Original quantitative research

Validating existing clinical cut-points for the parent-reported Strengths and Difficulties Questionnaire in a large sample of Canadian children and youth

Sarah E. Turner, MSc (1,2); Raelyne L. Dopko, PhD (1); Gary Goldfield, PhD (3); Paula Cloutier, MA (3,4); Kathleen Pajer, MD (3,5); Mohcene Abdessemed, MSc (3); Fatima Mougharbel, MSc (6); Michael Ranney, BSc (4); Matt D. Hoffmann, PhD (7); Justin J. Lang, PhD (1,8,9,10)

This article has been peer reviewed.

✕ Tweet this article

Abstract

Introduction: The Strengths and Difficulties Questionnaire (SDQ), for assessing behavioural and emotional difficulties, has been used internationally as a screening measure for mental health problems. Our objective was to validate the existing (British) SDQ cut-points in a sample of Canadian children and youth, and develop new Canadian SDQ cut-points if needed.

Methods: This study includes data from children and youth aged 6 to 17 years from the Canadian Health Measures Survey (n = 3435) and outpatient records from the Children's Hospital of Eastern Ontario (n = 1075). The parent-reported SDQ data were collected. We adjusted the existing SDQ cut-points using a distributional and receiver-operating characteristic (ROC) curve approach. We subsequently calculated the sensitivity, specificity and diagnostic odds ratio of the existing and new SDQ clinical cut-points to determine whether the new cut-points had better clinical utility, using both analytic approaches.

Results: Our data show differences in the screening effectiveness between the existing British and the Canadian-specific clinical cut-points. Specificity is maximized using the Canadian distributional cut-points, improving the likelihood of identifying true negative results. The total SDQ score met the threshold for clinical utility (diagnostic odds ratio > 20) using both the existing and new cut-points; however, the individual scales did not reach clinical utility threshold using either cut-points.

Conclusion: Future Canadian SDQ research should consider the new cut-points derived from our study population and the existing British cut-points to allow for historical and international comparisons.

Keywords: child and adolescent, mental health, validation, hyperactivity, peer problems, prosocial behaviour, conduct problems, emotional symptoms

Highlights

- This study validated the existing British SDQ cut-points in a large sample of Canadian children and youth and developed Canadian-specific cut-points using a distributional approach and receiver operating characteristic (ROC) curves.
- The Canadian-specific clinical cut-points (90th percentile) using the distributional approach demonstrated higher specificity than the ROC curve derived cut-points. For this reason, the distributional cut-points have better population-based utility.
- Both the existing British and the Canadian-specific clinical cut-points for the total difficulties score met the threshold for clinical utility to predict mental health diagnosis.

Introduction

In Canada, approximately 1.2 million children and youth are affected by mental illness, and a high percentage of children and youth are symptomatic, but do not meet full diagnostic criteria (i.e. they are symptomatic at a subclinical threshold).[1]

Author references:

1. Centre for Surveillance and Applied Research, Public Health Agency of Canada, Ottawa, Ontario, Canada
2. Department of Community Health Sciences, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada
3. CHEO Research Institute, Ottawa, Ontario, Canada
4. Mental Health Department, CHEO, Ottawa, Ontario, Canada
5. Department of Psychiatry, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
6. Interdisciplinary School of Health Sciences, Faculty of Health Sciences, University of Ottawa, Ottawa, Ontario, Canada
7. Department of Kinesiology, California State University, Fullerton, Fullerton, California, USA
8. School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
9. Alliance for Research in Exercise, Nutrition and Activity (ARENA), University of South Australia, Adelaide, South Australia, Australia
10. Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

Correspondence: Justin J. Lang, Centre for Surveillance and Applied Research, Public Health Agency of Canada, Ottawa, ON K1A 0K9; Tel: 613-293-9008; Email: justin.lang@phac-aspc.gc.ca

Compared to 2019, the proportion of children and youth aged 5 to 24 years hospitalized for mental health conditions rose by 2% in 2020, and nearly 1 in 4 of all hospitalizations in this age group were due to mental health problems.[2] Furthermore, up to 70% of adult mental health problems begin in childhood, highlighting the need to identify and treat mental health vulnerabilities in early life.[1] The increasing rates of mental health problems among Canadians necessitates access to screening in populations and clinical settings with psychometrically sound mental health measurement tools.

The Strengths and Difficulties Questionnaire (SDQ) is a widely used measure of children and youth social, emotional and behavioural difficulties.[3] In clinical settings and epidemiological studies, the clinical cut-points for the SDQ are used as a baseline screening tool for mental health problems.[4,5] The SDQ comprises five scales that measure conduct problems, emotional symptoms, hyperactivity, peer problems and prosocial behaviour, as well as a total difficulties score that sums the scores from all these scales except prosocial behaviour. Each scale has established cut-points for borderline and clinical SDQ scores that were originally identified in a small sample of 403 British children and youth aged 4 to 16 years. The cut-points were chosen so that roughly 80% of the scores from the community were considered "normal," 10% were "borderline" and 10% were "clinical."[3]

The original 1997 British SDQ cut-points[3] have been widely used in Canada and internationally. They have been compared to population data in high-income countries, including the United States (US), Japan and Germany.[6-8] In the US, the established cut-points were similar to the existing British values with two exceptions: the total difficulties score was 1 to 2 points lower than the British values for the normal, borderline and clinical cut-points; and the prosocial score was 1 point lower in the borderline and clinical categories than the British cut-points.[7] In Japan, the cut-points for the total difficulties score were 2 to 3 points lower across the categories (normal, borderline and clinical) for both boys and girls aged 10 to 15 years compared to the existing British cut-points;[6] however, the existing cut-points correctly classified Japanese boys aged 7 to 9 years. Finally, using German normative data for boys and girls aged 6

to 16 years, the cut-points for the total difficulties score were also 1 point lower than the existing British values.[8] None of the studies recommended changing the existing cut-points. These varying country-specific results highlight the need to investigate the validity of the 1997 British SDQ cut-points among a sample of Canadian children and youth.

Since 2007, the parent-reported SDQ has been collected as part of the Canadian Health Measures Survey (CHMS), a national survey of health and well-being. In 2020– the five-factor SDQ demonstrated sound psychometric properties using data from approximately 7500 children and youth who participated in the CHMS.[9] The five-factor SDQ (i.e. conduct problems, emotional symptoms, hyperactivity, peer problems and prosocial behaviour) showed good fit with the data using confirmatory factor analysis and was invariant across sex (male, female) and age (children, youth). However, the clinical cut-points have not been validated in a Canadian population.

The overall objective for this study was to validate the British SDQ cut-points in a large sample of Canadian children and youth. To attain this objective, we completed this study in two phases. The aim of the first phase was to determine if the British cut-points for the SDQ appropriately classified a national sample of Canadian children and youth aged 6 to 17 years. We also examined the cut-points using receiver operating characteristic (ROC) curves as a data-driven approach to identify clinical cut-points. Adjustments were made to the British SDQ cut-points when needed to create new Canadian cut-points. In the second phase, we compared the differences in screening effectiveness (i.e. sensitivity and specificity) using the British SDQ cut-points and the new Canadian cut-points.

Methods

General population sample

This study utilized data from children and youth aged 6 to 17 years from cycles 3 (2012–2013) and 4 (2014–2015) of the CHMS household questionnaire. The CHMS is a cross-sectional, nationally representative survey of Canadians living in the 10 provinces. The CHMS does not collect information from individuals living in the three territories or on reserves, full-time members of the Canadian Armed Forces or

those living in institutions (exclusions represented approximately 4% of the population).[10]

Ethics approval for data collection was obtained by Statistics Canada from the Health Canada and the Public Health Agency of Canada (PHAC) Research Ethics Board. Participation in the CHMS is voluntary.[11] Written informed consent prior to participation was obtained from the parent or guardians on behalf of the children aged 6 to 13 years. Assent from children aged 6 to 13 was also obtained.[11] Youth aged 14 to 17 years provided informed consent to participate. Further details about the CHMS are available elsewhere.[12]

In total, 3435 participants took part in this study—1720 individuals from cycle 3 (49.8% female) and 1715 individuals from cycle 4 (49.6% female). The SDQ was completed during the household interview by parents or guardians of children and youth aged 6 to 17 years (i.e. parent-reported SDQ). Slightly more than half (59.8%) of the sample were 6 to 11 years old, and the remaining 40.2% were 12 to 17 years old.

Clinical sample

We obtained clinical data of the children and youth aged 6 to 17 years who presented to an outpatient mental health clinic that was part of the Children's Hospital of Eastern Ontario (CHEO) between 25 January 2016 and 16 March 2020 (n = 1075). The majority of our sample was from the province of Ontario, with only a few out-of-province patients. The SDQ was completed as part of the mandatory baseline clinical assessment by the parent or caregiver during the first clinical appointment (i.e. parent-reported SDQ), and as a result, the response rate was greater than 85%.

Mental health diagnoses were made by a trained psychologist using ICD-10-CA codes.[13] Diagnoses recorded in the patient chart during the first clinical visit were used to classify children and youth into one or more diagnostic categories: mood disorders (ICD-10-CA:F30-39.*), anxiety disorders (ICD-10-CA:F40-49.*), pervasive developmental disorders (ICD-10-CA:80-89.*), conduct disorder (ICD-10-CA:F91.*) and attention deficit hyperactivity disorder (ADHD) (ICD-10-CA:F90.*). Patients with more than one diagnosis were retained in

the sample, and their data were used in multiple categories.

Ethics approval

Ethics approval for using the CHEO clinical sample was obtained from the CHEO Research Ethics Board (21/97X) and the PHAC Research Ethics Board (2021-032P). Written informed consent was obtained from parents or legal guardians, and assent was obtained from each child for their data to be used for research purposes. A formal data-sharing agreement was implemented between the CHEO Research Institute and PHAC to send clinical data to PHAC for this study. Clinical data from this study will be kept on a secure PHAC server for 7 years before being destroyed.

Measures

Strengths and Difficulties Questionnaire
The SDQ is a 25-item questionnaire designed to measure problematic behaviours, emotions and relationships.[3] It has demonstrated evidence of validity and reliability with Canadian children and youth.[9] All items are scored on a three-point Likert scale with the following response options: 0 ("not true"), 1 ("somewhat true") and 2 ("certainly true"). Higher scores indicate a greater difficulty for all scales except the prosocial behaviour scale, in which lower scores indicate greater difficulty. The conduct problems, emotional symptoms, hyperactivity and peer problems scores were summed to create a total difficulties score (and hence a scale).

Goodman (1997)[3] established score cut-points for normal, borderline and clinical mental health difficulties based on a sample of children and youth from London, England, United Kingdom.

Demographic variables
We summarized the characteristics of each sample with descriptive statistics. Demographic data in the general population sample comprised biological sex (male/female), age (years), highest level of household education (less than high school / high school or college / university), household income (less than $40 000 / $40 000 to $79 999 / $80 000 or more), self-perceived general health (poor or fair / good / very good / excellent) and self-perceived mental health (poor or fair / good / very good / excellent).

Self-perceived general and mental health were only available for youth aged 12 to 17 years. Age was the only demographic characteristic available for the clinical sample.

Statistical analyses
For the general population sample, we calculated descriptive statistics stratified by sex (male and female). We conducted sensitivity analyses to determine if SDQ scores, stratified by sex and age group, changed between cycle 3 and cycle 4 of the CHMS (data available on request from the authors). Few differences between groups justified combining data from cycles 3 and 4. We also combined SDQ data for all age and sex groups, in line with the approach originally conducted by Goodman.[3] For the full clinical sample, we calculated mean SDQ scores for each scale and the prevalence of each mental health diagnosis.

Analyses were conducted using SAS Enterprise Guide 7.1 (SAS Institute Inc., Cary, NC, US).

Phase 1: Establishing cut-points

Distributional technique
First, using the general population sample, we calculated the percentage with 95% confidence intervals (CI) of children and youth in the CHMS with scores that fell within the existing cut-points for the normal, borderline and clinical categories for each SDQ scale. In cases where the general population did not align with the 80%, 10%, 10% framework, we selected new Canadian cut-points based on visual inspection of density plots and manually adjusting cut-points to determine the best alignment to the framework. In cases where the percentages were either slightly below or above the target percentage, we chose scores below the target percentage while prioritising accuracy in the clinical group, following the example of Bourdon et al.[7] All distributional analyses used bootstrap and survey weights provided by Statistics Canada to generate nationally representative estimates.

ROC curve technique
We calculated ROC curves using the PROC LOGISTIC command in SAS Enterprise Guide with the "outroc" function. We used the SAS ROCPLOT macro (https://support.sas.com/kb/25/018.html) to calculate the sensitivity and specificity for each possible cut-point. We selected the

cut-point that maximized both sensitivity and specificity, otherwise known as Youden Index. The ROC curve analyses were not considered representative of the Canadian population because they were calculated using unweighted data.

Phase 2: Comparison of the existing British and the Canadian-specific cut-points

We used sensitivity and specificity calculations to validate the existing British and the Canadian-specific clinical (90th percentile) cut-points for each SDQ score. In previous studies, the 90th percentile cut-points were associated with 12 times higher odds of service use for a mental health difficulty[7] and 15 times higher odds of a diagnosed mental health disorder.[14]

Sensitivity (or true positive) is the proportion of the sample that is correctly identified as having a mental health diagnosis (CHEO clinical sample). Specificity (or true negative) is the proportion of the sample that is correctly identified as having no mental health diagnosis (general population sample). Scores above 0.5 indicate that the measure is better than chance at discriminating those with the outcome of interest.

Phase 2: Additional sensitivity analyses

We conducted additional sensitivity analyses to determine the accuracy of our results. First, we limited both datasets to those aged 12 to 17 years, and we retained those in the general population sample who had self-reported their mental health as being very good or excellent[15] (general population sample, n = 1021; clinical sample, n = 790). Limiting to very good or excellent mental health in the general population sample provides us with a more distinct positive mental health group to use for comparison.

Next, using the full age range (i.e. 6 to 17 years), we used the specific mental health diagnosis in the clinical sample to determine how well the existing British and the Canadian-specific cut-points discriminated between those with mood disorders, anxiety disorders, conduct disorder, ADHD or pervasive developmental disorders. For these specific mental health diagnoses, we also calculated the positive predictive value (PPV; the probability of having a mental health problem if meeting the clinical cut-point) and the negative

predictive value (NPV; the probability of not having a mental health problem if not meeting the clinical cut-point). In community-based screening, a test with high specificity or NPV will reduce the number of false positives and allow for monitoring or treatment to begin early if a positive result is detected.[16]

Finally, we calculated the diagnostic odds ratio (DOR) for candidate SDQ scales that aligned with specific clinical diagnoses. For instance, the conduct scale aligned with conduct disorders, the hyperactivity scale aligned with ADHD, the emotional symptoms scale aligned with mood or anxiety disorders, and the peer problems scale aligned with pervasive developmental disorder. Vugteveen et al.[5] have described using these candidate SDQ scales and their matched diagnoses. The DOR is a single measure that incorporates both sensitivity and specificity and is relatively independent of changes in the prevalence of mental health diagnoses.[17] DOR values greater than 20 identify a test that is potentially useful for influencing clinical decision-making.[18]

Results

Characteristics of general population sample

Descriptive statistics for the population sample, stratified by sex, are provided in Table 1. Based on overlapping 95% CIs, there were no differences between sex for any of the demographic variables. There were sex differences for three of the five SDQ scales. Females had higher (worse) emotional symptoms scores compared to males (2.30 [95% CI: 2.15–2.45] vs. 1.78 [95% CI: 1.63–1.93]) and higher (better) prosocial behaviour scores compared to males (9.25 [95% CI: 9.18–9.33] vs. 8.86 [95% CI: 8.74–8.98]). Males had higher hyperactivity scores compared to females (3.18 [95% CI: 2.96–3.39] vs. 2.48 [95% CI: 2.34–2.63]). Despite these differences, we decided not to stratify further analyses by sex to be consistent with the original SDQ development and other country-specific cut-points (as described in previous literature[3,7,8]).

Characteristics of the clinical sample

Descriptive statistics for the clinical sample are also provided in Table 1. Sex or gender was not obtained to maintain the sample anonymity, and so descriptive statistics are presented for the total population.

For all SDQ scales the clinical samples scored significantly worse than the general population sample. The mean scores for the emotional symptoms, peer problems and the total difficulties scales were in the clinical range according to the British cut-points. The most prevalent mental health diagnoses were anxiety disorders (59.1%), followed by ADHD (37.2%). Nearly half of the clinical sample presented with more than one diagnosis (43.9%).

Phase 1: Establishing cut-points

Table 2 shows the proportion of the general population sample that fell within the existing SDQ cut-points for each scale (i.e. the distributional technique). The existing cut-points for conduct problems and peer problems accurately classified the sample of Canadian children and youth into the borderline (80th percentile) and clinical (90th percentile) categories. For emotional symptoms and hyperactivity, the sample of Canadian children and youth were over-represented in the clinical category (13.9% and 13.0%, respectively). The existing cut-points for prosocial and total difficulties under-represented children and youth in the clinical category (1.5% and 8.1%, respectively).

We therefore created new Canadian cut-points for the three scales (emotional symptoms, hyperactivity and prosocial) and the total difficulties score that more accurately classified the sample of Canadian children and youth into the borderline and clinical categories (Table 3). The Canadian-specific clinical cut-points resulted in a range of 6 to 10 for emotional symptoms, 8 to 10 for hyperactivity, 0 to 6 for prosocial behaviour and 16 to 40 for total difficulties, differing by 1 to 2 points from the existing British clinical cut-points.

We also calculated clinical cut-points using ROC curves (Table 4). Apart from the prosocial behaviour scale, which was higher, the cut-points identified using this approach were consistently lower than those identified using the distributional approach. The cut-points identified using ROC curves had improved sensitivity and reduced specificity across all scales compared to the distributional approach.

Phase 2: Comparison of the existing British and the Canadian-specific cut-points

We calculated the sensitivity and specificity for each SDQ scale to determine if the Canadian-specific clinical cut-points (distributional and ROC curve), compared to the existing clinical cut-points, performed better at discriminating between the clinical and general population samples of children and youth (Table 4). The sensitivity for the existing British cut-points ranged from 0.17 to 0.74, the Canadian distributional cut-points ranged from 0.44 to 0.79, and the Canadian ROC curve cut-points ranged from 0.67 to 0.83. The specificity for the existing British cut-points ranged from 0.87 to 0.99, the Canadian distributional cut-points ranged from 0.90 to 0.94, and the ROC curve cut-points ranged from 0.75 to 0.87. We observed significant differences (calculated by non-overlapping 95% CIs) in the sensitivity and specificity scores between the existing and new distributional cut-points for emotional symptoms, hyperactivity and prosocial behaviour, but not for total difficulties. For the emotional symptoms and hyperactivity scores, sensitivity decreased and specificity increased for the Canadian cut-points. For the prosocial scale, the sensitivity increased while the specificity decreased for the Canadian compared to the existing cut-points. All the ROC curve Canadian cut-points were significantly different from the existing British cut-points.

Phase 2: Additional sensitivity analyses

After limiting the sample to those aged 12 to 17 years old and those reporting very good or excellent mental health in the general sample, the sensitivity and specificity results follow the same trend as the full dataset, but with slightly improved sensitivity and specificity values (data available on request from the authors).

Tables 5 and 6 show the screening effectiveness for the existing British and the Canadian-specific clinical cut-points for identifying those with mental health diagnoses in the clinical sample. The specificity was not reported because it is the same as the values reported in Table 4. There were limited differences in the ability of the existing or new clinical distributional cut-points to discriminate between individual mental health diagnoses. The ROC curve cut-points demonstrated significantly

TABLE 1
Descriptive statistics of the population and clinical samples

Variable	CHMS general population sample			CHEO clinical sample
	Males (n = 1729)	Females (n = 1706)	p value	(n = 1075)
	% or mean (95% CI)	% or mean (95% CI)		% or mean (95% CI)
Mean age, years	11.2 (11.1–11.3)	11.3 (11.1–11.4)	0.566	12.4 (12.2–12.5)
Parent education				
< High school	1.9 (0.7–3.2)[E]	3.63 (1.9–5.4)[E]		–
High school or college	13.0 (10.1–15.9)	9.64 (7.1–12.2)	0.050	–
University	85.1 (81.9–88.2)	86.7 (83.0–90.5)		–
Household income, $				
0–39 999	19.8 (15.1–24.5)	17.8 (14.3–21.3)		–
40 000–79 999	28.6 (22.5–34.8)	29.7 (24.0–35.4)	0.735	–
≥80 000	51.6 (45.9–57.2)	52.5 (45.7–59.4)		–
Self-perceived general health[a]				
Poor/fair	4.2 (2.6–5.8)[E]	4.8 (3.0–6.5)[E]		–
Good	21.1 (17.4–24.8)	18.9 (15.7–22.1)		–
Very good	38.8 (34.2–43.5)	37.5 (32.9–42.2)	0.478	–
Excellent	35.9 (31.8–39.9)	38.8 (35.1–42.5)		–
Self-perceived mental health[a]				
Poor/fair	6.0 (2.7–9.4)[E]	7.8 (4.4–11.2)[E]		–
Good	18.3 (13.9–22.7)	19.1 (13.0–25.2)		–
Very good	40.7 (36.1–45.2)	35.9 (30.6–41.2)	0.237	–
Excellent	35.0 (31.5–38.6)	37.2 (32.7–41.7)		–
SDQ scores (mean)				
Conduct problems	1.1 (1.0–1.2)	1.0 (0.9–1.1)	0.052	4.0 (3.8–4.1)
Emotional symptoms	1.8 (1.6–1.9)[E]	2.3 (2.2–2.5)	< 0.001	6.1 (5.9–6.2)
Hyperactivity	3.2 (3.0–3.4)	2.5 (2.3–2.6)	< 0.001	6.5 (6.3–6.6)
Peer problems	1.2 (1.1–1.4)	1.1 (1.0–1.3)	0.067	4.2 (4.1–4.4)
Prosocial behaviour	8.9 (8.7–9.0)	9.3 (9.2–9.3)	< 0.001	6.7 (6.6–6.9)
Total difficulties	7.3 (7.0–7.8)	7.0 (6.6–7.4)	0.064	20.7 (20.3–21.1)
Mental health diagnosis[b]				
Mood disorders	–	–		22.7 (–)
Anxiety disorders	–	–		59.1 (–)
Conduct disorder	–	–		14.7 (–)
ADHD	–	–	–	37.2 (–)
Pervasive developmental disorders	–	–		18.9 (–)
Comorbidity[c]	–	–		43.9 (–)

Abbreviations: ADHD, attention deficit hyperactivity disorder; CHEO, Children's Hospital of Eastern Ontario; CHMS, Canadian Health Measures Survey; CI, confidence interval; –, data not available.

[a] Self-perceived mental and general health are reported for youth aged 12–17 years only.

[b] Percentages include those with multiple diagnoses and, as a result, do not sum to 100%.

[c] % of children and youth diagnosed with more than one of the five listed mental disorders.

[E] Interpret with caution due to high sampling variability.

improved sensitivity across all mental health diagnosis groups. This translated into significantly lower positive predictive values across nearly all scales for all five mental health diagnosis groups, with only small, but significant improvements in the negative predictive values.

To determine the clinical utility of candidate SDQ scales for predicting mental disorders, we calculated the DOR for the existing British and the Canadian-specific distributional and ROC curve clinical cutpoints (data available on request from the authors). None of the candidate SDQ scales were useful for predicting their matched mental health diagnosis as determined by DOR of less than 20; however, the total difficulties score had clinical utility for predicting any of the five mental health diagnoses using either the existing British, new Canadian distributional or ROC curve

SDQ scale	Normal (80%)		Borderline (10%)		Clinical (10%)	
	British score[a]	CHMS % (95% CI)	British score[a]	CHMS % (95% CI)	British score[a]	CHMS % (95% CI)
Conduct problems	0–2	86.6 (84.6–88.7)	3	6.3 (4.6–8.0)	4–10	7.1 (5.5–8.6)
Emotional symptoms	0–3	77.9 (74.8–80.9)	4	8.2 (6.7–9.8)	5–10	13.9 (11.4–16.4)
Hyperactivity	0–5	82.4 (79.9–84.9)	6	4.6 (3.6–5.5)	7–10	13.0 (11.0–15.0)
Peer problems	0–2	83.6 (81.1–86.1)	3	6.9 (5.8–8.0)	4–10	9.5 (7.2–11.9)
Prosocial behaviour	6–10	96.8 (95.7–98.0)	5	1.7[E] (0.9–2.5)	0–4	1.5[E] (0.7–2.3)
Total difficulties	0–13	85.8 (83.5–88.1)	14–16	6.2 (5.0–7.3)	17–40	8.1 (6.1–10.0)

Abbreviations: CHMS, Canadian Health Measures Survey; CI, confidence interval; SDQ, Strengths and Difficulties Questionnaire.

[a] Existing British cut-points are from Goodman.[3]

[E] Interpret with caution due to high sampling variability.

clinical cut-points. The DOR for the total difficulties score for the British clinical cut-point ranged from 31.1 (95% CI: 20.5–50.0) to 46.0 (95% CI: 27.3–81.6); the DOR for the Canadian distributional clinical cut-point ranged from 31.9 (95% CI: 23.0–43.1) to 43.9 (95% CI: 25.6–74.2); and the ROC curve cut-points ranged from 34.8 (95% CI: 25.8–44.8) to 45.1 (95% CI: 27.7–77.0).

Discussion

In this study using a large sample of Canadian children and youth, we derived Canadian-specific distributional cut-points for three of the five SDQ scales and the total difficulties score. We also calculated new cut-points for each of the SDQ scales using a ROC curve technique; to the authors' knowledge, this is the first time this technique has been applied to SDQ data. We then tested the screening effectiveness by comparing the new cut-offs with the British cut-offs in a Canadian clinical sample. Our data demonstrated small differences in screening effectiveness

between the existing British and the Canadian-specific distributional clinical cut-points. Large differences were identified when using the ROC curve technique, which contributed to substantially reduced positive predictive values. When using the SDQ cut-points to screen for five different mental health diagnoses, we found that neither the existing British, nor the new Canadian-specific distributional or ROC curve clinical cut-points for the individual SDQ scales had a DOR above 20. This suggests that the individual SDQ scales may not be useful for screening those with mental health diagnoses. The total difficulties score was useful for predicting mental health diagnoses, indicated by DORs higher than 20, with no significant differences between the existing British and the Canadian-specific distributional or ROC curve clinical cut-points.

Phase 1: Establishing cut-points

The existing British SDQ cut-points did not accurately classify the sample of Canadian children and youth using the

distributional technique for the emotional symptoms, hyperactivity and prosocial behaviour scales and the total difficulties score. This general finding was consistent with other studies that used country-specific data.[6-8] Our results align with data from Germany and the US, which found that a cut-point of 16 or higher to be more accurate in identifying the 90th percentile of children and youth in the total difficulties score, compared to the existing cut-point of 17 or higher.[7,8] However, our results diverge slightly from the German and US data for the prosocial scale, which ranged from 0 to 4 and 0 to 5, respectively, for identifying the 90th percentile, compared to 0 to 6 in our study.

Comparison of the existing British and the Canadian-specific SDQ cut-points

The Canadian distributional cut-points provide a slightly better ability to rule out false positives (improved specificity) than the existing British cut-points for the emotional symptoms and hyperactivity subscales.

SDQ scale	Normal (80%)		Borderline (10%)		Clinical (10%)	
	CHMS score	% (95% CI)	CHMS score	% (95% CI)	CHMS score	% (95% CI)
Conduct problems	0–2	86.6 (84.6–88.7)	3	6.3 (4.6–8.0)	4–10	7.1 (5.5–8.6)
Emotional symptoms	0–3	77.9 (74.8–80.9)	4–5	13.8 (11.8–15.8)	6–10	8.3 (6.4–10.2)
Hyperactivity	0–5	82.4 (79.9–84.9)	6–7	8.8 (7.1–10.5)	8–10	8.8 (6.9–10.7)
Peer problems	0–2	83.6 (81.1–86.1)	3	6.9 (5.8–8.0)	4–10	9.5 (7.2–11.9)
Prosocial behaviour	8–10	86.7 (84.5–88.9)	7	6.8 (5.3–8.3)	0–6	6.5 (4.7–8.3)
Total difficulties	0–11	79.7 (76.7–82.7)	12–15	10.6 (8.8–12.5)	16–40	9.7 (7.6–11.8)

Abbreviations: CHMS, Canadian Health Measures Survey; CI, confidence interval; SDQ, Strengths and Difficulties Questionnaire.

SDQ scale	SDQ clinical cut-point	Sensitivity (95% CI)	Specificity (95% CI)
Conduct problems			
British/Canadian (Distributional)	≥ 4	0.54 (0.51–0.57)	0.94 (0.93–0.94)
Canadian (ROC curve)[b]	≥ 3	0.67 (0.64–0.70)[c]	0.87 (0.86–0.88)[c]
Emotional symptoms			
British	≥ 5	0.74 (0.71–0.76)	0.87 (0.85–0.88)
Canadian (Distributional)	≥ 6	0.61 (0.58–0.64)[c]	0.92 (0.91–0.93)[c]
Canadian (ROC curve)[d]	≥ 4	0.82 (0.80–0.85)[c]	0.78 (0.77–0.79)[c]
Hyperactivity			
British	≥ 7	0.54 (0.51–0.57)	0.87 (0.86–0.88)
Canadian (Distributional)	≥ 8	0.42 (0.39–0.45)[c]	0.91 (0.90–0.92)[c]
Canadian (ROC curve)[e]	≥ 5	0.76 (0.73–0.78)[c]	0.75 (0.73–0.76)[c]
Peer problems			
British/Canadian (Distributional)	≥ 4	0.60 (0.57–0.63)	0.91 (0.90–0.92)
Canadian (ROC curve)[f]	≥ 3	0.74 (0.71–0.76)[c]	0.84 (0.83–0.85)[c]
Prosocial behaviour			
British	≤ 4	0.17 (0.14–0.19)	0.99 (0.98–0.99)
Canadian (Distributional)	≤ 6	0.44 (0.41–0.47)[c]	0.94 (0.93–0.95)[c]
Canadian (ROC curve)[g]	≤ 8	0.73 (0.70–0.75)[c]	0.76 (0.74–0.77)[c]
Total difficulties			
British	≥ 17	0.74 (0.71–0.77)	0.92 (0.91–0.93)
Canadian (Distributional)	≥ 16	0.79 (0.76–0.81)	0.90 (0.89–0.91)
Canadian (ROC curve)[h]	≥ 14	0.86 (0.84–0.88)[c]	0.86 (0.85–0.87)[c]

Abbreviations: CI, confidence interval; ROC, receiver operating characteristic; SDQ, Strengths and Difficulties Questionnaire.

[a] Outcome is clinical sample compared to general population sample. General sample n = 3435; clinical sample n = 1091.

[b] Area under the curve = 0.827.

[c] Significantly different from the British cut-points. Sensitivity is the proportion of true positives identified by the cut-point. Specificity is the proportion of true negatives identified by the cut-point.

[d] Area under the curve = 0.877.

[e] Area under the curve = 0.815.

[f] Area under the curve = 0.853.

[g] Area under the curve = 0.814.

[h] Area under the curve = 0.928.

However, the sensitivity for both scales were reduced. Compared to a previous study using clinical data from a Dutch sample and the existing British cut-offs for each score, both the existing and Canadian distributional cut-points in the current study had better specificity and slightly poorer sensitivity for combinations of candidate SDQ scales and mental health diagnoses.[5] The Canadian ROC curve cut-points demonstrated reduced specificity for all scales, and a substantially lower positive predictive value across all five mental health diagnoses groups. Strong specificity reduces the risk of misclassifying children

and youth not at risk for mental health problems and allows those who test positive to go on for further assessment and treatment. For this reason, we believe the cut-points identified using the distributional technique provides better population-based utility.

Similar to previous work, the DORs for the combinations of candidate SDQ scales and mental health diagnoses did not reach the threshold for clinical utility (> 20).[7] However, the DOR for mood disorder, anxiety disorder, pervasive developmental disorder and conduct disorder all have

95% CIs that cross 20, indicating that the reported DOR is not significantly different from the greater-than-20 threshold. In addition, our results perform better than the results reported by Vugteveen et al.,[5] who found DORs between 3.82 and 5.79 for the same candidate SDQ scales and mental health diagnosis combinations. The predictive ability of the SDQ could be improved by including multiple informants instead of only the parent-reported SDQ scores included in our study. A previous study with a community sample showed better sensitivity when using a combination of parent, teacher and self-report SDQ scores compared to only the parent-reported SDQ scores.[19]

Strengths and limitations

This is the first study to investigate the effectiveness of the existing British SDQ cut-points in a large sample of Canadian children and youth to determine if they appropriately categorized the population into normal, borderline and clinical SDQ categories. This study also applied ROC curves to identify new SDQ cut-points, a novel approach in this area, based on the literature.

The use of a large, population-based sample allows for greater generalizability to the population of children and youth in Canada compared to using a small or convenience sample. Our study is also strengthened by combining a general population sample with a large clinical sample of diagnosed mental health conditions to validate the cut-points. We also compared two validated methods of quantifying cut-points, making the internal validity of our results more robust.

This study also has limitations. First, the original cut-points from Goodman[3] were developed using a sample of children and youth aged 4 to 16 years, while we used a sample aged 6 to 17 years. While these age ranges only differ slightly, they may account for some of the prevalence differences observed between the cut-points.

Second, we used data collected from the general population aged between 7 and 10 years old at the time of our data analysis and when creating the Canadian cut-points. It is possible that the prevalence of clinical-level symptoms on the SDQ scales has increased over the past 10 years. The existing British clinical SDQ cut-points for the emotional symptoms and hyperactivity

TABLE 5

Screening efficiency for existing and Canadian SDQ clinical cut-points from the distributional and ROC curve techniques: mood, anxiety and pervasive developmental disorders in the clinical population compared to the general population of children and youth[a]

SDQ scale	SDQ clinical cut-point	Sensitivity (95% CI)	PPV (95% CI)	NPV (95% CI)
Mood disorder				
Conduct problems				
British/Canadian (Distributional)	≥ 4	0.61 (0.54–0.67)	0.40 (0.35–0.45)	0.97 (0.96–0.98)
Canadian (ROC curve)	≥ 3	0.72 (0.66–0.77)	0.28 (0.24–0.31)[b]	0.98 (0.97–0.98)
Emotional symptoms				
British	≥ 5	0.73 (0.67–0.78)	0.28 (0.24–0.31)	0.98 (0.97–0.98)
Canadian (Distributional)	≥ 6	0.63 (0.57–0.69)	0.35 (0.31–0.40)	0.97 (0.97–0.98)
Canadian (ROC curve)	≥ 4	0.80 (0.75–0.85)	0.21 (0.18–0.23)[b]	0.98 (0.98–0.99)[b]
Hyperactivity				
British	≥ 7	0.59 (0.52–0.65)	0.24 (0.21–0.28)	0.97 (0.96–0.97)
Canadian (distributional)	≥ 8	0.49 (0.43–0.56)	0.28 (0.24–0.33)[b]	0.96 (0.95–0.97)
Canadian (ROC curve)	≥ 5	0.80 (0.75–0.85)[b]	0.18 (0.16–0.21)[b]	0.98 (0.98–0.99)[b]
Peer problems				
British/Canadian (distributional)	≥ 4	0.60 (0.53–0.66)	0.33 (0.28–0.37)	0.97 (0.96–0.98)
Canadian (ROC curve)	≥ 3	0.75 (0.69–0.80)[b]	0.25 (0.22–0.28)[b]	0.98 (0.97–0.98)
Prosocial behaviour				
British	≤ 4	0.16 (0.12–0.22)	0.45 (0.34–0.56)	0.94 (0.94–0.95)
Canadian (distributional)	≤ 6	0.45 (0.38–0.51)[b]	0.34 (0.29–0.39)	0.96 (0.95–0.97)
Canadian (ROC curve)	≤ 8	0.74 (0.68–0.79)[b]	0.18 (0.15–0.20)[b]	0.98 (0.97–0.98)[b]
Total difficulties				
British	≥ 17	0.78 (0.73–0.83)	0.42 (0.37–0.46)	0.98 (0.98–0.99)
Canadian (Distributional)	≥ 16	0.82 (0.77–0.87)	0.38 (0.33–0.42)	0.99 (0.98–0.99)
Canadian (ROC curve)	≥ 14	0.87 (0.83–0.91)[b]	0.31 (0.27–0.34)[b]	0.99 (0.99–0.99)[b]
Anxiety disorder				
Conduct problems				
British/Canadian	≥ 4	0.52 (0.48–0.56)	0.60 (0.56–0.64)	0.91 (0.90–0.92)
Canadian (ROC curve)	≥ 3	0.65 (0.61–0.69)[b]	0.48 (0.44–0.51)[b]	0.93 (0.92–0.94)[b]
Emotional symptoms				
British	≥ 5	0.73 (0.69–0.76)	0.50 (0.47–0.53)	0.95 (0.94–0.95)
Canadian (Distributional)	≥ 6	0.62 (0.58–0.65)	0.58 (0.54–0.62)[b]	0.93 (0.92–0.94)
Canadian (ROC curve)	≥ 4	0.83 (0.80–0.86)[b]	0.41 (0.38–0.44)[b]	0.96 (0.95–0.97)[b]
Hyperactivity				
British	≥ 7	0.52 (0.48–0.56)	0.42 (0.39–0.46)	0.91 (0.90–0.92)
Canadian (Distributional)	≥ 8	0.41 (0.37–0.45)	0.46 (0.42–0.50)	0.89 (0.88–0.90)
Canadian (ROC curve)	≥ 5	0.75 (0.71–0.78)[b]	0.35 (0.33–0.38)[b]	0.94 (0.93–0.95)[b]
Peer problems				
British/Canadian	≥ 4	0.59 (0.55–0.63)	0.55 (0.52–0.59)	0.92 (0.91–0.93)
Canadian (ROC curve)	≥ 3	0.72 (0.68–0.75)[b]	0.45 (0.42–0.48)[b]	0.94 (0.93–0.95)[b]
Prosocial behaviour				
British	≤ 4	0.16 (0.13–0.19)	0.68 (0.60–0.75)	0.86 (0.85–0.87)
Canadian (Distributional)	≤ 6	0.43 (0.39–0.47)[b]	0.56 (0.52–0.61)	0.90 (0.89–0.91)
Canadian (ROC curve)	≤ 8	0.72 (0.68–0.75)[b]	0.35 (0.33–0.38)[b]	0.94 (0.93–0.95)[b]

TABLE 5 (continued)

Screening efficiency for existing and Canadian SDQ clinical cut-points from the distributional and ROC curve techniques: mood, anxiety and pervasive developmental disorders in the clinical population compared to the general population of children and youth[a]

SDQ scale	SDQ clinical cut-point	Sensitivity (95% CI)	PPV (95% CI)	NPV (95% CI)
Total difficulties				
British	≥ 17	0.73 (0.69–0.76)	0.63 (0.60–0.67)	0.95 (0.94–0.96)
Canadian (Distributional)	≥ 16	0.78 (0.74–0.81)	0.60 (0.56–0.63)	0.96 (0.95–0.96)
Canadian (ROC curve)	≥ 14	0.85 (0.82–0.87)[b]	0.53 (0.50–0.56)[b]	0.97 (0.96–0.97)[b]
Pervasive developmental disorder				
Conduct problems				
British/Canadian (Distributional)	≥ 4	0.51 (0.44–0.58)	0.32 (0.27–0.37)	0.97 (0.96–0.98)
Canadian (ROC curve)	≥ 3	0.66 (0.59–0.73)[b]	0.23 (0.19–0.26)[b]	0.98 (0.97–0.98)
Emotional symptoms				
British	≥ 5	0.75 (0.68–0.81)	0.25 (0.21–0.28)	0.98 (0.98–0.99)
Canadian (Distributional)	≥ 6	0.61 (0.54–0.68)	0.30 (0.26–0.35)	0.98 (0.97–0.98)
Canadian (ROC curve)	≥ 4	0.87 (0.82–0.91)[b]	0.19 (0.16–0.21)[b]	0.99 (0.99–0.99)[b]
Hyperactivity				
British	≥ 7	0.54 (0.47–0.61)	0.19 (0.16–0.23)	0.97 (0.96–0.98)
Canadian (Distributional)	≥ 8	0.38 (0.32–0.45)	0.20 (0.17–0.25)	0.96 (0.95–0.97)
Canadian (ROC curve)	≥ 5	0.75 (0.69–0.81)[b]	0.15 (0.13–0.17)	0.98 (0.98–0.99)[b]
Peer problems				
British/Canadian (Distributional)	≥ 4	0.65 (0.58–0.72)	0.31 (0.26–0.35)	0.98 (0.97–0.98)
Canadian (ROC curve)	≥ 3	0.78 (0.72–0.84)[b]	0.22 (0.19–0.25)[b]	0.98 (0.98–0.99)[b]
Prosocial behaviour				
British	≤ 4	0.16 (0.11–0.22)	0.40 (0.30–0.52)	0.95 (0.94–0.96)
Canadian (Distributional)	≤ 6	0.43 (0.36–0.50)[b]	0.29 (0.24–0.34)	0.97 (0.96–0.97)
Canadian (ROC curve)	≤ 8	0.75 (0.69–0.81)[b]	0.16 (0.13–0.18)[b]	0.98 (0.98–0.99)[b]
Total difficulties				
British	≥ 17	0.73 (0.67–0.79)	0.36 (0.31–0.40)	0.98 (0.98–0.99)
Canadian (Distributional)	≥ 16	0.79 (0.73–0.84)	0.33 (0.28–0.37)	0.99 (0.98–0.99)
Canadian (ROC curve)	≥ 14	0.88 (0.83–0.92)[b]	0.27 (0.24–0.31)[b]	0.99 (0.99–0.99)[b]

Abbreviations: CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristic; SDQ, Strengths and Difficulties Questionnaire.

[a] Outcome is clinical sample with specified mental disorder compared to general population sample.

[b] Significantly different from the British cut-points. Sensitivity is the proportion of true positives identified by the cut-point. PPV is the probability of having a mental health problem if meeting the clinical cut-point. NPV is the probability of not having a mental health problem if not meeting the clinical cut-point.

scales included 13.9% and 13.0% of the sample, reflecting the rising prevalence of mental health symptoms among Canadian children and youth, even 7 to 10 years ago. Therefore, it is likely that the existing clinical cut-points underestimate the true prevalence of mental health disorders in Canadian children and youth.

Third, the general population sample excluded those living in the territories and on reserves, whereas the clinical sample may have included these individuals. This reflects the differences in sampling techniques used in both samples. Fourth, we only used clinical data from a single institution, which may have contributed to the differences we observed. Together, limitations three and four limit the generalizability of our findings. Finally, the response rate for the CHMS was low. Despite applying survey weights to adjust for non-response bias, effects of residual confounding due to non-response bias may still exist.

Conclusion

The current study presents Canada-specific SDQ cut-points that more accurately categorizes the sample of Canadian children and youth. However, the existing British and the Canadian-specific distributional cut-points have small differences in screening effectiveness to predict mental health diagnoses in children and youth. Although we identified new Canadian cut-points using ROC curves, we do not recommend their use in practice due to lower specificity compared to the distributional approach. Future SDQ users may consider using the new Canadian distributional cut-points, to maximize specificity of the emotional symptoms and hyperactivity subscales, and the existing British cut-points to allow for historical and international comparisons.

TABLE 6
Screening efficiency for existing and Canadian SDQ clinical cut-points from the distributional and ROC curve techniques: conduct disorder and ADHD in the clinical population compared to the general population of children and youth[a]

SDQ scale	SDQ clinical cut-point	Sensitivity (95% CI)	PPV (95% CI)	NPV (95% CI)
Conduct disorder				
Conduct problems				
British/Canadian (Distributional)	≥ 4	0.61 (0.53–0.69)	0.30 (0.25–0.36)	0.98 (0.98–0.99)
Canadian (ROC curve)	≥ 3	0.72 (0.64–0.79)	0.20 (0.17–0.23)[b]	0.99 (0.98–0.99)
Emotional symptoms				
British	≥ 5	0.75 (0.67–0.81)	0.20 (0.17–0.24)	0.99 (0.98–0.99)
Canadian (Distributional)	≥ 6	0.61 (0.53–0.68)	0.25 (0.21–0.30)	0.98 (0.98–0.99)
Canadian (ROC curve)	≥ 4	0.81 (0.75–0.87)	0.15 (0.12–0.17)[b]	0.99 (0.99–0.99)[b]
Hyperactivity				
British	≥ 7	0.56 (0.48–0.64)	0.16 (0.13–0.20)	0.98 (0.97–0.98)
Canadian (Distributional)	≥ 8	0.46 (0.38–0.54)	0.19 (0.15–0.24)	0.97 (0.97–0.98)
Canadian (ROC curve)	≥ 5	0.77 (0.71–0.84)[b]	0.12 (0.10–0.14)	0.99 (0.98–0.99)
Peer problems				
British/Canadian (Distributional)	≥ 4	0.64 (0.56–0.71)	0.25 (0.21–0.30)	0.98 (0.98–0.99)
Canadian (ROC curve)	≥ 3	0.76 (0.69–0.83)	0.18 (0.15–0.21)[b]	0.99 (0.98–0.99)
Prosocial behaviour				
British	≤ 4	0.14 (0.09–0.20)	0.31 (0.21–0.43)	0.96 (0.95–0.97)
Canadian (Distributional)	≤ 6	0.46 (0.38–0.54)[b]	0.25 (0.20–0.31)	0.97 (0.97–0.98)
Canadian (ROC curve)	≤ 8	0.80 (0.73–0.86)[b]	0.13 (0.11–0.15)[b]	0.99 (0.98–0.99)[b]
Total difficulties				
British	≥ 17	0.80 (0.73–0.86)	0.32 (0.27–0.37)	0.99 (0.99–0.99)
Canadian (Distributional)	≥ 16	0.83 (0.76–0.88)	0.28 (0.24–0.33)	0.99 (0.99–0.99)
Canadian (ROC curve)	≥ 14	0.87 (0.82–0.93)	0.22 (0.19–0.26)[b]	0.99 (0.99–1.00)[b]
ADHD				
Conduct problems				
British/Canadian (Distributional)	≥ 4	0.53 (0.48–0.58)	0.49 (0.44–0.54)	0.95 (0.94–0.95)
Canadian (ROC curve)	≥ 3	0.68 (0.64–0.73)[b]	0.38 (0.34–0.41)[b]	0.96 (0.95–0.97)[b]
Emotional symptoms				
British	≥ 5	0.77 (0.72–0.81)	0.40 (0.36–0.43)	0.97 (0.96–0.98)
Canadian (Distributional)	≥ 6	0.62 (0.57–0.66)[b]	0.46 (0.42–0.51)	0.95 (0.95–0.96)
Canadian (ROC curve)	≥ 4	0.84 (0.80–0.88)	0.31 (0.28–0.34)[b]	0.98 (0.97–0.98)
Hyperactivity				
British	≥ 7	0.54 (0.49–0.59)	0.32 (0.29–0.36)	0.94 (0.93–0.95)
Canadian (Distributional)	≥ 8	0.41 (0.36–0.46)[b]	0.35 (0.31–0.40)	0.93 (0.92–0.94)
Canadian (ROC curve)	≥ 5	0.77 (0.73–0.81)[b]	0.26 (0.24–0.29)[b]	0.97 (0.96–0.97)[b]
Peer problems				
British/Canadian (Distributional)	≥ 4	0.62 (0.57–0.66)	0.45 (0.41–0.49)	0.95 (0.95–0.96)
Canadian (ROC curve)	≥ 3	0.77 (0.73–0.81)[b]	0.36 (0.33–0.39)[b]	0.97 (0.96–0.98)[b]
Prosocial behaviour				
British	≤ 4	0.18 (0.14–0.22)	0.59 (0.49–0.68)	0.91 (0.90–0.92)
Canadian (Distributional)	≤ 6	0.48 (0.43–0.53)[b]	0.47 (0.42,0.52)	0.94 (0.93–0.95)[b]
Canadian (ROC curve)	≤ 8	0.77 (0.73–0.81)[b]	0.27 (0.24–0.30)[b]	0.97 (0.96–0.97)[b]

Continued on the following page

Acknowledgements

Conflicts of interest

The authors report no conflicts of interest related to this study.

Authors' contributions and statement

JJL – Conceptualization, Project administration, Supervision, Methodology, Writing – Review & Editing, Data curation, Formal analysis.

SET – Validation, Methodology, Writing – Original draft, Data curation.

RLD – Validation, Supervision, Writing – Review & Editing, Methodology.

GG – Conceptualization, Project administration, Validation, Writing – Review & Editing, Methodology.

All authors – Validation, Writing – Review & Editing.

All authors approved the manuscript for publication.

The content and views expressed in this article are those of the authors and do not necessarily reflect those of the Government of Canada.

References

1. Mental Health Commission of Canada. Children and youth [Internet]. Ottawa (ON): MHCC; 2017 [cited 2022 Nov 13]. Available from: https://mentalhealthcommission.ca/what-we-do/children-and-youth/

2. Canadian Institute of Health Information. Children and Youth Mental Health in Canada [Internet]. Ottawa (ON): CIHI; 2022 [cited 2022 Nov 13]. Available from: https://www.cihi.ca/en/children-and-youth-mental-health-in-canada

3. Goodman R. The Strengths and Difficulties Questionnaire: a research note. J Child Psychol Psychiatry. 1997;38(5):581-6. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

4. Pastor PN, Reuben CA, Duran CR. Identifying emotional and behavioral problems in children aged 4-17 Years: United States, 2001-2007. Natl Health Stat Rep. 2012;48:1-17.

5. Vugteveen J, De Bildt A, Hartman CT, Timmerman ME. Using the Dutch multi-informant Strengths and Difficulties Questionnaire (SDQ) to predict adolescent psychiatric diagnoses. Eur Child Adolesc Psychiatry. 2018;27(10):1347-59. https://doi.org/10.1007/s00787-018-1127-y

6. Moriwaki A, Kamio Y. Normative data and psychometric properties of the Strengths and Difficulties Questionnaire among Japanese school-aged children. Child Adolesc Psychiatry Ment Health. 2014;8(1):1-12. https://doi.org/10.1186/1753-2000-8-1

7. Bourdon KH, Goodman R, Rae DS, Simpson G, Koretz DS. The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. J Am Acad Child Adolesc Psychiatry. 2005;44(6):557-64. https://doi.org/10.1097/01.chi.0000159157.57075.c8

8. Woerner W, Becker A, Rothenberger A. Normative data and scale properties of the German parent SDQ. Eur Child Adolesc Psychiatry. 2004;13(Suppl 2): ii3-10. https://doi.org/10.1007/s00787-004-2002-6

9. Hoffmann MD, Lang JJ, Guerrero MD, et al. Evaluating the psychometric properties of the parent-rated Strengths and Difficulties Questionnaire in a nationally representative sample of Canadian children and adolescents age 6 to 17 years. Health Rep. 2020; 31(8):13-20. https://doi.org/10.25318/82-003-x202000800002-eng

10. Statistics Canada. Canadian Health Measures Survey (CHMS) Data User Guide: Cycle 4. Ottawa (ON): Statistics Canada; 2017.

11. Day B, Langlois R, Tremblay M, Knoppers B.-M. Canadian Health Measures Survey: ethical, legal and social issues. Health Rep. 2007;18 Suppl:37-51.

12. Tremblay M, Wolfson M, Connor Gorber S. Canadian Health Measures Survey: rationale, background and overview. Health Rep. 2007;18 Suppl: 7-20.

13. Canadian Institute for Health Information. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10-CA). Ottawa (ON): CIHI; 2022.

14. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. J Am Acad Child Adolesc Psychiatry. 2001;40(11):1337-45. https://doi.org/10.1097/00004583-200111000-00015

15. Public Health Agency of Canada, Centre for Surveillance and Applied Research. Positive Mental Health Indicator Framework Quick Statistics, youth (12 to 17 years of age), Canada, 2022 Edition. Ottawa (ON): PHAC; 2022.

16. Trevethan R. Sensitivity, specificity, and predictive values: Foundations, pliabilities, and pitfalls in research and practice. Front Public Health. 2017;5(307):307. https://doi.org/10.3389/fpubh.2017.00307

17. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol. 2003;56: 1129-35. https://doi.org/10.1016/S0895-4356(03)00177-X

18. Fischer JE, Bachmann LM, Jaeschke R. A reader's guide to the interpretation of diagnostic test properties: clinical example of sepsis. Intensive Care Med. 2003;29(7):1043-51. https://doi.org/10.1007/s00134-003-1761-8

19. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. Br J Psychiatry. 2000;177:534-9. https://doi.org/10.1192/bjp.177.6.534