevalence of psychiatric disorder, functional status, and the use of health services among Canadians.[14] Using multistage, stratified, cluster sampling, a representative sample of respondents ≥15 years of age were enrolled ($N = 25{,}113$). In-person or telephone computer-assisted interviews were conducted resulting in a combined household and person response of 69%.[14] In all selected households, a knowledgeable household member was asked to supply basic demographic information on all residents residing in the household. One respondent ≥15 years of age was then selected for a more in-depth interview. To ensure data quality, interviewers were instructed to make every effort to conduct the interview with the selected respondent in privacy. In situations where this was unavoidable, the respondent was interviewed with another person present. Participation in the CCHS-MH was voluntary and confidentiality guaranteed by Statistics Canada.

The public use file used for this study was developed from the Statistics Canada master file such that the most useful data possible are provided while protecting participant confidentiality.[15] Variables most likely to lead to identification of an individual are deleted or collapsed to broader categories. For this study, relevant variables were available for analysis, including those that described the complex sampling (i.e., stratified sampling by province), and given the target population, categorization of participants into a 15- to 19-year-old age bin was acceptable.

Sample

The sample was stratified into 2024 youth (15 to 19 years) and 16,972 adults (20 to 64 years). Individuals in older adulthood (≥65 years) were removed due to differences in the phenomenology of mental disorder compared to younger individuals.[16] Missing K6 data were minimal, and these individuals were removed from the study: $n = 14$ (0.7%) youth and $n = 79$ (0.5%) adults. There were no correlates of missingness for either group. The number of youth available for analysis was 2010 ($n_{\text{weighted}} = 2015$). A simple random sample of adults ($n = 2010$; $n_{\text{weighted}} = 2084$) was selected for testing invariance of the K6 with youth. Compared to this random sample of adults, youth were more likely to be born in Canada ($\chi^2[1] = 51.08$; $P < 0.001$), be nonwhite ($\chi^2[1] = 6.18$; $P = 0.013$), have lower educational attainment ($\chi^2[3] = 796.63$; $P < 0.001$), and have better perceived health ($\chi^2[4] = 22.41$; $P < 0.001$). Given that age is correlated with these characteristics, such differences are not unexpected and do not indicate a meaningful discrepancy between the groups. There were no differences with regards to the distribution of sex ($\chi^2[1] = 1.02$; $P = 0.312$) or household income ($\chi^2[4] = 5.79$; $P = 0.212$) between groups.

Measures

The K6 is a nonspecific measure of psychological distress consisting of 6 questions asking participants if they had felt nervous, hopeless, restless, or fidgety; so depressed that nothing could cheer you up; that everything was an effort; and worthless.[1] Preceded by the statement, "During the past month, that from (date one month ago) to yesterday, about how often did you feel . . .," participants were asked to respond to each of the items using a 5-point scale that ranged from 1 (*all of the time*) to 5 (*none of the time*). Responses were then recoded (0 to 4) such that higher scores indicated more psychological distress (range, 0 to 24).

The WHO version of the Composite International Diagnostic Interview 3.0 (WHO-CIDI) was used to measure the presence of major depressive episode, generalized anxiety disorder, and bipolar disorder in the 1 year immediately prior to the CCHS-MH interview according to the *DSM-IV* and *International Classification of Diseases, Tenth Revision*.[17] The WHO-CIDI consists of a screening and lifetime review module that allows early termination for individuals who show no evidence of lifetime psychopathology, thereby reducing burden. The reliability and validity of the WHO-CIDI are well established.[17,18]

Demographic characteristics relating to ethnicity (white/nonwhite), immigrant status (born/not born in Canada), education (currently enrolled school or not), household income (<$20,000 to ≥$80,000 in $20,000 increments), and perceived health (poor, fair, good, very good, excellent) were measured in the CCHS-MH and compared between sexes in the sample of youth in this study.

Statistical Analysis

Male and female youth were compared across demographic characteristics using Rao-Scott χ^2 tests. K6 percentiles and mean item and total scores were computed, stratified by sex. Male and female mean K6 scores were compared using *t* tests. Due to its ordinal response pattern, polychoric correlation matrices and ordinal reliability coefficients were calculated to assess internal consistency of the K6.[19]

Multiple-group confirmatory factor analysis employing a weighted least squares means and variance-adjusted estimator was used to test for measurement invariance of the K6

Table 1. Characteristics of the Study Sample.

Characteristic	All Youth ($n = 2015$)	Males ($n = 1018$)	Females ($n = 997$)	χ^2 (P Value)
White	1392 (69.3)	674 (66.3)	719 (72.3)	3.73 (.053)
Immigrant	246 (12.2)	124 (12.2)	122 (12.3)	0.01 (.994)
Currently in school	1682 (83.6)	835 (82.1)	847 (85.0)	1.63 (.203)
Household income				8.02 (.091)
<$20,000	108 (5.4)	50 (5.0)	58 (5.8)	
$20,000-$39,999	238 (11.8)	115 (11.3)	123 (12.4)	
$40,000-$59,999	292 (14.5)	137 (13.5)	154 (15.5)	
$60,000-$79,999	347 (17.2)	151 (14.9)	196 (19.7)	
≥$80,000	1028 (51.1)	563 (55.4)	465 (46.7)	
Perceived health				3.91 (.418)
Poor	7 (0.4)	2 (0.2)	5 (0.6)	
Fair	74 (3.7)	34 (3.4)	40 (4.0)	
Good	576 (28.6)	302 (29.6)	274 (27.5)	
Very good	871 (43.3)	420 (41.3)	452 (45.4)	
Excellent	485 (24.1)	260 (25.5)	225 (22.6)	

Data are reported as weighted frequency (percent).

between youth and the random adult sample, as well as test for invariance between sexes among the youth sample.[20] Increasingly stringent equality constraints were specified for model parameters between groups (i.e., youth vs. adults; males vs. females). First, configural invariance imposed no equality constraints on parameters and was the origin for subsequent tests.[21] Second, weak invariance examined the extent to which the factor loadings for particular items were equivalent between groups. This is a prerequisite for making valid comparisons.[22] Third, strong invariance tested for evidence that item thresholds were equivalent between groups.[23] Strong invariance verifies whether mean differences at the item level are fully explained by mean differences at the factor level. Fourth, strict invariance was performed to determine whether the variances of the regression equations for each item were equivalent across groups. Strict invariance is required for defensible item score comparisons (i.e., average item scores) between groups.[24] Model fit was assessed using the comparative fit index (CFI; ≥0.95 considered adequate) and the root mean square error of approximation (RMSEA; ≤0.08 considered adequate). Measurement invariance was considered established when there was a nonsubstantial worsening of model fit according to the following criteria: ΔCFI ≤–0.01 or ΔRMSEA ≥0.015.[25] Both criteria needed to be satisfied. Unlike the χ^2 (which is included only for completeness), changes in the CFI and RMSEA are not influenced by sample size.[25]

Using receiver operating characteristic curve analysis, area under the curve and associated 95% confidence intervals (CIs) were computed to determine the extent to which the K6 predicted a positive screen for past-year major depressive episode, generalized anxiety disorder, or bipolar disorder. Relative sampling weights were used to ensure comparability between the sample and the Canadian population and were applied to all analyses. Taylor series linearization, specifying province as the stratification factor, was used to adjust estimates of variance due to the complex sampling for the CCHS-MH. Data were analyzed using SAS 9.4 and Mplus 6.11.

Results

Sample Characteristics

There was an equal distribution of males (50.5%) and females in the youth sample. Over two-thirds (69.3%) were white and 12.2% were immigrants. The majority (83.6%) were currently enrolled in school and 51.1% reported annual household incomes ≥$80,000. The majority of youth perceived their health as "very good" or "excellent." There were no statistically significant differences in these characteristics between males and females. Characteristics of the sample, stratified by sex, are shown in Table 1.

Distribution of K6 Scores

Figure 1 shows the distribution of K6 scores and cumulative prevalence by sex. A typical J-shape distribution was observed with 56.0% of youth scoring ≤3 (60.0% for males; 51.7% for females). Mean item and total scores were all significantly higher among females compared to males (4.43 vs. 3.43; Table 2).

K6 Reliability

Polychoric correlations among K6 items were moderate to strong for males (0.31 to 0.76) and females (0.36 to 0.78; Table 3). Internal consistency was high in the sample ($\alpha = 0.86$) and in both males ($\alpha = 0.85$) and females ($\alpha = 0.87$) specifically.

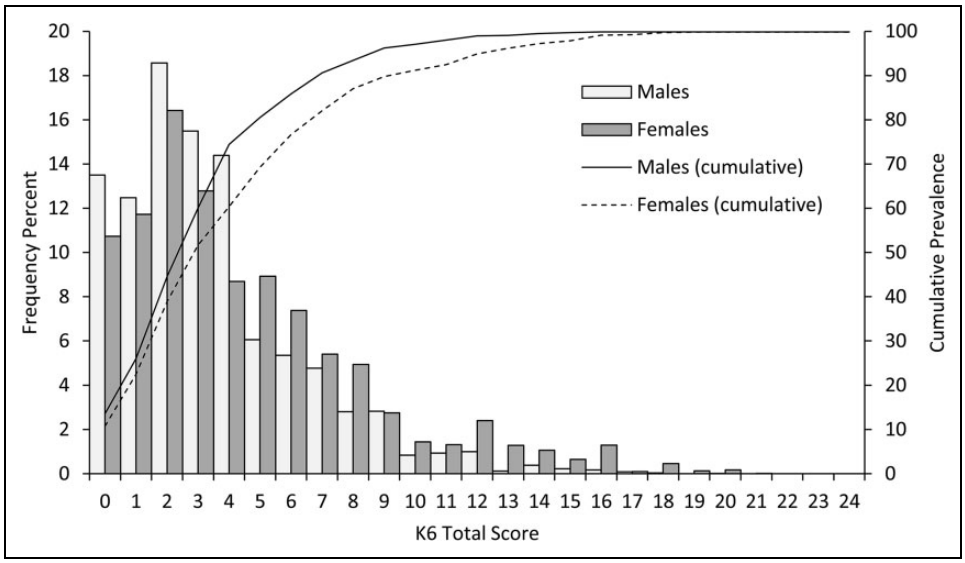

Figure 1. Distribution and cumulative prevalence of K6 scores.

Table 2. Percentiles and Mean K6 Item and Total Scores.

Percentile	All Youth	Males	Females
10th	0	0	0
25th	2	1	2
50th	3	3	3
75th	6	5	6
90th	8	7	10
95th	11	9	13
99th	16	13	16
Mean (SD)			
Nervous	3.87 (0.93)	4.00 (0.90)	3.74 (0.93)
Hopeless	4.65 (0.72)	4.74 (0.59)	4.55 (0.81)
Restless	3.97 (1.02)	4.04 (1.00)	3.89 (1.03)
Depressed	4.75 (0.63)	4.82 (0.53)	4.68 (0.71)
Effort	4.13 (1.03)	4.17 (1.02)	4.08 (1.03)
Worthless	4.71 (0.69)	4.80 (0.57)	4.62 (0.77)
Total	3.93 (3.44)	3.43 (2.97)	4.43 (3.77)

Values in the top panel represent K6 scores that correspond to the listed percentile. All mean item and total K6 score comparisons were statistically significant at $P < 0.001$, with the exception of "Effort," which was significant at $P < 0.05$.

Measurement Invariance of the K6

Given established evidence that the K6 best fits a one-factor model, we proceeded directly to fitting the confirmatory factor model (i.e., exploratory analyses of alternative factor models were not tested). As shown in Table 4, the fit of the one-factor model was good for the youth (CFI = 0.993; RMSEA = 0.052 [0.040 to 0.066]) and adult samples (CFI = 0.999; RMSEA = 0.028 [0.014 to 0.042]). Equality constraints placed on the factor loadings (weak model) did not substantially worsen model fit: ΔCFI = –0.001 and ΔRMSEA = –0.003. Similar results were found when constraining item thresholds (strong model: ΔCFI = –0.001; ΔRMSEA = –0.010) and residuals (strict model: ΔCFI = –0.008; ΔRMSEA = 0.011).

Having established full invariance between youth and adults, sex invariance was tested among the youth sample (Table 5). Baseline models fit the data well for both male youth (CFI = 0.996; RMSEA = 0.047 [0.029 to 0.066]) and female youth (CFI = 0.989; RMSEA [0.059 [0.041 to 0.079]). The configural model (no equality constraints) demonstrated adequate fit to the data (CFI = 0.986; RMSEA = 0.074 [0.061 to 0.086]). Constraints imposed at the weak, strong, and strict levels did not result in substantial worsening of model fit, and full sex invariance was demonstrated. The final strict model had good fit (CFI = 0.982; RMSEA = 0.053 [0.045 to 0.061]), and all parameters were statistically significant for both males and females (Table 6).

K6 and WHO-CIDI Diagnoses

The prevalence of past-year major depressive episode, generalized anxiety disorder, and bipolar disorder is shown in Table 7. The K6 provided good predictions of these disorders with AUCs ranging from 0.847 to 0.853 in the overall sample. Only the AUC for predicting generalized anxiety disorder among male youth was considered fair.[26] In all cases, AUCs were higher among females compared to males, although these differences were not statistically significant. Receiver operating characteristic curves for the full sample, males, and females, are shown in Figures 2-4.

Discussion

Summary of Findings

The K6 was found to be valid and reliable in an epidemiological sample of Canadian youth. The scale demonstrated measurement invariance by age, suggesting that youth interpret items on the K6 similarly to adults—the population for whom the K6 was originally developed—and by

Table 3. Polychoric Correlation Matrix of the K6.

	Nervous	Hopeless	Restless	Depressed	Effort	Worthless
Nervous	—	0.44	0.41	0.53	0.31	0.49
Hopeless	0.54	—	0.40	0.72	0.36	0.76
Restless	0.37	0.46	—	0.40	0.31	0.43
Depressed	0.50	0.74	0.41	—	0.44	0.78
Effort	0.36	0.57	0.37	0.60	—	0.46
Worthless	0.48	0.78	0.44	0.76	0.59	—

Correlations above the diagonal are for males and below the diagonal are for females. All correlations were statistically significant at $P < 0.001$.

Table 4. Measurement Invariance of the K6 between Youth and Adults.

Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	$\Delta\chi^2$ (*df*)	ΔCFI	ΔRMSEA
Baseline						
Youth	58.81 (9)	0.993	0.052 (0.040 to 0.066)	—	—	—
Adult	23.15 (9)	0.999	0.028 (0.014 to 0.042)	—	—	—
Configural	159.48 (19)	0.992	0.061 (0.052 to 0.070)	—	—	—
Weak	181.31 (24)	0.991	0.057 (0.049 to 0.065)	33.47 (5)[a]	–0.001	–0.003
Strong	222.98 (41)	0.990	0.047 (0.041 to 0.053)	53.10 (17)[a]	–0.001	–0.010
Strict	367.47 (47)	0.982	0.058 (0.053 to 0.064)	129.63 (6)[a]	–0.008	0.011

CFI, comparative fit index; CI, confidence interval; RMSEA, root mean square error of approximation; —, change scores for fit indices are not computed for baseline or configural models.
[a]$P < 0.001$.

Table 5. Measurement Invariance of the K6 between Male and Female Youth.

Model	χ^2 (*df*)	CFI	RMSEA (90% CI)	$\Delta\chi^2$ (*df*)	ΔCFI	ΔRMSEA
Baseline						
Male	29.73 (9)	0.996	0.047 (0.029 to 0.066)	—	—	—
Female	39.04 (9)	0.989	0.059 (0.041 to 0.079)	—	—	—
Configural	122.18 (19)	0.986	0.074 (0.061 to 0.086)	—	—	—
Weak	115.09 (24)	0.988	0.061 (0.050 to 0.073)	10.23 (5)	0.002	–0.013
Strong	155.25 (41)	0.985	0.053 (0.044 to 0.062)	44.35 (17)[a]	–0.003	–0.008
Strict	179.87 (47)	0.982	0.053 (0.045 to 0.061)	28.37 (6)[a]	–0.003	0.000

CFI, comparative fit index; CI, confidence interval; RMSEA, root mean square error of approximation; —, Change scores for fit indices are not computed for baseline or configural models.
[a]$P < 0.001$.

Table 6. Parameter Estimates of the Final Invariant K6 Model.

Parameter	Nervous	Hopeless	Restless	Depressed	Effort	Worthless
Males						
Loading	.665	.834	.494	.832	.558	.865
Threshold 1	–2.151	–2.667	–1.983	–2.690	–1.961	–2.631
Threshold 2	–1.350	–1.963	–1.245	–2.160	–1.264	–1.972
Threshold 3	–0.310	–1.270	–0.444	–1.454	–0.525	–1.387
Threshold 4	0.720	–0.574	0.407	–0.863	0.222	–0.754
Residual	.558	.305	.756	.309	.689	.252
Females						
Loading	.707	.862	.538	.860	.603	.888
Threshold 1	–2.036	–2.453	–1.923	–2.475	–1.886	–2.406
Threshold 2	–1.278	–1.805	–1.207	–1.988	–1.216	–1.803
Threshold 3	–0.294	–1.168	–0.430	–1.338	–0.505	–1.268
Threshold 4	0.682	–0.528	0.394	–0.794	0.213	–0.689
Residual	.500	.258	.710	.261	.637	.211

All estimates are standardized and statistically significant at $P < 0.001$.

Table 7. Area under the Curve for the K6 and Past-Year Mental Disorder (WHO-CIDI 3.0).

	Total Sample			Males			Females		
Mental Disorder	*n* (%)	AUC	95% CI	*n* (%)	AUC	95% CI	*n* (%)	AUC	95% CI
Major depression	240 (5.4)	0.848	0.814 to 0.882	71 (3.3)	0.804	0.742 to 0.866	168 (7.5)	0.864	0.823 to 0.905
Generalized anxiety	125 (2.9)	0.847	0.781 to 0.912	37 (1.7)	0.736	0.578 to 0.895	89 (4.0)	0.878	0.815 to 0.941
Bipolar disorder	73 (1.7)	0.853	0.796 to 0.910	32 (1.5)	0.833	0.729 to 0.936	41 (1.8)	0.862	0.796 to 0.928
Any disorder[a]	281 (6.4)	0.849	0.817 to 0.880	95 (4.4)	0.810	0.753 to 0.866	186 (8.3)	0.864	0.827 to 0.902

AUC, area under the curve; CI, confidence interval.
[a]Refers to any of the three disorders listed in the table.

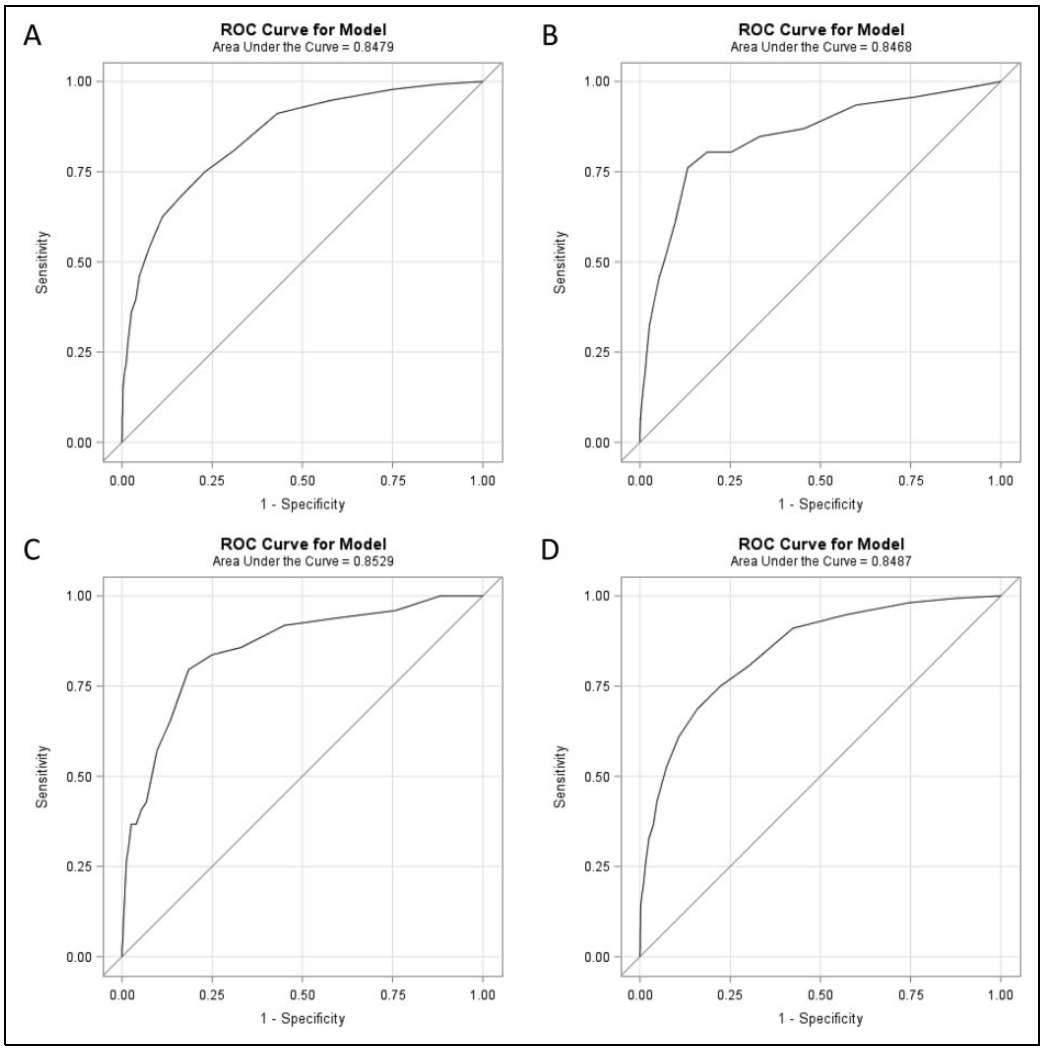

Figure 2. Receiver operating characteristic (ROC) curves for the K6 and past-year mental disorder (full sample). (A) Major depressive episode, (B) generalized anxiety disorder, (C) bipolar disorder, and (D) any mental disorder.

sex. Findings also showed that the K6 performed equally well in predicting *DSM-IV* disorders, as measured by the WHO-CIDI.

The J-shaped distribution of K6 scores was similar to that reported in previous adult[1,2] and youth studies,[4-6] with the majority of our sample having no or low scores. Levels of psychological distress were higher among females compared to males, perhaps related to the fact that the items represent emotion-oriented, as opposed to behaviour-oriented, symptoms. Evidence shows well-defined sex differences with regards to the types of psychiatric disorder experienced.[27,28]

Factor models and reliability analyses confirmed the one-factor model of the K6,[3,5,7] and no modifications to the factor structure were needed to demonstrate full strict

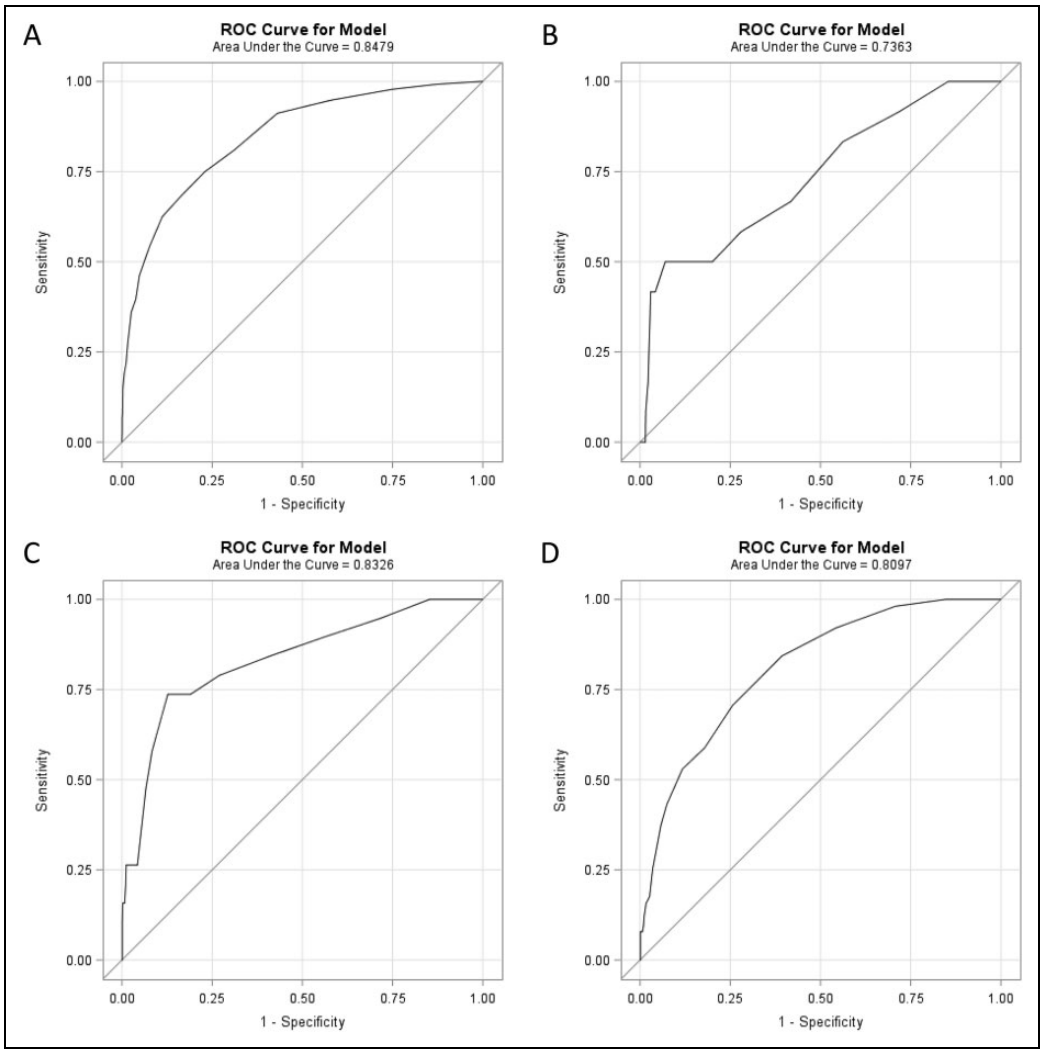

Figure 3. Receiver operating characteristic (ROC) curves for the K6 and past-year mental disorder (males). (A) Major depressive episode, (B) generalized anxiety disorder, (C) bipolar disorder, and (D) any mental disorder.

invariance of the K6 across age and sex. These findings reflect that the items on the K6 were purposively selected to minimize redundancy.[1] The absence of noninvariant parameters instills confidence that K6 item and composite score comparisons across age and sex are valid—observed differences are likely to represent true differences between groups and are not an artifact of differential interpretation of the K6. Findings from our invariance testing are comparable to a recent study[8] but contrast previous research in other adult and youth samples that required modifications to the factor structure of the K6 (i.e., correlated residuals) to achieve adequate model fit,[5,6] as well as the release of constraints on noninvariant parameters at the strong and strict levels to establish partial measurement invariance.[5,29] Previous studies have used RMSEA values indicative of poor fit to justify the inclusion of modifications to the K6.[5,6] In those studies, CFI or Tucker-Lewis index values suggested fit was adequate, and thus the inclusion of correlated residuals prior to invariance testing may have been unnecessary. Evidence is available suggesting that the RMSEA underperforms in models with small degrees of freedom.[30]

The ability of the K6 to predict psychiatric disorder was comparable to previous findings in youth populations with regards to the magnitude of the AUC[3,7] and the finding that prediction was better for females compared to males.[5] It noteworthy that AUCs in the current study were larger than those reported in the NCS-A.[4] Methodological differences may explain this finding. First, the NCS-A included a younger sample of youth (13 to 17 years) compared to the CCHS-MH (15 to 19 years). Second, whereas the CCHS-MH relied on a single informant (i.e., the youth), the NCS-A included both parent- and youth-reported CIDI and recorded symptoms as present, if endorsed by either parent or youth. Our findings further confirm the utility of the K6 in epidemiological studies of youth, particularly in the Canadian context, to estimate the prevalence of psychological distress.

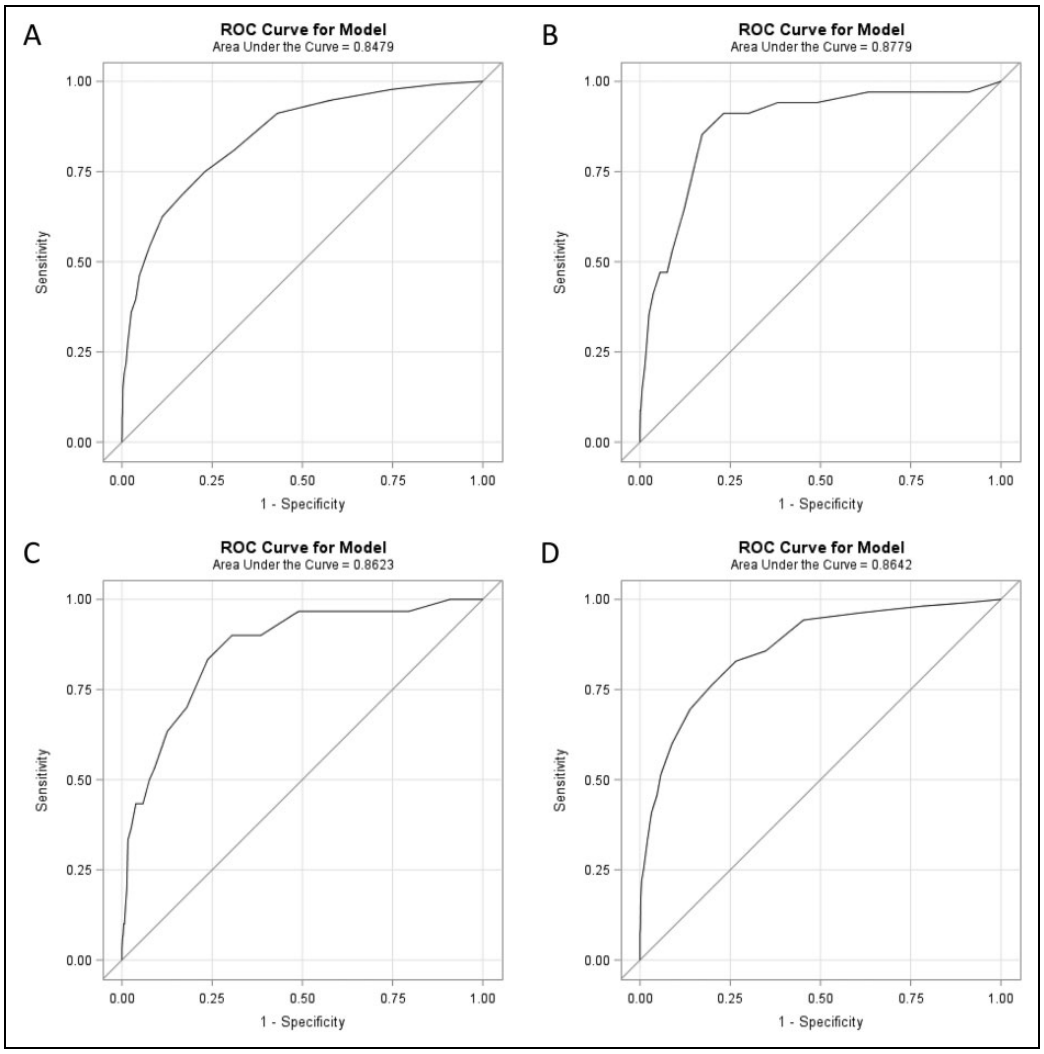

Figure 4. Receiver operating characteristic (ROC) curves for the K6 and past-year mental disorder (females). (A) Major depressive episode, (B) generalized anxiety disorder, (C) bipolar disorder, and (D) any mental disorder.

Additionally, these findings support the broad applicability of the K6 in community, educational, and clinical settings to screen and identify youth with elevated psychological distress who require more comprehensive assessment of their mental health.

Future Research Directions

Research validating the K6 should begin to examine the longitudinal invariance of the scale. While we show that the K6 is invariant between adults and youth, these are independent samples. Tests of longitudinal invariance would inform the extent to which the K6 is invariant in the same individuals over time and, if noninvariance is observed, could help identify developmental periods whence it occurs. Such research is needed to ensure that research investigating trajectories of psychological distress, as measured by the K6, are not biased by differential interpretation of its items or the construct of psychological distress as individuals mature over the life course. Furthermore, researchers should consider opportunities to examine the psychometric properties of the K6 in younger children. Tests of agreement and invariance between children and their parents can extend the utility of the K6 throughout life. Finally, validation of the K6 should extend to other youth populations, particularly in Europe, South America, and Africa, as well as the extent to which the K6 predicts externalizing disorders, where there is currently a paucity of research.

Limitations

Findings from this study should be considered in the context of the following limitations. First, the K6 measured past-month psychological distress, whereas the WHO-CIDI measured past-year psychiatric disorder. As noted in previous studies,[4] this discrepancy in the recall period likely results in an underestimate of the AUC. In a related vein, while the WHO-CIDI has been shown to be valid and reliable, positive

screens do not necessarily correspond to clinical diagnoses. Second, only 3 mental disorders were examined in the CCHS-MH, none of which were externalizing disorders. Thus, it is possible that individuals with other disorders not assessed in the CCHS-MH are included in the "no mental disorder" group. Third, individuals residing in the Canadian territories, Indigenous reserves and settlements, institutions, and full-time members of the Canadian Armed Forces, which comprise approximately 3% of the total Canadian population, were excluded. Thus, findings may not generalize to these groups. Fourth, while childhood typically ends at 18 years of age, the CCHS-MH public use file categorized participants into 5-year bins, with the youngest bin consisting of youth aged 15 to 19 years. Including youth aged 19 years is suitable for a Canadian context given that several provinces legislate this as the age of majority. Furthermore, there is evidence of the extension of adolescence/emerging adulthood into the 20s.[31,32]

Conclusions

Despite these limitations, these findings provide robust evidence validating the K6 among Canadian youth. Researchers and health professionals can be confident that group differences in K6 scores represent true differences. The strong predictive power of the K6 indicates that youth reporting elevated symptoms of psychological distress should be referred for more comprehensive assessment for psychiatric disorder. Strategies for early identification of psychiatric disorder should include screening with the K6 to ensure at-risk youth are provided with the health care and supportive resources needed to ensure the best possible mental health outcomes.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: This work was conducted using data collected by Statistics Canada. However, the analyses presented here were conducted by the author and the interpretations presented in this paper do not reflect the interpretations or opinions of Statistics Canada. Dr. Ferro serves on the editorial board of the *Canadian Journal of Psychiatry*.

Data Access

The data used in this study are available to researchers through application to the Canadian Research Data Centre Network.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study received no funding. Dr. Ferro holds the Canada Research Chair in Youth Mental Health and is the recipient of an Early Researcher Award from the Ontario Ministry of Research, Innovation and Science.

References

1. Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med. 2002;32(6):959-976.
2. Kessler RC, Green JG, Gruber MJ, et al. Screening for serious mental illness in the general population with the K6 screening scale: results from the WHO World Mental Health (WMH) survey initiative. Int J Methods Psychiatr Res. 2010;19(suppl 1):4-22.
3. Chan SM, Fung TCT. Reliability and validity of K10 and K6 in screening depressive symptoms in Hong Kong adolescents. Vulnerable Child Youth Stud. 2014;9:75-85.
4. Green JG, Gruber MJ, Sampson NA, et al. Improving the K6 short scale to predict serious emotional disturbance in adolescents in the USA. Int J Methods Psychiatr Res. 2010; 19(suppl 1):23-35.
5. Mewton L, Kessler RC, Slade T, et al. The psychometric properties of the Kessler Psychological Distress Scale (K6) in a general population sample of adolescents. Psychol Assess. 2016;28(10):1232-1242.
6. Peiper N, Clayton R, Wilson R, et al. The performance of the K6 scale in a large school sample. Psychol Assess. 2015;27(1): 228-238.
7. Kang YK, Guo WJ, Xu H, et al. The 6-item Kessler psychological distress scale to survey serious mental illness among Chinese undergraduates: psychometric properties and prevalence estimate. Compr Psychiatry. 2015;63:105-112.
8. Peiper N, Lee A, Lindsay S, et al. The performance of the K6 scale in a large school sample: a follow-up study evaluating measurement invariance on the Idaho youth prevention survey. Psychol Assess. 2016;28(6):775-779.
9. Ferro MA. Major depressive disorder, suicidal behaviour, bipolar disorder, and generalised anxiety disorder among emerging adults with and without chronic health conditions. Epidemiol Psychiatr Sci. 2016;25(5):462-474.
10. Ferro MA, Gorter JW, Boyle MH. Trajectories of depressive symptoms in Canadian emerging adults. Am J Public Health. 2015;105(11):2322-2327.
11. Kessler RC, Avenevoli S, Costello EJ, et al. Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the national comorbidity survey replication adolescent supplement. Arch Gen Psychiatry. 2012;69(4):372-380.
12. Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2197-2223.
13. Kim-Cohen J, Caspi A, Moffitt TE, et al. Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. Arch Gen Psychiatry. 2003;60(7):709-717.
14. Statistics Canada. Canadian Community Health Survey (CCHS)–Mental Health. User Guide: Microdata Files. Ottawa: Statistics Canada; 2013.
15. Statistics Canada, Canadian Community Health Survey–Mental Health (CCHS). 2013 [cited 2018 Nov 10]. http://www23.

statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5015.
16. Hegeman JM, Kok RM, van der Mast RC, et al. Phenomenology of depression in older compared with younger adults: meta-analysis. Br J Psychiatry. 2012;200(4):275-281.
17. Kessler RC, Ustun TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). Int J Methods Psychiatr Res. 2004;13(2):93-121.
18. Kessler RC, Abelson J, Demler O, et al. Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMHCIDI). Int J Methods Psychiatr Res. 2004;13(2):122-139.
19. Gadermann A, Guhn M, Zumbo BD. An empirical comparison of Cronbach's alpha with ordinal reliability coefficients alpha and theta. Int J Psychol. 2008;43:55.
20. Millsap RE, Yun-Tein J. Assessing factorial invariance in ordered-categorical measures. Multivariate Behav Res. 2004; 39(3):479-515.
21. Byrne BM. Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming. New York: Taylor & Francis; 2012.
22. Bollen KA. Structural Equations with Latent Variables. New York: John Wiley; 1989.
23. Brown TA. Confirmatory Factor Analysis for Applied Research. 2nd ed. New York: Guilford; 2015.
24. Steinmetz H, Schmidt P, Tina-Booh A, et al. Testing measurement invariance using multiple group CFA: differences between educational groups in human values measurement. Qual Quant. 2009;43:599-616.
25. Chen FF. Sensitivity of goodness of fit indices to lack of measurement invariance. Struct Equat Model. 2007;14:464-504.
26. Cicchetti DV. The precision of reliability and validity estimates re-visited: distinguishing between clinical and statistical significance of sample size requirements. J Clin Exp Neuropsychol. 2001;23(5):695-700.
27. Seedat S, Scott KM, Angermeyer MC, et al. Cross-national associations between gender and mental disorders in the World Health Organization world mental health surveys. Arch Gen Psychiat. 2009;66(7):785-795.
28. Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol. 2011;21(9):655-679.
29. Drapeau A, Beaulieu-Prevost D, Marchand A, et al. A life-course and time perspective on the construct validity of psychological distress in women and men: measurement invariance of the K6 across gender. BMC Med Res Methodol. 2010; 10:68.
30. Kenny DA, Kaniskan B, McCoach DB. The performance of RMSEA in models with small degrees of freedom. Sociol Methods Res. 2015;44:486-507.
31. Arnett JJ. Emerging adulthood: a theory of development from the late teens through the twenties. Am Psychol. 2000;55(5): 469-480.
32. Qadeer RA, Georgiades K, Boyle MH, et al. An epidemiological study of substance use disorders among emerging and young adults. Can J Psychiatry. 2018. doi: 10.1177/0706743718792189.