BMJ Paediatrics Open

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

SCHOLARONE™
Manuscripts

MyHEARTSMAP: Development and evaluation of a psychosocial self-assessment tool, for

and by youth

Punit Virk[1]; Samara Laskin[2] MD, FRCPC; Rebecca Gokiert[3] PhD; Chris Richardson[1] PhD;

Mandi Newton[4,5] PhD RN; Rob Stenstrom[6] MD; Bruce Wright[4,5] MD; Tyler Black[7] MD; Quynh

Doan[1,2,6,8] MDCM, MHSc, PhD

Affiliations:

1. School of Population and Public Health, Faculty of Medicine, University of British

   Columbia; Vancouver, British Columbia, Canada

2. Department of Pediatrics, Faculty of Medicine, University of British Columbia;

   Vancouver, British Columbia, Canada

3. Faculty of Extension, University of Alberta; Edmonton, Alberta, Canada

4. Department of Pediatrics, Faculty of Medicine and Dentistry University of Alberta;

   Edmonton, Alberta, Canada

5. Women and Children's Hospital Research Institute; Edmonton, Alberta, Canada

6. Department of Emergency Medicine, Faculty of Medicine, University of British

   Columbia; Vancouver, British Columbia, Canada

7. Department of Psychiatry, University of British Columbia, Faculty of Medicine,

   University of British Columbia; Vancouver, British Columbia, Canada

8. BC Children's Hospital Research Institute; Vancouver, British Columbia, Canada

Word count: 2499

Correspondence to:

Quynh Doan, Faculty of Medicine, Department of Pediatrics, University of British Columbia,

4480 Oak Street, Vancouver, BC, CAN V6T 1Z3, [qdoan@bcchr.ca], 604-875-2345 ext. 3691,

fax: 604-875-2366

ABSTRACT

Background: Pediatric mental health-related visits to the emergency department are rising.

However, few tools exist to identify concerns early and connect youth with appropriate mental

health care. Our objective was to develop a digital youth psychosocial assessment and

management tool (MyHEARTSMAP) and evaluate its interrater reliability when self-

administered by a community-based sample of youth and parents.

Methods: We conducted a multi-phasic, multi-method study with community-based youth and

parents. In phase one, focus group sessions were used to inform tool development, through an

iterative modification process. In phase two, a cross-sectional study that involved 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 one focus groups with youth and parents (n=38). Among phase two participants

(n=30), the tool showed moderate to excellent agreement across all psychosocial sections

(κ=0.76 to 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.

What is known about the subject?

- Mental health concerns in youth often go unrecognized, 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 hopes to add?

- 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.

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 management (ED).(3) Among youth presenting with non-mental health related complaints to the ED, 20-50% are found upon screening to have mild to severe unrecognized or unmanaged mental health conditions.(4,5) These conditions may complicate management of physical complaints, (6) and increase emergency services utilization.(7)

Early recognition of mental health conditions can lead to timely access to mental health services, thus, improve health outcomes and utilization of care.(8) While the American Academy of Pediatrics has recommended universal screening for mental health conditions amongst youth,(3) this has yet to be effectively implemented. Rising pediatric 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 interrater reliability and accuracy in predicting in-patient psychiatric admission.(14) HEARTSMAP is an expanded, but brief assessment and management tool for ED clinicians, that distinguishes psychiatric, social, and behavioural concerns. This tool has good interrater reliability among diverse ED clinician-types(15) and good predictive validity for in-patient psychiatric admissions.(16)

Universal screening implementation barriers include ED clinicians' inadequate mental health training,(17) time constraints,(18) integration into existing practices,(19) strained hospital resources, and limited awareness of community care.(14) An online self-assessment could help reduce screening burden on clinicians and minimally impact ED flow.(20) Youth may prefer disclosing sensitive information over electronic interfaces versus face-to-face interaction.(21) Digital screening offers patients privacy, time to effectively articulate concerns, and a sense of control over managing their well-being without clinician judgement.(22) 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 interrater reliability among them.

METHODS

<u>Design</u>

We conducted a multi-phasic, multi-method study. In phase one, 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 two, we engaged a cross-section of youth and parents to evaluate 25 fictional clinical vignettes, to evaluate MyHEARTSMAP interrater reliability. This study was approved by our local institutional ethics review board.

<u>Recruitment</u>

A convenience sample of community-based youth and parents was recruited through the support of a mental health non-profit organization, 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 two sample size was based on an intraclass correlation (ICC) power analysis,(23) equivalent to quadratically weighted kappas.(24) 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 & drugs, Relationship & bullying, Thoughts & anxiety, Safety, Mood, Abuse, Professional resources. Sections map to general domains: Social, Functional, Youth health, 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 is measured on a separate 2-point scale (yes or no). Input from both scales feed into a built-in algorithm, triggering service recommendations with suggested time frames of access.(15,16) Scoring options on each severity scale have descriptive statements expanding on each score's conditions, helping clinicians decide on appropriate scores.

Study Procedures

Phase One Focus Groups:

Sixty-minute focus groups were held with up to five youth and three parents per group, in separate but simultaneously 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 ten 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 scales 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 familiarized 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 Two Interrater Reliability Evaluation:

Participants completed MyHEARTSMAP for 25 fictional clinical vignettes, describing a range of pediatric psychosocial visits to the ED, from none to severe issues. Individually, participants completed a 45-60-minute telephone or in-person training session with a research

assistant prior to reviewing vignettes. Training included a 3-minute instructional video and presentation overviewing MyHEARTSMAP sections, scoring guidelines, and application to fictional cases. Participants also completed 2-3 training cases, scoring tool sections and sought clarification when necessary. Upon 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 REDCap,(25) 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 ten 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 (e.g., medical jargon, acronyms, word choice).

Analytic approach:

Focus groups

We used qualitative content analysis to evaluate focus group transcripts.(26) Data saturation was reached when no new constructive feedback or tool modifications were proposed. Transcripts were coded, summarized 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 on the independent test case, to measure changes in scoring consistency with iterative tool modifications. We compared average agreement between the first and second

group of youth using Fischer's Exact test. We compared overall agreement across tool sections, using a Chi-Square test.

Interrater Reliability Evaluation:

We used quadratically weighted kappa statistics to measure overall interrater agreement on tool sections and domains. We also conducted sub-group analyses, measuring section and domain agreement among participating youth and parents. The mean of all pairwise kappas was used as our index of agreement.(24) Statistical comparisons of kappas between or within each round of evaluation were carried out using Welch's t-test, Chi-Square test and Fischer's Exact test, with significance level at p=0.05. We report 95% confidence intervals for all tests. Analyses were conducted using Microsoft Excel 2010 Data Analysis Toolpak (Microsoft, Redmond, Washington) and STATA 15.0 (Stata Corporation, 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% of participants were female. The median age for participating youth was 16.0 years ranging, from 10-17 years. All participants had some lived experience with mental health concerns. Additional details are summarized in table 1. Qualitative content analysis revealed two feedback categories—MyHEARTSMAP's approachability (covering relatability and accessibility) and interpretability.

Table 1. Demographic characteristics of study participants in phase one (focus groups) and two (interrater session).

	Phase one: Focus group sessions	Phase two: Interrater sessions
Total N (%)	38	30
Sex (female)	27 (71.0%)	21 (70.0%)
Parents	9 (23.7%)	10 (33.3%)
Youth	29 (76.3%)	20 (66.7%)
Median age, IQR[a] (years)	16.0 (3)	14.5 (2)
Ethnicity		
Caucasian	19 (50.0%)	13 (43.3%)
Visible minority[b]	19 (50.0%)	3 (10.0%)
Aboriginal	-	1 (3.30%)
Refused to answer	-	13 (43.3%)
Past mental health experiences[c]		
Yes	38 (100%)	5 (16.7%)
No	-	12 (40.0%)
Refused to answer	-	13 (43.3%)

[a]Interquartile range of participating youth's age.

[b]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".

[c]Participants were asked whether they experienced mental health concerns in the past, regardless of a clinical diagnosis.

Approachability of MyHEARTSMAP

Participants evaluating version 1-2 (sessions 1-4) stressed the importance of being able to answer tool items honestly, without judgment from themselves or others (table 2) and being reluctant to choose a scoring option labeled 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 comfortably score. Reviewers also suggested descriptors be inclusive of youth with different lifestyles, such as "homeschooled youth" and "different romantic relationships." Versions 3 and onward 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 re-affirming. Youth described the tool as "easy to understand" and that it "makes sense." Figure 2 displays an example of progressive tool changes.

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

Category	Tool version & Sample Feedback[a]		Tool Modifications
Approachability	1	The title of the answer options in each section ('no', 'mild', 'moderate', 'severe' concern), imply judgment 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 (e.g. 'caregiver' was changed to parent/guardian)
	2	There is a sense of judgement associated with certain words/statements (e.g. good grades)	Terminology with a potentially judgment connotation was removed (e.g. changed 'good grade' to 'passing grades')
		Kids may perceive a specific behavior 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 behavior
		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 (e.g. trigger)	Terminology with other common meanings were 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 (e.g. consensual, recreational, abuse)	Complex language was simplified (e.g. 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 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 (e.g. 'out of the blue')	Idioms were removed
		Some of the vocabulary is challenging (e.g. 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'
		The word 'harm' may be difficult for some participants to understand	The term 'harm' was changed to 'hurt'
	7	Everything was really clear and straightforward	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

[a]Sample feedback corresponding to the specific version of MyHEARTSMAP that participants reviewed.

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 ($p < 0.001$).

Interrater reliability evaluation

We recruited and trained 32 participants, however 2 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 12-17 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 section 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 ($p < 0.001$).

Table 3. Quadratically weighted kappa statistics (95% confidence intervals) measuring MyHEARTSMAP sectional agreement when applied by parents and youth (N=30) to a set of 25 fictional vignettes during phase two 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
	(0.81-0.84)	(0.88-0.90)	(0.79-0.83)	(0.85-0.89)	(0.83-0.87)	(0.89-0.92)
Education & activities	0.79	0.81	0.82	0.8	0.73	0.83
	(0.77-0.81)	(0.79-0.83)	(0.80-0.84)	(0.77-0.83)	(0.66-0.80)	(0.79-0.89)
Alcohol & drugs	0.9	0.98	0.9	0.98	0.93	0.98
	(0.89-0.91)	(0.97-0.98)	(0.88-0.91)	(0.97-0.98)	(0.90-0.95)	(0.97-1.00)
Relationships & bullying	0.85	0.91	0.85	0.9	0.84	0.95
	(0.84-0.86)	(0.90-0.92)	(0.83-0.87)	(0.88-0.91)	(0.80-0.87)	(0.93-0.97)
Thoughts & anxiety	0.81	0.88	0.79	0.91	0.83	0.86
	(0.79-0.82)	(0.86-0.89)	(0.76-0.81)	(0.90-0.92)	(0.79-0.87)	(0.83-0.90)
Safety	0.85	0.85	0.84	0.84	0.88	0.86
	(0.83-0.85)	(0.83-0.87)	(0.82-0.85)	(0.81-0.87)	(0.86-0.90)	(0.81-0.91)

Sexual health	0.86	0.98	0.87	0.98	0.81	0.96
	(0.83-0.88)	(0.97-0.99)	(0.84-0.91)	(0.97-0.99)	(0.72-0.89)	(0.94-0.99)
Mood	0.8	0.94	0.79	0.93	0.81	0.95
	(0.78-0.81)	(0.93-0.94)	(0.76-0.82)	(0.92-0.94)	(0.74-0.87)	(0.93-0.96)
Abuse	0.8	0.95	0.81	0.93	0.78	1
	(0.77-0.84)	(0.93-0.98)	(0.76-0.86)	(0.89-0.97)	(0.61-0.96)	
Professionals & services	0.3	0.76	0.18	0.72	0.58	0.83
	(0.23-0.36)	(0.73-0.79)	(0.09-0.27)	(0.68-0.77)	(0.47-0.69)	(0.78-0.88)

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 to 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 judgment. The tool displayed strong interrater 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.(27)

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 pediatric 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.(28) KIDSCREEN-27 has been broadly validated and shares similar content and completion time (5-10-minutes) to MyHEARTSMAP.(29,30) KIDSCREEN-27 studies have shown inconsistent agreement with child-parent agreement ICC's ranging from 0.46 (poor-fair) to 0.74 (good). (29,30)

Variable and generally low agreement between youth and parents on psychosocial subscales in the above studies may reflect inherent tool properties (e.g., response format, item content), or parental misperceptions. Youth can better assess their own experiences of internalizing behaviours such as anxiety and depression compared to parents.(31) Parents as key informants may introduce discrepancies in assessing youth's mental health status. By providing all raters standardized vignettes on a fictional youth's psychosocial status, we eliminated the

need for parental inference about their own child,(32) and found higher levels of agreement that may more closely reflect rater precision in applying and scoring with MyHEARTSMAP. However, agreement comparisons made with KIDSCREEN-27 are made cautiously, given the different study populations, and kappa and ICC sensitivity to sample heterogeneity and prevalence.(33) Quadratically weighted kappa's 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,(34) infrequently reported in interrater studies of psychosocial measures.(35) A self-administered psychosocial tool (YouthCHAT) for opportunistic primary care screening also had end-users inform tool development. (36) While we received similar positive feedback for MyHEARTSMAP's ease-of-use and simplicity, our unique iterative approach allowed us to make on-going 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 health care recommendations made in the ED, report greater ED visit satisfaction,(37) and are more likely to remain connected.(38) Generally, participants spent 5-10 minutes 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, interrater agreement estimates may vary depending on tool application to patients or vignettes,(39) vignette use required rater training to ensure participants could comfortably score psychosocial information of fictional patients. While vignettes have been used in interrater studies and offer diverse, realistic, ED mental health presentations,(40)an ongoing cohort study will evaluate whether scoring reliability differs when youth self-report with MyHEARTSMAP.

MyHEARTSMAP demonstrates good content and face validity and interrater 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 pediatric 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.

PATIENT CONSENT FOR PUBLICATION

Not required.

FUNDING

COMPETING INTEREST

None to declare.

DATA ACCESS

Data will not be made available to protect participant identity, as confidentiality cannot be fully guaranteed, given the small sample size which was collected in a fixed time period through specific institutions.

REFERENCES

1. Mental Health Commission of Canada. Making the Case for Investing in Mental Health in Canada. 2016. https://www.mentalhealthcommission.ca/sites/default/files/2016-06/Investing_in_Mental_Health_FINAL_Version_ENG.pdf (accessed 18 March 2019)

2. Perou R, Bitsko RH, Blumberg SJ, Pastor P, Ghandour RM, Gfroerer JC, et al. Mental health surveillance among children--United States, 2005-2011. *MMWR Suppl.* 2013 May 17;62(2):1–35. https://www.cdc.gov/mmwr/pdf/other/su6202.pdf (accessed 18 March 2019)

3. Dolan MA, Fein JA, Committee on Pediatric Emergency Medicine. Pediatric and Adolescent Mental Health Emergencies in the Emergency Medical Services System. *Pediatrics*. 2011 May 1;127(5):e1356–66. http://www.ncbi.nlm.nih.gov/pubmed/21518712 (accessed 18 March 2019)

4. Ramsawh HJ, Chavira DA, Kanegaye JT, Ancoli-Israel S, Madati PJ, Stein MB. Screening for Adolescent Anxiety Disorders in a Pediatric Emergency Department. *Pediatr Emerg Care*. 2012 Oct;28(10):1041–7. http://www.scopus.com/inward/record.url?eid=2-s2.0-84867399707&partnerID=tZOtx3y1 (accessed 18 March 2019)

5. Scott EG, Luxmore B, Alexander H, et al. Christopher NC. Screening for Adolescent

Depression in a Pediatric Emergency Department. *Acad Emerg Med*. 2006;13(5):537–42.

6.   Shefer G, Henderson C, Howard LM, at 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. Dekel S, editor.

PLoS One. 2014 Nov 4;9(11). http://www.ncbi.nlm.nih.gov/pubmed/25369130 (accessed

18 March 2019)

7.   Brennan JJ, Chan TC, Hsia RY, et al. Emergency Department Utilization Among Frequent

Users With Psychiatric Visits. Mycyk M, editor. *Acad Emerg Med*. 2014 Sep;21(9):1015–

22. http://www.ncbi.nlm.nih.gov/pubmed/25269582 (accessed 18 March 2019)

8.   National Research Council and Institute of Medicine. Preventing Mental, Emotional, and

Behavioral Disorders Among Young People: Progress and Possibilites. O'Connell ME,

Boat T, Warner KE, editors. Washington, D.C.: The National Academies Press; 2009.

Available from: http://www.nap.edu/catalog/12480

9.   Mapelli E, Black T, Doan Q. Trends in Pediatric Emergency Department Utilization for

Mental Health-Related Visits. *J Pediatr*. 2015 Oct;167(4):905–10.

http://linkinghub.elsevier.com/retrieve/pii/S0022347615007192 (accessed 18 March 2019)

10. Wilson KM, Klein JD. Adolescents Who Use the Emergency Department as Their Usual

Source of Care. *Arch Pediatr Adolesc Med*. 2000 Apr 1;154(4):361–5.

http://archpedi.jamanetwork.com/article.aspx?doi=10.1001/archpedi.154.4.361 (accessed

18 March 2019)

11. Klein JD, Woods AH, Wilson KM, et al. Homeless and runaway youths' access to health

care. *J Adolesc Heal*. 2000 Nov;27(5):331–9.

http://www.ncbi.nlm.nih.gov/pubmed/11044705 (accessed 18 March 2019)

12. Horowitz LM, Ballard ED, Pao M. Suicide screening in schools, primary care and emergency departments. *Curr Opin Pediatr*. 2009 Oct;21(5):620–7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2879582/ (accessed 18 March 2019)

13. van Amstel LL, Lafleur DL, Blake K. Raising Our HEADSS: Adolescent Psychosocial Documentation in the Emergency Department. *Acad Emerg Med*. 2004 Jun;11(6):648–55. http://doi.wiley.com/10.1197/j.aem.2003.12.022 (accessed 18 March 2019)

14. 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 Feb;26(2):99–106. https://www.ncbi.nlm.nih.gov/pubmed/20094002 (accessed 18 March 2019)

15. 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 Mar;1–6. https://academic.oup.com/pch/article/23/8/503/4922416 (accessed 18 March 2019)

16. 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 Mar 28;1–10.

17. Zun L. Care of Psychiatric Patients: The Challenge to Emergency Physicians. *West J Emerg Med*. 2016 Mar 8;17(2):173–6. http://escholarship.org/uc/item/3bc0r3vs (accessed 18 March 2019)

18. Habis A, Tall L, Smith J, et al. Pediatric Emergency Medicine Physicians' Current Practices and Beliefs Regarding Mental Health Screening. *Pediatr Emerg Care*.2007 Jun;23(6):387–93. https://www.ncbi.nlm.nih.gov/pubmed/17572523 (accessed 18 March 2019)

19.   Betz ME, Boudreaux ED. Managing Suicidal Patients in the Emergency Department. *Ann Emerg Med*. 2016;67(2):276–82. https://www.annemergmed.com/article/S0196-0644(15)01264-0/pdf (accessed 18 March 2019)

20.   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(1):1199-1216

21.   Bradford S, Rickwood D. Young People's Views on Electronic Mental Health Assessment: Prefer to Type than Talk? *J Child Fam Stud*. 2015;24(5):1213–21.

22.   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 Dec 1;15(1):305. http://dx.doi.org/10.1186/s12888-015-0694-4 (accessed 18 March 2019)

23.   Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998 Jan 15;17(1):101–10. http://doi.wiley.com/10.1002/(SICI)1097-0258(19980115)17:1%3C101::AID-SIM727%3E3.0.CO;2-E (accessed 18 March 2019)

24.   Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An overview and Tutorial. *Tutor Quant Methods Psychol*. 2012;8(1):23–34.

25.   Harris P, Taylor R, Thielke R, et al. Research Electronic Data Capture (REDCap) - A metadata driven methodology and workflow process for providing translational research informatict support. *J Biomed Inform*. 2009;42(2):377–81. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2700030/ (accessed 18 March 2019)

26.   Mayan M. Essentials of Qualitative Inquiry. Morse J, editor. Qualitative Research. New York: Routledge; 2016: 93-98. https://www.taylorfrancis.com/books/9781315429243

(accessed 18 March 2019)

27. Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ*. 2004

Sep;38(9):1006–12. http://www.ncbi.nlm.nih.gov/pubmed/15327684 (accessed 18 March

2019)

28. 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(3):791–803.

29. 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 Heal*. 2015;18(2):315–33.

http://dx.doi.org/10.1016/j.jval.2014.12.006 (accessed 18 March 2019)

30. 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

Heal*. 2015;18(2):334–45. http://dx.doi.org/10.1016/j.jval.2015.01.004 (accessed 18

March 2019)

31. 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(5):630–45.

32. 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(7):878–86.

http://dx.doi.org/10.1097/01.chi.0000125092.35109.c5 (accessed 18 March 2019)

33. de Vet HCW, Terwee CB, Knol DL, et al. When to use agreement versus reliability

measures. *J Clin Epidemiol*. 2006;59(10):1033–9.

34.   Shweta, Bajpai RC, 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(3):20–7.

35.   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 Oct;161(1):126–30. http://linkinghub.elsevier.com/retrieve/pii/S0165178108000620 (accessed 18 March 2019)

36.   Goodyear-Smith F, Corter A, 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):1–8. http://dx.doi.org/10.1186/s12911-016-0379-z (accessed 18 March 2019)

37.   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 Dec 4;0(0):1–12. https://www.cambridge.org/core/product/identifier/S148180351700416X/type/journal_article (accessed 18 March 2019)

38.   Frosch E, DosReis S, Maloney K. Connections to outpatient mental health care of youths with repeat emergency department visits for psychiatric crises. *Psychiatr Serv*. 2011;62(6):646–9.

39.   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(1):85–94.

40.   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(6):632–8.
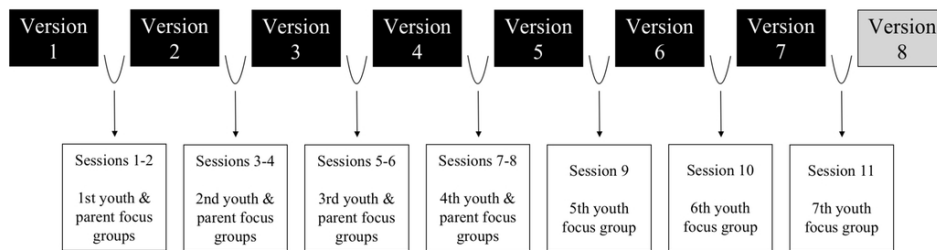
Figure legend

Figure 1. Schematic diagram showing the process of iterative modification that

MyHEARTSMAP underwent in phase one (1a) and phase two (1b), with corresponding tool

versions, sessions/rounds, and participants involved.

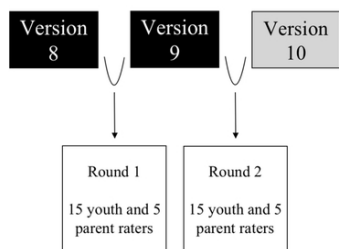Figure 2: Progression and transformation of MyHEARTSMAP's 'Mood' section, in accordance

with tool versions shown in figure 1.

92x69mm (300 x 300 DPI)

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