MyHEARTSMAP: development and evaluation of a psychosocial self-assessment tool, for and by youth

Punit Virk,[1] Samara Laskin,[2] Rebecca Gokiert,[3] Chris Richardson,[1] Mandi Newton,[4] Rob Stenstrom,[5] Bruce Wright,[4] Tyler Black,[6] Quynh Doan [2,7]

[1]School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada
[2]Pediatrics, University of British Columbia, Vancouver, British of Columbia, Canada
[3]Faculty of Extension, University of Alberta, Edmonton, Alberta, Canada
[4]Pediatrics, University of Alberta, Edmonton, Alberta, Canada
[5]Emergency Medicine, University of British Columbia, Vancouver, British Columbia, Canada
[6]Psychiatry, University of British Columbia, Vancouver, British Columbia, Canada
[7]BC Children's Hospital Research Institute, Vancouver, British Columbia, Canada

Correspondence to
Dr Quynh Doan; qdoan@bcchr.ca

ABSTRACT

Background Paediatric mental health-related visits to the emergency department are rising. However, few tools exist to identify concerns early and connect youth with appropriate mental healthcare. Our objective was to develop a digital youth psychosocial assessment and management tool (MyHEARTSMAP) and evaluate its inter-rater reliability when self-administered by a community-based sample of youth and parents.

Methods We conducted a multiphasic, multimethod study. In phase 1, focus group sessions were used to inform tool development, through an iterative modification process. In phase 2, a cross-sectional study was conducted in two rounds of evaluation, where participants used MyHEARTSMAP to assess 25 fictional cases.

Results MyHEARTSMAP displays good face and content validity, as supported by feedback from phase 1 focus groups with youth and parents (n=38). Among phase 2 participants (n=30), the tool showed moderate to excellent agreement across all psychosocial sections (κ=0.76–0.98).

Conclusions Our findings show that MyHEARTSMAP is an approachable and interpretable psychosocial assessment and management tool that can be reliably applied by a diverse community sample of youth and parents.

INTRODUCTION

Mental health conditions affect approximately 13%–23% of North American youth.[1 2] Delayed identification of mental health conditions may lead to crises and reliance on emergency department (ED) management.[3] Among youth presenting with non-mental health-related complaints to the ED, 20%–50% are found on screening to have mild to severe unrecognised or unmanaged mental health conditions.[4 5] These conditions may complicate management of physical complaints,[6] and increase emergency services utilisation.[7]

Early recognition of mental health conditions can lead to timely access to mental health services, thus improve health outcomes and utilisation of care.[8] While the American Academy of Pediatrics has recommended universal screening for mental health conditions among youth,[3] this has yet to be effectively implemented. Rising paediatric visits,[9] coupled with the ED's access to vulnerable populations,[10 11] and ability to manage acute screening results, make EDs a promising universal screening venue.[12] The ED provides an opportunity to evaluate broader psychosocial health, including substance use, education and other lifestyle factors.[13] Existing assessments include HEADS-ED, a clinician-administered evaluation of youths need for immediate intervention, with good inter-rater reliability and accuracy in predicting inpatient psychiatric admission.[14 15] HEARTSMAP is an expanded, but brief assessment and management tool for ED clinicians, which distinguishes psychiatric, social and behavioural concerns. This tool has good inter-rater reliability among diverse ED clinician types[16] and good predictive validity for inpatient psychiatric admissions.[17]

Universal screening implementation barriers include ED clinicians' inadequate mental health training,[18] time constraints,[19] integration into existing practices,[20] strained hospital resources and limited awareness of community care.[21] An online self-assessment could help reduce screening burden on clinicians and minimally impact ED flow.[22] Youth may prefer disclosing sensitive information over electronic interfaces versus face-to-face interaction.[23] Digital screening offers patients privacy, time to effectively articulate concerns and a sense of control over managing their well-being, without clinician judgement.[24] In the ED, electronic self-assessment is time and resource efficient, which may facilitate screening uptake.

To enable universal mental health self-screening in the ED, we proposed modifying HEARTSMAP for use as a self-administered online assessment by youth and family members (MyHEARTSMAP), and to evaluate its inter-rater reliability among them.

METHODS

Design

We conducted a multiphasic, multimethod study. In phase 1, we used qualitative methods to develop MyHEARTSMAP, a youth and family version of the clinical HEARTSMAP emergency assessment and management guiding tool. We used focus groups with youth and parents to establish tool content and face validity, and ensure tool structure, readability and content appropriateness. In phase 2, we engaged a cross-section of youth and parents to evaluate 25 fictional clinical vignettes, to evaluate MyHEARTSMAP inter-rater reliability.

Recruitment

A convenience sample of community-based youth and parents was recruited through the support of a mental health non-profit organisation, posters at a children's hospital and postings on the study's and non-profit partner's social media. We excluded youth with severe overall disability and non-English speakers. Phase 2 sample size was based on an intraclass correlation coefficient (ICC) power analysis,[25] equivalent to quadratically weighted kappas.[26] Thirty parent and youth raters were required to achieve a power of 80% to detect a kappa of 0.60 (substantial agreement) under the alternative hypothesis, assuming a kappa of 0.42 (moderate agreement) under the null hypothesis.

Instrument

The HEARTSMAP clinical tool served as a template in developing MyHEARTSMAP. The tool has clinicians' report across 10 psychosocial sections: home, education, alcohol and drugs, relationship and bullying, thoughts and anxiety, safety, sexual health, mood, abuse, and professional resources. Sections map to general domains: social, functional, youth health and psychiatry. For each section, concern severity is measured on a 4-point Likert-type scale from 0 (no concern) to 3 (severe concern), and services already accessed are measured on a separate 2-point scale (yes or no). Inputs from both scales feed into a built-in algorithm, triggering service recommendations with suggested time frames of access.[16 17] Scoring options on each severity scale have descriptive statements expanding on each score's conditions, helping clinicians decide on appropriate scores.

Study procedures

Phase 1 focus groups

Sixty-minute focus groups were held with up to five youth and three parents per group, in separate but simultaneous sessions. Smaller more numerous focus groups were used to facilitate in-depth discussion, and gain more varied input.

Each session followed the same structure. All participants had the opportunity to review the tool and inform its modification. A moderator introduced the tool's purpose and thoroughly reviewed its 10 psychosocial sections while a research assistant took comprehensive notes on group discussions. The first youth and parent focus groups reviewed an expanded version of the clinical tool. Modifications were made after each set of simultaneous youth and parent sessions, subsequent groups were presented with the up-to-date version, as shown in figure 1A.

First, participants went through each tool section, reviewing guiding questions, severity and resource scoring scale descriptors, with focus on improving usability. For each tool section, open-ended questions were used to assess participant's understanding of tool components, whether they felt the sections were important to youth their age (or other parents), if they could place themselves (or their child) on the scoring scale, and ways the tool could be improved. Each session ended with participants applying the reviewed MyHEARTSMAP version to three fictional vignettes. The first two cases familiarised participants with the tool and were completed as a group or independently with the opportunity to ask questions. We retained responses from the independently completed final case, reflecting participants' ability to use the tool.

Phase 2 inter-rater reliability evaluation

Participants completed MyHEARTSMAP for 25 fictional clinical vignettes, describing a range of paediatric psychosocial visits to the ED, from none to severe issues. Individually, participants completed a 45–60 min telephone or in-person training session with a research assistant prior to reviewing vignettes. Training included a 3 min instructional video and presentation overviewing MyHEARTSMAP sections, scoring guidelines and application to fictional cases. Participants also completed two to three training cases, scoring tool sections, and sought clarification when necessary. On training completion, vignettes were emailed in sets of five for remote completion at a self-directed pace, under parental supervision (youth participants). Vignette responses were captured in Research Electronic Data Capture (REDCap),[27] an online survey system. REDCap's activity logging feature was used to monitor duration, to ensure participants did not complete cases with unreasonable speed. After the first 10 cases, participants received a generic email highlighting close-reading strategies.

Procedures above were carried out in two consecutive rounds of evaluation shown in figure 1B. Between the rounds, participant feedback was incorporated into the tool version and vignettes, allowing further vignette and tool understandability refinement (eg, medical jargon, acronyms, word choice).

Analytical approach

Focus groups

We used qualitative content analysis to evaluate focus group transcripts.[28] Data saturation was reached when no new constructive feedback or tool modifications were proposed. Transcripts were coded, summarised into categories and reviewed by the study team to make tool modifications prior to subsequent groups. We compared average percent agreement for tool sections and domains
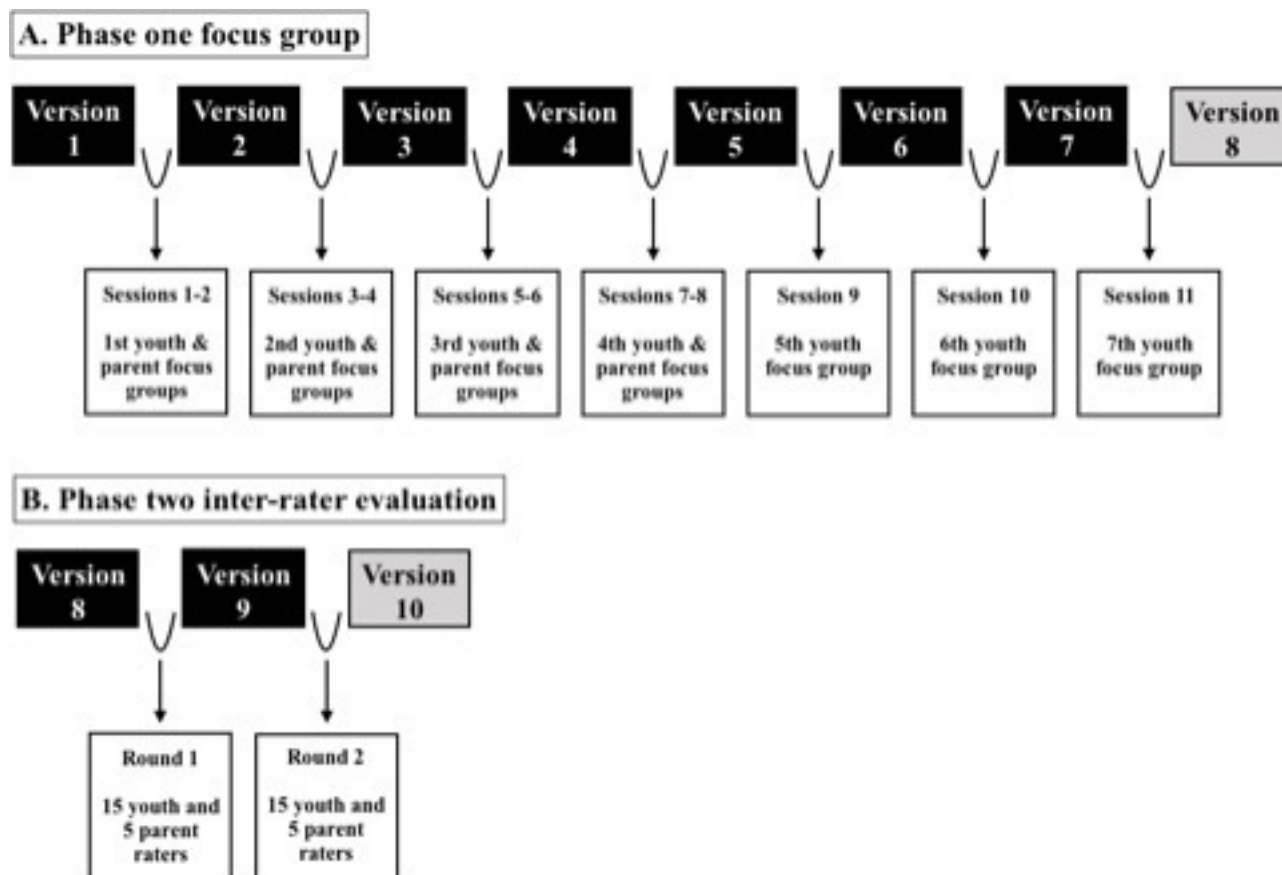
Figure 1 Schematic diagram showing the process of iterative modification that MyHEARTSMAP underwent in phase 1 (A) and phase 2 (B), with corresponding tool versions, sessions/rounds and participants involved.

on the independent test case, to measure changes in scoring consistency with iterative tool modifications. We compared average agreement between the first and second groups of youth using Fisher's exact test. We compared overall agreement across tool sections using a χ^2 test.

Inter-rater reliability evaluation

We used quadratically weighted kappa statistics to measure overall inter-rater agreement on tool sections and domains. We also conducted subgroup analyses, measuring section and domain agreement among participating youth and parents. The mean of all pairwise kappas was used as our index of agreement.[26] Statistical comparisons of kappas between or within each round of evaluation were carried out using Welch's t-test, χ^2 test and Fisher's exact test, with significance level at p=0.05. We report 95% CIs for all tests. Analyses were conducted using Microsoft Excel 2010 Data Analysis ToolPak (Microsoft, Redmond, Washington) and STATA V.15.0 (StataCorp, College Station, Texas).

Patient and public involvement

No patients were involved in the design, data collection or analysis of this study.

RESULTS
Focus groups

We recruited 38 participants, 9 parents and 29 youth into 11 focus groups, 7 with youth and 4 with parents. Sixteen were youth–parent dyad members and 22 were independent. A total of 71.1% of participants were female. The median age for participating youth was 16.0 years ranging from 10 to 17 years. All participants had some lived experience with mental health concerns. Additional details are summarised in table 1. Qualitative content analysis revealed two feedback categories—MyHEARTSMAP's approachability (covering relatability and accessibility) and interpretability.

Approachability of MyHEARTSMAP

Participants evaluating versions 1–2 (sessions 1–4) stressed the importance of being able to answer tool items honestly, without judgement from themselves or others (table 2) and being reluctant to choose a scoring option labelled as 'major concern.' Thus, Likert scale labels were changed to only include 0–3 numbering. Scoring descriptors were kept so participants could understand the general severity of each option. However, sometimes, score descriptors were only partially applicable, therefore an 'or' was introduced between statements allowing flexibility. Participants felt adding 'or' helped them more

Table 1 Demographic characteristics of study participants in phase 1 (focus groups) and phase 2 (inter-rater session)

	Phase 1: focus group sessions	Phase 2: inter-rater sessions
Total n (%)	38	30
Sex (female)	27 (71.1%)	21 (70.0%)
Parents	9 (23.7%)	10 (33.3%)
Youth	29 (76.3%)	20 (66.7%)
Median age, IQR* (years)	16.0 (3)	14.5 (2)
Ethnicity		
Caucasian	19 (50.0%)	13 (43.3%)
Visible minority†	19 (50.0%)	3 (10.0%)
Aboriginal	–	1 (3.33%)
Refused to answer	–	13 (43.3%)
Past mental health experiences‡		
Yes	38 (100%)	5 (16.7%)
No	–	12 (40.0%)
Refused to answer	–	13 (43.3%)

*IQR of participating youth's age.
†A visible minority, as defined by Statistics Canada, are 'persons, other than aboriginal peoples, who are non-Caucasian in race or non-white in colour'.[43]
‡Participants were asked whether they have had lived mental health experiences, regardless of a clinical diagnosis.

comfortably score. Reviewers also suggested descriptors be inclusive of youth with different experiences, such as 'homeschooled youth' and 'different romantic relationships.' Version 3 and onwards showed no new feedback with respect to how well participants related to the tool.

Interpretability of MyHEARTSMAP

On versions 3–6, feedback shifted towards tool language. Youth reviewing version 3 suggested some words might have multiple meanings, while on version 4, participants noted that idioms and terms such as 'contraception' and 'consensual' might be difficult for youth to understand. With these corrections, most comments on versions 5–7 (sessions 5–7) were reaffirming. Youth described the tool as 'easy to understand' and that it 'makes sense.' Figure 2 displays an example of progressive tool changes.

Test case

Overall agreement of focus group participants on MyHEARTSMAP sections ranged from 55% (Safety) to 97% (Abuse), with similar agreement patterns between youth and parents. Across sessions, sectional and domain scoring distributions varied significantly (p<0.001).

Inter-rater reliability evaluation

We recruited and trained 32 participants; however, two youth withdrew after training, prior to case review, leaving 10 parents and 20 youth. Participating youth's median age was 14.5 years, ranging from 12 to 17 years.

Table 1 displays their demographic information. Only 57% responded to questions about ethnicity and mental health experience. Among respondents, 10% identified as visible minorities, and 17% as having past mental health experiences.

Overall, we report high weighted kappa, displaying substantial to almost perfect agreement in both rounds (table 3). Significant (p<0.001) improvements were seen in nearly all sections between rounds 1 and 2. Clinically meaningful and statistically significant improvement was observed for 'Professionals & services', where agreement level rose from slight to substantial. Higher sectional kappas in round 2 were found when stratified by youth and parents; domain scores and tool-triggered recommendations also improved significantly (p<0.001).

DISCUSSION

MyHEARTSMAP was developed through an iterative process to be a psychosocial self-assessment and management guiding application. We saw excellent face and content validity in a diverse community sample of youth and families. Participants valued the tool's need to be easily interpretable, approachable for users, reflect different backgrounds and situations and reduce fears of judgement. The tool displayed strong inter-rater reliability when applied to fictional cases. Scoring consensus and significant improvements between evaluation rounds are quality indicators of MyHEARTSMAP assessment data and sources of evidence for tool reliability.[29]

There are few valid, reliable and brief tools for youth mental health self-assessment in the ED. The Behavioural Health Screen has been evaluated for acceptability and feasibility in the paediatric ED, where it saw an uptake rate of 33%, however it was not validated for ED use. While not specific to acute care, KIDSCREEN-27 is a European self-reporting tool for routine mental health monitoring and screening in school, home or clinical settings, for healthy and chronically ill youth.[30] KIDSCREEN-27 has been broadly validated and shares similar content and completion time (5–10 min) to MyHEARTSMAP.[31 32] KIDSCREEN-27 studies have shown inconsistent agreement with child–parent agreement ICCs ranging from 0.46 (poor-fair) to 0.74 (good).[31 32]

Variable and generally low agreement between youth and parents on psychosocial subscales in the above studies may reflect inherent tool properties (eg, response format, item content), or parental misperceptions. Youth can better assess their own experiences of internalising behaviours such as anxiety and depression compared with parents.[33] Parents as key informants may introduce discrepancies in assessing youth's mental health status. By providing all raters standardised vignettes on a fictional youth's psychosocial status, we eliminated the need for parental inference about their own child,[34] and found higher levels of agreement that may more closely reflect rater precision in applying and scoring with MyHEARTSMAP. However, agreement comparisons

Table 2 Summary of key categories, feedback and tool modifications from phase 1 parent and youth focus group sessions

Category		Tool version and sample feedback*	Tool modifications
Approachability	1	The title of the answer options in each section ('no', 'mild', 'moderate', 'severe' concern) implies judgement; I felt embarrassed to choose 'major concern'.	Scoring descriptors were limited to an ordinal number scale (0–3).
		Statements need to be more inclusive, for example, the 'Education' section should include homeschooled kids.	Scoring descriptors in the 'Education' section were updated to include homeschooled youth.
		Some kids may feel uncomfortable choosing a scoring option, because the category may have details that are not important to them, for example, someone may have anxiety but no mind tricks.	An 'or' was placed between statements in each scoring description, so youth do not need to meet all criteria mentioned to make a selection.
		Some words are confusing, when I read 'caregiver' I think about a housemaid or living support staff.	Terminology was simplified (eg, 'caregiver' was changed to parent/guardian).
	2	There is a sense of judgement associated with certain words/statements (eg, good grades).	Terminology with a potentially judgement connotation was removed (eg, changed 'good grade' to 'passing grades').
		Kids may perceive a specific behaviour to be acceptable if it is put in the zero score category.	Descriptors in the zero category were reviewed to ensure they represent age-appropriate and acceptable behaviour.
		In the 'Relationship and bullying' section, it is missing romantic partnerships kids may be in.	Romantic partners were included in the 'Relationship and bullying' section.
		The 'Professionals and resources' section should distinguish youth who have 'long-term' support from those who sought occasional or one-time help.	Long-term mental health support was explicitly mentioned in the 'Professionals and resources' section.
Interpretability	3	Some of the words used in the tool have other meanings (eg, trigger).	Terminology with other common meanings was removed and replaced.
		The scoring descriptions are too verbose.	Sentences were made shorter, less wordy, with emphasis on key points.
		Some of the vocabulary is too advanced for younger kids to understand (eg, consensual, recreational, abuse).	Complex language was simplified (eg, consensual was changed to 'agreed to'; abuse was changed to 'threatened or hurt').
		There need to be more examples to make some of the statements easier to understand, like giving broad examples where it says, 'practicing steps to end one's own life', so it's clear this is referring to suicide.	Examples were added to further clarify complex issues, for example, 'for practicing steps to end one's life,' examples such as 'holding rope around neck' were added.
	4	Where and how would the tool be used? And who would see the results?	
		Idioms may not be understood by other kids (eg, 'out of the blue').	Idioms were removed.
		Some of the vocabulary is challenging (eg, contraception).	The term 'contraception' was changed to 'protection'.
		This tool is very exciting.	
	5	The word 'isolated' may be difficult for some participants to understand.	The term 'isolated' was changed to 'alone'.
		Overall, it is really well written and easy to understand.	
		The examples used in the tool are helpful.	
	6	The tool makes sense and is easy to understand.	
		In the 'Relationship and bullying' section, 'fighting' with a romantic partner could be verbal or physical.	In the 'Relationship and bullying' section the term 'fight' was changed to 'argue'.
	7	The word 'harm' may be difficult for some participants to understand.	The term 'harm' was changed to 'hurt'.
		Everything was really clear and straightforward.	

*Sample feedback corresponding to the specific version of MyHEARTSMAP that participants reviewed.

Version 1 (expanded clinical tool version)

Guiding questions
- How would you rate your mood, with '0' being as low as possible, and '10' being perfectly happy?
- Do you feel down or depressed recently?
- Do you feel really happy or energetic lately?
- How often are you getting into trouble?

Severity scoring scale

0 (No Concerns or N/A)	I don't feel depressed or low.
1 (Mild Concerns)	Sometimes I feel sad or low, but I am generally in a good mood.
2 (Moderate Concerns)	I often feel sad or low or am cranky, and I act differently when I feel this way. Sometimes my mood gets better.
3 (Major Concerns)	I feel sad or low most of the time, and I act a lot differently when I feel this way. Sometimes, I have extreme mood swings (go from happy to sad really fast) for no reason.

Resource scoring scale

I have seen a healthcare provider for my mood concerns and there is a plan to deal with these problems.
I have not seen a healthcare provider for my mood concerns.

Version 5

Guiding questions
- How would I rate my mood, with '0' being the lowest possible, and '10' being perfectly happy?
- Have I felt down or depressed recently?
- Have I been having a lot of mood swings lately?
- How often am I getting into trouble with my parents or other authorities?
- Is my sadness affecting my ability to participate in normal day to day activities?
- Do I have abnormal behaviours (eg. needing to repeat actions multiple times, frequently lying)? Are these behaviours disruptive to others or to my own ability to function?

Severity scoring scale

0	I don't feel depressed or low, except for when bad things happen.
1	Sometimes I feel sad or low without a reason, but I am generally in a good mood.
2	I often feel sad or low, and I act differently when I feel this way. OR People around me have noticed a change in my personality, activity level, or quality of interactions
3	I feel sad or low most of the time, and it is affecting my sleep or eating habits. OR Sometimes, I have extreme mood swings (eg. go from happy to sad really fast) for no reason.

Resource scoring scale

I have seen a doctor/counsellor about my mood concerns and we have a diagnosis/plan to help me.
I have not seen a doctor/counsellor about my mood concerns.

Version 10 (latest version)

Guiding questions
- How would I rate my mood, with '0' being the lowest possible, and '10' being perfectly happy?
- Have I been having a lot of mood swings lately?
- How often am I getting into trouble with my parents or other authorities?
- Is my sadness affecting my ability to participate in normal activities?
- Do I have abnormal behaviours (e.g. needing to repeat specific actions multiple times)? Are these behaviours disruptive to others or to my own ability to function?

Severity scoring scale

0	I don't feel depressed or low, except for when bad things happen.
1	Sometimes I feel sad or low without a reason, but I am generally in a good mood.
2	I often feel sad or low, and I act differently when I feel this way. OR People around me have noticed a change in my personality or activity level.
3	I feel sad or low most of the time and it is affecting my sleep or eating habits. OR Sometimes, I have extreme mood swings (e.g. go from happy to sad really fast) for no reason.

Resource scoring scale

I have seen and been helped/treated by a doctor/counsellor about my mood concerns.
I have not seen and been helped/treated by a doctor/counsellor about my mood concerns.

Figure 2 Progression and transformation of MyHEARTSMAP's 'Mood' section, in accordance with tool versions shown in figure 1. N/A, not applicable.

Table 3 Quadratically weighted kappa statistics (95% CIs) measuring MyHEARTSMAP sectional agreement when applied by parents and youth (n=30) to a set of 25 fictional vignettes during phase 2 of the study

MyHEARTSMAP section	All participants (n=30)		Youth only (n=20)		Parent only (n=10)	
	Session 1	Session 2	Session 1	Session 2	Session 1	Session 2
Home	0.83	0.89	0.81	0.87	0.85	0.9
	(0.81 to 0.84)	(0.88 to 0.90)	(0.79 to 0.83)	(0.85 to 0.89)	(0.83 to 0.87)	(0.89 to 0.92)
Education and activities	0.79	0.81	0.82	0.8	0.73	0.83
	(0.77 to 0.81)	(0.79 to 0.83)	(0.80 to 0.84)	(0.77 to 0.83)	(0.66 to 0.80)	(0.79 to 0.89)
Alcohol and drugs	0.9	0.98	0.9	0.98	0.93	0.98
	(0.89 to 0.91)	(0.97 to 0.98)	(0.88 to 0.91)	(0.97 to 0.98)	(0.90 to 0.95)	(0.97 to 1.00)
Relationships and bullying	0.85	0.91	0.85	0.9	0.84	0.95
	(0.84 to 0.86)	(0.90 to 0.92)	(0.83 to 0.87)	(0.88 to 0.91)	(0.80 to 0.87)	(0.93 to 0.97)
Thoughts and anxiety	0.81	0.88	0.79	0.91	0.83	0.86
	(0.79 to 0.82)	(0.86 to 0.89)	(0.76 to 0.81)	(0.90 to 0.92)	(0.79 to 0.87)	(0.83 to 0.90)
Safety	0.85	0.85	0.84	0.84	0.88	0.86
	(0.83 to 0.85)	(0.83 to 0.87)	(0.82 to 0.85)	(0.81 to 0.87)	(0.86 to 0.90)	(0.81 to 0.91)
Sexual health	0.86	0.98	0.87	0.98	0.81	0.96
	(0.83 to 0.88)	(0.97 to 0.99)	(0.84 to 0.91)	(0.97 to 0.99)	(0.72 to 0.89)	(0.94 to 0.99)
Mood	0.8	0.94	0.79	0.93	0.81	0.95
	(0.78 to 0.81)	(0.93 to 0.94)	(0.76 to 0.82)	(0.92 to 0.94)	(0.74 to 0.87)	(0.93 to 0.96)
Abuse	0.8	0.95	0.81	0.93	0.78	1
	(0.77 to 0.84)	(0.93 to 0.98)	(0.76 to 0.86)	(0.89 to 0.97)	(0.61 to 0.96)	
Professionals and services	0.3	0.76	0.18	0.72	0.58	0.83
	(0.23 to 0.36)	(0.73 to 0.79)	(0.09 to 0.27)	(0.68 to 0.77)	(0.47 to 0.69)	(0.78 to 0.88)

made with KIDSCREEN-27 are made cautiously, given the different study populations, and kappa and ICC sensitivity to sample heterogeneity and prevalence.[35] Quadratically weighted kappas offer practical comparability to ICCs used in KIDSCREEN-27 studies. The primary outcome measure in these studies was between child–parent agreement, we measured overall sectional agreement on MyHEARTSMAP. However, our values were comparable to these other studies, as we saw nearly identical overall and among-group kappas.

Our study is strengthened by its methodological considerations for tool administration, using rater training and accountability measures for thoughtful scoring,[36] infrequently reported in inter-rater studies of psychosocial measures.[37] A self-administered psychosocial tool (Youth-CHAT) for opportunistic primary care screening also had end-users inform tool development.[38] While we received similar positive feedback for MyHEARTSMAP's ease of use and simplicity, our unique iterative approach allowed us to make ongoing modifications to address participant concerns, raised in both study phases, regarding item difficulty and need for age-appropriate language. MyHEARTSMAP's ability to reliably recommend management options is a novel addition to standard psychosocial self-assessment. Patients receiving and connecting with mental healthcare recommendations made in the ED

report generally higher healthcare satisfaction,[39] and are more likely to remain connected to care.[40] Generally, our participants spent 5–10 min on each case. However, as the tool is intended for self-assessment, evaluation of time spent self-reporting with MyHEARTSMAP will be conducted in an ongoing cohort study.

Study limitations include using note taking for focus group data collection instead of audio-recording discussions, preventing us from producing verbatim transcripts, but provided sufficient documentation for MyHEARTSMAP modifications without potentially stressing participants with audio-recording. We did not evaluate MyHEARTSMAP for reading level and while diverse, the small number of participants may not display reading comprehension issues more substantive in the general population. Furthermore, inter-rater agreement estimates may vary depending on tool application to patients or vignettes,[41] vignette use required rater training to ensure participants could comfortably score psychosocial information of fictional patients. While vignettes have been used in inter-rater studies and offer diverse, realistic,[42] ED mental health presentations an ongoing cohort study will evaluate whether scoring reliability differs when youth self-report with MyHEARTSMAP.

MyHEARTSMAP demonstrates good content and face validity and inter-rater reliability comparable, if not higher,

than similar tools. Following prospective evaluation of its predictive validity, we intend for MyHEARTSMAP to be accessible to youth and families visiting acute and paediatric primary care settings as a downloadable application. Clinicians may offer MyHEARTSMAP on a mobile device or stationary computer in waiting rooms, for universal screening and discuss appropriate mental health services recommendations as needed.

What is known about the subject?

► Mental health concerns in youth often go unrecognised, leading to poor health outcomes, and crisis-driven management in acute care settings.
► Universal screening has been recommended, but not implemented due to lack of reliable, effective and efficient methods.

What this study adds?

► A digital self-administered psychosocial assessment and management tool (MyHEARTSMAP) was developed and evaluated for use by youth and parents in emergency care.
► MyHEARTSMAP is well positioned for evaluation for universal screening in primary and acute care settings that see youth with or without identified mental health concerns.

REFERENCES
1. Mental Health Commission of Canada. Making the case for investing in mental health in Canada, 2016. Available: https://www.mentalhealthcommission.ca/sites/default/files/2016-06/Investing_in_Mental_Health_FINAL_Version_ENG.pdf [Accessed 18 Mar 2019].
2. Perou R, Bitsko RH, Blumberg SJ, *et al*. Mental health surveillance among children-United states, 2005-2011. *MMWR* 2013;62:1–35.
3. Dolan MA, Fein JA. Committee on pediatric emergency medicine. pediatric and adolescent mental health emergencies in the emergency medical services system. *Pediatrics* 2011;127:e1356–66.
4. Ramsawh HJ, Chavira DA, Kanegaye JT, *et al*. Screening for adolescent anxiety disorders in a pediatric emergency department. *Pediatr Emerg Care* 2012;28:1041–7.
5. Scott EG, Luxmore B, Alexander H, *et al*. Screening for adolescent depression in a pediatric emergency department. *Acad Emerg Med* 2006;13:537–42.
6. Shefer G, Henderson C, Howard LM, *et al*. Diagnostic Overshadowing and other challenges involved in the diagnostic process of patients with mental illness who present in emergency departments with physical symptoms – a qualitative study. *PLoS One* 2014;9:11.
7. Brennan JJ, Chan TC, Hsia RY, *et al*. Emergency department utilization among frequent users with psychiatric visits. *Acad Emerg Med* 2014;21:1015–22.
8. O'Connell ME, Boat T, Warner KE. *Preventing mental, emotional, and behavioral disorders among young people: progress and possibilites*. Washington D C: The National Academies Press, 2009.
9. Mapelli E, Black T, Doan Q. Trends in pediatric emergency department utilization for mental health-related visits. *J Pediatr* 2015;167:905–10.
10. Wilson KM, Klein JD. Adolescents who use the emergency department as their usual source of care. *Arch Pediatr Adolesc Med* 2000;154:361–5.
11. Klein JD, Woods AH, Wilson KM, *et al*. Homeless and runaway youths' access to health care. *J Adolesc Health* 2000;27:331–9.
12. Horowitz LM, Ballard ED, Pao M. Suicide screening in schools, primary care and emergency departments. *Curr Opin Pediatr* 2009;21:620–7.
13. Van Amstel LL, Lafleur DL, Blake K. Raising our HEADSS: adolescent psychosocial documentation in the emergency department. *Acad Emerg Med* 2004;11:648–55.
14. Cappelli M, Gray C, Zemek R, *et al*. The HEADS-ED: a rapid mental health screening tool for pediatric patients in the emergency department. *Pediatrics* 2012;130:e321–7.
15. Cappelli M, zemek R, Polihronis C. The HEADS-ED: evaluating the clinical use of aBrief, Action-Oriented, pediatric mental health screening tool. *Pediatr Emerg Care* 2017.
16. Virk P, Stenstrom R, Doan Q. Reliability testing of the HEARTSMAP psychosocial assessment tool for multidisciplinary use and in diverse emergency settings. *Paediatr Child Health* 2018:1–6.
17. Lee A, Deevska M, Stillwell K, *et al*. A psychosocial assessment and management tool for children and youth in crisis. *Can J Emerg Med* 2018:1–10.
18. Zun L. Care of psychiatric patients: the challenge to emergency physicians. *West J Emerg Med* 2016;17:173–6.
19. Habis A, Tall L, Smith J, *et al*. Pediatric emergency medicine physicians' current practices and beliefs regarding mental health screening. *Pediatr Emerg Care* 2007;23:387–93.
20. Betz ME, Boudreaux ED. Managing suicidal patients in the emergency department. *Ann Emerg Med* 2016;67:276–82.
21. Cloutier P, Kennedy A, Maysenhoelder H, *et al*. Pediatric mental health concerns in the emergency department: caregiver and youth perceptions and expectations. *Pediatr Emerg Care* 2010;26:99–106.
22. Chun TH, Duffy SJ, Linakis JG. Emergency department screening for adolescent mental health disorders: the who, what, when, where, why and how it could and should be done. *Clin Pediatr Emerg Med* 2013;14:3–11.
23. Bradford S, Rickwood D. Young people's views on electronic mental health assessment: prefer to type than talk? *J Child Fam Stud* 2015;24:1213–21.
24. Bradford S, Rickwood D. Acceptability and utility of an electronic psychosocial assessment (myAssessment) to increase self-disclosure in youth mental healthcare: a quasi-experimental study. *BMC Psychiatry* 2015;15:305.
25. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998;17:101–10.
26. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol* 2012;8:23–34.
27. Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
28. Mayan M. Essentials of Qualitative Inquiry. In: Morse J, ed. *Qualitative research*. New York: Routledge, 2016: 93–8.
29. Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ* 2004;38:1006–12.
30. Ravens-Sieberer U, Herdman M, Devine J, *et al*. The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances. *Qual Life Res* 2014;23:791–803.
31. Janssens A, Thompson Coon J, Rogers M, *et al*. A systematic review of generic multidimensional patient-reported outcome measures for children, part I: descriptive characteristics. *Value Health* 2015;18:315–33.
32. Janssens A, Rogers M, Thompson Coon J, *et al*. A systematic review of generic multidimensional patient-reported outcome measures for children, part II: evaluation of psychometric performance of English-language versions in a general population. *Value Health* 2015;18:334–45.

33. DiBartolo PM, Grills AE. Who is best at predicting children's anxiety in response to a social evaluative task? A comparison of child, parent, and teacher reports. *J Anxiety Disord* 2006;20:630–45.
34. Comer JS, Kendall PC. A symptom-level examination of parent-child agreement in the diagnosis of anxious youths. *J Am Acad Child Adolesc Psychiatry* 2004;43:878–86.
35. de Vet HCW, Terwee CB, Knol DL, *et al*. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006;59:1033–9.
36. Shweta BRC, Chaturvedi HK. Evaluation of inter-rater agreement and inter-rater reliability for observational data: an overview of concepts and methods. *J Indian Acad Appl Psychol* 2015;41:20–7.
37. Rosen J, Mulsant BH, Marino P, *et al*. Web-Based training and interrater reliability testing for scoring the Hamilton depression rating scale. *Psychiatry Res* 2008;161:126–30.
38. Goodyear-Smith F, Corter A, Suh H, *et al*. Electronic screening for lifestyle issues and mental health in youth: a community-based participatory research approach. *BMC Med Inform Decis Mak* 2016;16:1–8.
39. Cappelli M, Cloutier P, Newton AS, *et al*. Evaluating mental health service use during and after emergency department visits in a multisite cohort of Canadian children and youth. *Can J Emerg Med* 2017:1–12.
40. Frosch E, dosReis S, Maloney K. Connections to outpatient mental health care of youths with repeat emergency department visits for psychiatric crises. *PS* 2011;62:646–9.
41. Brunner E, Probst M, Meichtry A, *et al*. Comparison of clinical vignettes and standardized patients as measures of physiotherapists' activity and work recommendations in patients with non-specific low back pain. *Clin Rehabil* 2016;30:85–94.
42. Hjortsø S, Butler B, Clemmesen L, *et al*. The use of case vignettes in studies of interrater reliability of psychiatric target syndromes and diagnoses. A comparison of ICD-8, ICD-10 and DSM-III. *Acta Psychiatr Scand* 1989;80:632–8.
43. Statistics Canada. Visible Minority and Population Group Reference Guide, 2016. Available: from:http://www12.statcan.gc.ca/censusrecensement/2016/ref/guides/006/98-500-x2016006-eng.cfm