



ELSEVIER

Contents lists available at ScienceDirect

General Hospital Psychiatry

journal homepage: www.elsevier.com/locate/genhospsych

Development and pilot study of simple suicide risk rulers for use in the emergency department

Laurel O'Connor^{a,*}, Celine Larkin^b, Ameer F. Ibrahim^b, Michael Allen^c, Bo Wang^b,
Edwin D. Boudreaux^b

^a Department of Emergency Medicine, Boston Medical Center, One Boston Medical Center Place, Boston, MA 02118, United States of America

^b Department of Emergency Medicine, University of Massachusetts, 55 Lake Avenue North, Worcester, MA 01655, United States of America

^c Department of Psychiatry, University of Colorado at Denver, 500 - 13001 E. 17th Place, Bldg. 500, Level 2, Rm. E2308, Aurora, CO 80045, United States of America

ARTICLE INFO

Keywords:
Suicide
Screening
Emergency department
Risk

ABSTRACT

Background: Many patients treated in the emergency department (ED) for non-psychiatric complaints have elevated suicide risk. Universal screening can detect occult suicide risk, but gold standard risk measurement tools, such as the Beck Scale for Suicidal Ideation (BSS), are too long and cumbersome for ED use.

Objective: To test the performance of seven novel 0- to 10-point suicide risk “rulers” against the BSS.

Method: 399 patients from three EDs completed seven novel risk rulers, traditional binary screening items, and the BSS. Using BSS criterion references, we tested the diagnostic performance of each risk ruler and examined correlations between the rulers and BSS scores.

Results: By varying thresholds on the risk rulers, high levels of sensitivity and specificity were obtained. A threshold of 3 on the “sadness” ruler gave 89% sensitivity for the BSS criterion reference, and a threshold of 1 on the “wish to be dead” ruler provided 94–97% specificity.

Conclusion: Our novel risk rulers may be an efficient way to detect risk and triage potentially suicidal patients, showing good concurrent validity with the BSS. Clinicians can obtain high sensitivity and high specificity using just two rulers. Further research should examine the rulers' ability to predict independent clinician risk ratings and prospective suicidal behavior.

1. Introduction

Most adult patients treated in emergency departments (EDs) present with non-psychiatric chief complaints [1] and are not routinely evaluated for suicidal ideation or behavior [2]. Prior studies have reported that up to 8% of these non-psychiatric ED patients, however, do endorse active suicidal ideation when screened [2,3,4,5], and those who die by suicide are more likely to have made a non-psychiatric than psychiatric ED visit in the weeks before death [6]. Universal suicide screening has been successful in detecting ED patients with “hidden” suicide risk [7,8] and is a best practice recommended by the Joint Commission [9]. Despite its promise, universal suicide screening is still under-implemented, perhaps because it challenges the existing culture: the ED is a fast-paced and resource-limited environment [10] typically focused on rapid evaluation and provision of acute medical, rather than psychiatric, care [11]. To be acceptable to clinicians and administrators, evaluations and screening tools must be brief and easily administered by medical staff [10]. Gold standard tools for suicide evaluation used in research, such

as the Beck Scale for Suicidal Ideation [12], are not well-suited to the ED: they tend to be lengthy, require paper and pencil administration and clinician scoring, and are not easily memorized. As well as being brief, validated, and acceptable, a universal suicide risk-screening tool should accurately and efficiently distinguish patients with negligible risk from those with some suicide risk, who need further evaluation or intervention. The tool's utility and likely adoption would be strengthened if the tool could further stratify those with elevated risk, such as those with moderate risk, who require some level of intervention but do not need a full psychiatric evaluation during the ED visit, from those with high risk, who do require such a full psychiatric evaluation.

Brief suicide screeners for ED settings have been previously developed. The Patient Safety Screener (PSS) [13] is a three-item screening tool that assesses two-week presence of depression and active suicidal ideation, and any previous suicide attempt. This tool has strong agreement with the BSS [13]. The PSS-3 was implemented in the landmark Emergency Department Safety Assessment and Follow-up Evaluation (ED-SAFE) study: it was shown to be feasible to incorporate

* Corresponding author.

E-mail address: laurel.o'connor@bmc.org (L. O'Connor).

<https://doi.org/10.1016/j.genhospsych.2018.08.004>

Received 15 February 2018; Received in revised form 3 August 2018; Accepted 8 August 2018

0163-8343/ © 2018 Elsevier Inc. All rights reserved.

into routine clinical care and led to almost a doubling of risk detection [7]. Even with this success, full fidelity to the three-item screener was not achieved: up to one-third of patients with negative screens documented in their medical record reported to research assistants later in the ED visit that they had *not* been asked the suicidal ideation or behavior items [7]: it is possible that clinicians document negative screenings without administering the screener as intended. While brief measures such as the PSS are an improvement over longer measures like the BSS, challenges remain in terms of high-fidelity implementation in the ED setting.

During qualitative interviews with emergency nurses and physicians as part of the ED-SAFE study, clinicians identified a common impediment to suicide risk screening [7]. They expressed reluctance to ask about active suicidal ideation of all patients, especially those who were not there for a psychiatric chief complaint, and did not like having to ask multiple items. Additionally, they pointed out that a single item 0- to 10-point ruler, similar to the commonly used pain scale, might be a better fit with clinical practice [7]. In contrast to the inconsistent delivery of suicide risk screening, simple 0- to 10-point pain rulers have become widely used in clinical practice. Such a scale is easy to remember, administer, and explain, and it lends itself well to development of user-friendly visuals. Numerical rating scales for pain have been successfully applied in the ED setting [15,16]. Redesigning suicide risk screeners with the 0- to 10-point ruler conventions better align suicide screening with current ED practice and may therefore foster better fidelity by clinicians.

The objective of the current study was to develop seven newly created 0- to 10-point rulers and perform a pilot study of their operating characteristics to assess suicide risk in adults against criterion references from the BSS.

2. Methods

2.1. Setting and participants

The study was conducted in three EDs affiliated with a large urban academic medical center from December 2013 to March 2014. Consecutive patients 18 years or older were considered for enrollment, regardless of presenting complaint.

2.2. Procedure

Research assistants (RAs) approached potential subjects in the ED and screened them for inclusion and exclusion criteria. If a subject was eligible, the RA explained the study's risks and benefits. If a subject consented to participate, the RAs then instructed them to read and sign a written consent form. All participants were asked to complete seven 0- to 10-point rulers, binary screening questions, and the BSS. The full language of all rulers is shown in Table 4 and in Supplemental Fig. 1. The measures were completed as a written questionnaire unless patients requested to complete the questions verbally. The Institutional Review Board approved this study.

2.3. Measures

Based on the widely used form factor of the 0- to 10-point pain ruler, our team created seven rulers ranging from a global well-being item to items on sadness, hopelessness, wish to sleep and not wake up, wish to be dead, wish to kill self, and how suicidal the patient feels (Table 4 and Supplementary Fig. 1). All items have a 0- to 10-point scale with appropriate anchors and a time reference of the past 2 weeks, making the time reference the same as other commonly used screeners, like the Patient Health Questionnaire [17] and the PSS-3 [13]. This timeframe served to narrow the focus on recent experiences, which is most germane to the ED setting. Supplementary Fig. 1 depicts the rulers with their anchor: rulers were administered in the order shown. Each

ruler asks subjects to rate a single specific mood or behavior. They begin by requesting ratings about feelings of sadness and hopelessness, progressing to ratings of frank thoughts of wishing to die or wishing to kill themselves. The items were created with existing screeners and clinical practice in mind, with an emphasis on clinician acceptability. Simple terms that subjects who could read on an eighth-grade level could reliably recognize were chosen deliberately and the rulers were intended to increase in the directness with which they address suicidal ideations. The authors executed a pilot study of several rulers with the aim of identifying a smaller number of the best performing rulers for further assessment in a longitudinal study.

In addition to the rulers, patients completed several traditional binary "Yes/No" items drawn from the PSS-3 [13] and ASQ [14] assessing passive suicidal ideation ("At any time during the past 2 weeks, including today, have you wished you were dead?" and "At any time during the past 2 weeks, including today, have you felt that you or your family would be better off if you were dead?") and active suicidal ideation ("At any time during the past 2 weeks, including today, have you had any thoughts of killing yourself?"). The active ideation and lifetime attempt items were used to assess whether a person would meet the criteria for a positive screen on the PSS-3. These items were included to illustrate how they performed in predicting BSS criterion references and provide context for interpretation of the rulers' performance.

Finally, subjects were administered the BSS [12]. The BSS has been validated as a predictor of suicidal behavior in research [18] although it is not recommended for clinical use. It consists of 21 items, each with three descriptors related to suicidal ideation and past suicide attempts. The respondent reads the three descriptors for each item and chooses the one that best represents him/her. It serves as the gold standard criterion measure for the current study because of its widespread adoption, association with suicidal outcomes, and good psychometric properties [19]. In addition to a continuous score calculated by summing the 21 items, the BSS has several options to create categorical positive vs. negative suicide risk classifications. The first five items of the BSS measure passive ideation (lack of desire to save oneself in a life threatening situation) and active ideation (active desire to kill oneself). The definition of a positive screen on the BSS according to the manual is a positive response to items 4 or 5, which pertain to active suicidal ideation, a classification hereafter referred to as "BSS Active Ideation." Second, researchers have used items 4 and 5 along with the suicide attempt item to define a positive screen, such that a positive is anyone who screens positive according to the first criteria described above *and/or* who has a lifetime attempt, hereafter referred to as "BSS Active Ideation or Attempt." Finally, an empirical cut-off of > 2 on the BSS has been established through prediction of future suicidal behavior [13], hereafter referred to as "BSS > 2 ." We dichotomized our sample using these three criterion references (BSS Active Ideation, BSS Active Ideation or Attempt, and BSS > 2) and report the performance of our novel rulers against each classification. The BSS appeared to have good internal consistency in the current sample, with a Cronbach's alpha in the current study of $\alpha = 0.93$: this echoes the good psychometric properties observed in previous studies [18,19].

Research assistants also recorded patients' sex, age, race, ethnicity, and the nature of their chief complaints.

2.4. Data analyses

We conducted power analysis using PASS 15. A sample of 32 from the positive group and 371 from the negative group achieves 82% power to detect a difference of 0.1500 between the area under the ROC curve (AUC) under the null hypothesis of 0.65 and an AUC under the alternative hypothesis of 0.80 using a two-sided z-test at a significance level of 0.05.

Descriptive statistics were calculated on all measures. Kolmogorov-Smirnov tests of normality were conducted to assess the distribution of

the ruler scores. We calculated receiver operating characteristics (ROC), sensitivity, and specificity of each of the rulers for the three criterion references described above. An optimized score was calculated for each criterion reference for each ruler, defined as the score that achieves overall best sensitivity and specificity for a given outcome. This was accomplished by maximizing Youden's index (Calculated as Maximum = Sensitivity + Specificity - 1). The area underneath the ROC (AUC) curve of each ruler represented the ruler's discrimination [20]. For each optimized score, the positive and negative predictive values for each of the BSS screening definitions were calculated.

For each ruler, the threshold that afforded 95% sensitivity to predict the criterion and the threshold that afforded 95% specificity were identified. This, in effect, allowed creation of two thresholds for each ruler that could be used to divide respondents into three groups for each of the three criterion references. A "lower scoring" group for a criterion reference (for example BSS > 2) comprises of those scoring below the 95% sensitivity threshold (i.e., highly confident that the subject would have a BSS score ≤ 2). Taking into account other aspects of the clinical situation, a clinician could postulate that a patient who reported a ruler score highly predictive of a BSS score ≤ 2 may not require further suicide assessment. The "high scoring" group is defined as those scoring above the threshold defining the 95% specificity threshold (i.e., highly confident that the subject would have a BSS score > 2), representing patients that are very likely to need additional clinical evaluation for suicide risk. Such a determination could be used to help the treating clinician determine that more extensive suicide screening is needed. Patients who fall between two thresholds can be considered moderate risk of scoring a BSS > 2. Pearson correlations between each ruler score and a total sum score of all of the rulers, and the total BSS score were computed to assess their association as continuous scores.

Finally, to examine the performance of the PSS-3 and traditional binary screening items for context, their associations with the BSS criterion references were examined by calculating sensitivity, specificity, and positive and negative predictive values for each of the three BSS criterion references.

3. Results

3.1. Descriptive statistics

Of the 872 patients approached by RAs, 472 (54%) declined to participate. The remaining 400 subjects (45% of those actually approached) were administered the survey. One subject was subsequently withdrawn from the study per her own request, leaving 399 subjects included in analysis. Of all subjects, 361 (90.5%) had a chief complaint that was medical and 38 (9.5%) of whom had a chief complaint that was psychiatric. Prior suicide attempt was endorsed by 14.3% of the sample; 37 (10.4%) in the medical chief complaint cohort and 20 (52.6%) in the psychiatric chief complaint cohort. Table 1 describes the sample. Kolmogorov-Smirnov tests of normality indicated that none of the rulers were normally distributed, so the median and inter-quartile range for each ruler and the ruler total are presented.

Most patients (n = 298; 74.7%) completed the assessment in written form while a minority of patients chose to complete it as an interview with the RA (n = 101; 25.3%). Mann Whitney U tests showed that median scores on the rulers did not differ significantly between those who gave written versus verbal responses. Chi square analysis also showed no difference between written and verbal responses in terms of the proportion of patients screening positive on the three BSS criterion references.

3.2. Performance of rulers and traditional binary screening items

Although the rulers were not designed to operate as a scale, they demonstrated good internal consistency: Cronbach's alpha was 0.86 across the seven rulers, and increased to 0.91 when the first ruler ("How

Table 1
Descriptive statistics of the sample (n = 399).

| | N | % |
|--|--------|-------|
| Female | 202 | 50.6 |
| White | 342 | 85.7 |
| Black/African American | 35 | 8.8 |
| Other/not documented | 26 | 11.5 |
| Hispanic ethnicity | 48 | 12.0 |
| Presenting complaint involved psychiatric behavior | 38 | 9.5 |
| Method of survey administration | | |
| Interview by RA | 101 | 25.3 |
| Paper and pencil | 298 | 74.7 |
| Binary items | | |
| Wished were dead (Yes) ^a | 38 | 9.5 |
| Felt family would be better off if patient was dead (Yes) ^a | 28 | 7.0 |
| Any thoughts of killing yourself (Yes) ^a | 24 | 6.0 |
| Lifetime suicide attempt (Yes) | 57 | 14.3 |
| Lifetime suicide attempt (Yes) medical chief complaint | 37 | 10.4 |
| Lifetime suicide attempt (Yes) psychiatric chief complaint | 20 | 52.6% |
| Suicide attempt in past year (Yes) | 9 | 2.3 |
| BSS active ideation (BSS 4 or 5 > 0) | 27 | 7.4 |
| BSS active ideation or attempt (BSS 4 or 5 > 0 or lifetime attempt) | 64 | 16.8 |
| BSS > 2 (BSS sum > 2) | 33 | 8.6 |
| | Median | (IQR) |
| Age | 45 | 29-59 |
| On a scale from 1 to 10... | | |
| ...how has your life been going? (n = 397) | 7 | 5-9 |
| ... how sad have you been over the past two weeks? (n = 393) | 2 | 0-5 |
| ...how hopeless have you been over the past two weeks? (n = 393) | 0 | 0-4 |
| ...how much have you wished you could go to sleep and not wake up during the past two weeks? (n = 395) | 0 | 0-0 |
| ... have you wished you were dead over the past two weeks? (n = 395) | 0 | 0-0 |
| ...how much have you thought about killing yourself over the past two weeks? (n = 396) | 0 | 0-0 |
| ...how suicidal have you been over the past two weeks? (n = 395) | 0 | 0-0 |
| Total BSS Score | 0 | 0-0 |

^a Time reference past 2 weeks.

has your life been going?") was excluded. Table 2 presents the rulers' sensitivities, specificities, optimized scores, AUC (determinations), and the scores for each ruler required to achieve 95% sensitivity and 95% specificity on the three criterion references (BSS Active Ideation, BSS Active Ideation or Attempt, and BSS > 2 respectively). The ROC curves for each of the three criterion references are illustrated in Supplementary Figs. 2, 3, and 4.

At their optimized scores, the sensitivity values of the rulers assessing "sadness," "hopelessness," "wishing to go to sleep and not wake up," and "wishing to be dead" were higher than the PSS-3 positive screen across each of the criterion references (Table 3). While the PSS-3 positive screen had higher specificity in general than the rulers (Table 3), the highest performing rulers still achieved specificities of > 95%.

The rulers with the highest AUCs were those assessing "thoughts of killing oneself," "thoughts of suicide," "sadness," "wishing to go to sleep and not wake up," "hopelessness," and "wishing to be dead," regardless of which BSS criterion reference they were tested against.

For the criterion reference of BSS Active Ideation (Table 2), the rulers assessing hopelessness and sadness had unique potential as stratifying questions: their sensitivity was > 95% for a score threshold of > 2 and 1 respectively and their specificity was > 95% for a score threshold of > 8. "Thoughts of killing oneself" performed similarly with score thresholds of 2 and 8 respectively. The rulers assessing sadness and hopelessness performed in a similar manner for the criterion reference of BSS > 2.

Table 2
Diagnostic values of rulers in predicting three criterion references from BSS.^a

| | Optimized single score (sensitivity, specificity) | Optimized paired scores for 95% sensitivity, 95% specificity | PPV for optimized score | NPV for optimized score | Area under ROC |
|---|--|---|----------------------------|----------------------------|-------------------|
| Criterion reference 1: "BSS Active Ideation" (BSS 4 or 5 > 0) (n = 27 positive, 336 negative) | | | | | |
| On a scale from 0 to 10... ^b | | | | | |
| ...how has your life been going? | 1 (0.64, 0.03) | N/A, N/A | 0.05 | 0.52 | 0.13 |
| ... how sad have you been? | 3 (0.89, 0.67) | 1, 8 | 0.18 | 0.99 | 0.88 |
| ...how hopeless have you been? | 2 (0.96, 0.74) | 2, 8 | 0.23 | 0.99 | 0.9 |
| ...how much have you wished you could go to sleep and not wake up? | 1 (0.86, 0.87) | N/A, 5 | 0.35 | 0.99 | 0.89 |
| ... have you wished you were dead? | 1 (0.82, 0.94) | 1, 2 | 0.52 | 0.99 | 0.96 |
| ...how much have you thought about killing yourself? | 1 (0.71, 0.98) | 2, 8 | 0.71 | 0.98 | 0.84 |
| ...how suicidal have you been? | 1 (0.57, 0.98) | N/A, 1 | 0.67 | 0.97 | 0.86 |
| Criterion reference 2: "BSS Active Ideation or Attempt" (n = 64 positive, 315 negative) | | | | | |
| On a scale from 1 to 10... ^b | | | | | |
| ...how has your life been going? | 1 (0.75, 0.02) | N/A, N/A | 0.14 | 0.3 | 0.22 |
| ... how sad have you been? | 3 (0.80, 0.71) | N/A, 8 | 0.37 | 0.95 | 0.82 |
| ...how hopeless have you been? | 2 (0.74, 0.77) | N/A, 8 | 0.4 | 0.94 | 0.78 |
| ...how much have you wished you could go to sleep and not wake up? | 1 (0.65, 0.90) | N/A, 3 | 0.58 | 0.93 | 0.8 |
| ... have you wished you were dead? | 1 (0.62, 0.97) | N/A, 1 | 0.82 | 0.92 | 0.85 |
| ...how much have you thought about killing yourself? | 1 (0.43, 0.98) | N/A, 1 | 0.85 | 0.98 | 0.73 |
| ...how suicidal have you been? | 1 (0.38, 0.99) | N/A, 1 | 0.86 | 0.89 | 0.73 |
| Criterion reference 3: "BSS > 2" (BSS sum score of > 2) (n = 33 positive, 347 negative) | | | | | |
| On a scale from 0 to 10... ^b | | | | | |
| ...how has your life been going? | 1 (0.63, 0.03) | N/A, N/A | 0.06 | 0.46 | 0.1 |
| ... how sad have you been? | 4 (0.71, 0.78) | 3, 8 | 0.4 | 0.93 | 0.91 |
| ...how hopeless have you been? | 3 (0.85, 0.74) | 1, 8 | 0.23 | 0.98 | 0.91 |
| ...how much have you wished you could go to sleep and not wake up? | 2 (0.89, 0.78) | N/A, 4 | 0.28 | 0.97 | 0.91 |
| ... have you wished you were dead? | 1 (0.89, 0.94) | N/A, 2 | 0.6 | 0.99 | 0.98 |
| ...how much have you thought about killing yourself? | 1 (0.74, 0.98) | N/A, 1 | 0.76 | 0.98 | 0.87 |
| ...how suicidal have you been? | 1 (0.63, 0.98) | N/A, 1 | 0.76 | 0.96 | 0.87 |

^a A reported score greater than and not including the first value defines 95% sensitivity, and a reported score less than and not including the second value, defines 95% specificity. For example, "how hopeless have you been," is interpreted as such: values of > 1 but not including 1 are at least 95% sensitive for the criterion (BSS Active Ideation positive) and values > 8 are at least 95% specific for the criterion (BSS Active Ideation positive).

^b Time reference past 2 weeks.

Table 3
Diagnostic values of binary items and Patient Safety Screener-3 positive screen in predicting three criterion references from BSS.

| | Sensitivity, specificity | PPV | NPV |
|---|--------------------------|------|------|
| Criterion reference 1: BSS Active Ideation (n = 27 positive, 336 negative) | | | |
| Wished were dead | 0.71, 0.96 | 0.65 | 0.98 |
| Felt family would be better off if patient was dead | 0.61, 0.97 | 0.67 | 0.97 |
| Any thoughts of killing self | 0.57, 0.99 | 0.8 | 0.97 |
| Lifetime attempt | 0.59, 0.90 | 0.31 | 0.97 |
| Positive PSS-3 | 0.60, 0.99 | 0.8 | 0.97 |
| Criterion reference 2: BSS Active Ideation or Attempt (n = 64 positive, 315 negative) | | | |
| Wished were dead | 0.49, 0.98 | 0.86 | 0.91 |
| Felt family would be better off if patient was dead | 0.39, 0.99 | 0.89 | 0.89 |
| Any thoughts of killing self | 0.31, 0.99 | 0.96 | 0.88 |
| Lifetime attempt | 0.83, 0.99 | 0.95 | 0.97 |
| Positive PSS-3 | 0.37, 0.99 | 0.96 | 0.89 |
| Criterion reference 3: BSS > (n = 33 positive, 347 negative) | | | |
| Wished were dead | 0.79, 0.96 | 0.71 | 0.98 |
| Felt family would be better off if patient was dead | 0.58, 0.97 | 0.64 | 0.96 |
| Any thoughts of killing self | 0.60, 0.99 | 0.88 | 0.96 |
| Lifetime attempt | 0.62, 0.90 | 0.37 | 0.96 |
| Positive PSS-3 | 0.65, 0.99 | 0.71 | 0.98 |

3.3. Continuous scores on rulers

All Pearson correlations between individual and total ruler scores and the total BSS score were statistically significant (Table 4). Four

correlations with total BSS score were strong to very strong: "wishing to be dead" ($r = 0.78$), "thoughts of killing self" ($r = 0.81$), "suicidal" ($r = 0.87$) and total ruler score ($r = 0.76$).

4. Discussion

4.1. Ruler performance

In this pilot study, ED patients' responses to 0- to 10-point rulers assessing several suicide-related constructs, from sadness to active suicidal ideation, had good concurrent validity with the longer gold standard Beck Scale for Suicidal Ideation. This is particularly important because ED patients, even those presenting with non-psychiatric complaints, are more likely to have elevated suicide risk than the general population [2] but universal screening for suicide in patients with non-psychiatric presentations is rarely performed.

4.2. Screening for suicide in the ED

There remains a significant need for the validation of brief universal assessment tools for suicidal thought and behaviors in the emergency department; indeed, the American College of Emergency Physicians recently identified a need for better screening tools as an essential barrier to universal suicide screening [24]. Although the Joint Commission's Sentinel Alert 56 recommends the use of universal suicide screening in EDs and there have been some new, promising screeners described in the literature such as the PSS-3 [13], the clinical

Table 4
Pearson correlations between BSS total sum score and each of the rulers and the total ruler score.

| | BSS total sum score (r) | p |
|---|-------------------------|---------|
| On a scale from 1 to 10... | | |
| ... how has your life been going over the past two weeks, with 0 being Terrible and 10 being Terrific? | −0.42 | < 0.001 |
| ... how sad have you been over the past two weeks, with 0 being Not at All and 10 being Extremely? | 0.45 | < 0.001 |
| ... how hopeless have you been over the last two weeks, with 0 being Not at All and 10 being Extremely? | 0.46 | < 0.001 |
| ... how much have you wished you could go to sleep and not wake up over the past two weeks, with 0 being Not at All and 10 being All the Time | 0.67 | < 0.001 |
| ... how much have you wished you were dead over the past two weeks, with 0 being Not at All and 10 being All the Time? | 0.78 | < 0.001 |
| ... how much have you thought about killing yourself over the past two weeks, with 0 being Not at All and 10 being All the Time? | 0.81 | < 0.001 |
| ... how suicidal have you been over the past two weeks with 0 being Not at All and 10 being Extremely suicidal? | 0.87 | < 0.001 |
| Sum across seven rulers | 0.76 | < 0.001 |

acceptability of universal screening continues to be hampered by a multifaceted set of barriers. In part, screening may be hindered by perceived shortage of acceptable screening tools [21]. An ED-based suicide screening tool should be brief, easy for a health care provider or ancillary staff member to administer in a standardized fashion, easy to understand, and predictive of suicide risk with both high sensitivity and specificity. Many existing tools, such as the Manchester Self-Harm Rule [22] and ReACT rule [23] have high sensitivity but poor specificity, and apply only to those presenting with a current suicide attempt.

In addition, some clinicians and hospital administrators balk at using tools with frank suicide-related content as the first step of screening, such as asking all individuals if they have had thoughts of killing themselves in the past 2 weeks, as happens with the PSS-3 [13]. They might prefer to build a branched screening process using the sadness or hopelessness rulers we created. Asking suicide-specific questions without a “softer” segue (for example, an item about emotional health) or in the absence of clinically relevant context (such as presenting with overt psychiatric symptoms) can be perceived as abrupt or inappropriate. It has been shown that clinicians may skip suicide-specific questions in real-world implementation [7].

One benefit of a ruler-based approach to screening is that it may help to avoid biased framing of binary suicide items by providers, characterized by phrasing like “you haven't been feeling suicidal, have you?”; in primary care, such negative framing has been shown to increase the risk that patients will deny suicidality [24]. Additionally, these rulers may reduce the patient experience of stigma associated with having to provide a binary answer endorsing suicidality. If a visual aid were also used, it could help to provide increased privacy in a typically crowded ED. Because these new rulers operate on a continuum but yet are in a format that is familiar to non-behavioral health clinicians, they hold significant promise as nuanced but disseminable tools to assess suicide risk. A theoretical use of the rulers, after appropriate validation study, is to use them in a stepwise manner by administering an initial high-sensitivity ruler to all patients. Those who scored below a threshold of appropriate sensitivity may be considered “low risk”, not requiring additional screening; the remainder of the population would be administered a second, high specificity ruler, further risk-stratifying patients into a moderate “and” high risk populations, thus flagging them for different degrees of provider intervention. Such a step-wise model would require significant prospective validation prior to deployment to aid in clinical decision-making.

It should be noted the purpose of these ruler-based screening tools is not to replace a full psychiatric evaluation; rather, they are intended to be used by frontline health care providers who are not mental health providers to guide whether patients who screen positive require more intensive evaluation by an emergency mental health professional prior to being discharged from a health care setting. Clearly, other factors help make this determination, such as the presenting complaint, and the rulers would be a complement to, not a replacement of, these other clinical factors.

4.3. Limitations

Its innovation notwithstanding, there were several limitations in this study, including a small sample size and a lack of longitudinal follow up to determine prospective suicidal behavior. The enrollment rate was under 50%, which introduces the possibility of selection bias to the study. The rulers were compared against the BSS, and should be tested against other criterion references predictive of suicidal behavior, such as an independent administration of the Columbia Suicide Severity Rating Scale [25]. It is possible that the concurrence found across the measures tested may be in part because subjects' memory of previously answered questions impacted their responses later in the questionnaire. Additionally, subjects knew that the study was a mental health assessment, and were informed prior to taking the survey that a positive screen might result in further evaluation. This may have led patients be less forthcoming about their symptoms. This is a common issue with self-report tools. Enrollment was limited to adults who were medically, emotionally, and cognitively healthy enough to participate.

4.4. Future research and conclusions

Given that our rulers have demonstrated concurrent validity here in pilot study against several BSS benchmarks in a cross-sectional design, future projects should be designed to validate the rulers against other suicide risk evaluations, comprehensive clinician assessment and future suicidal behavior. Such a study would require a prospective design with high enrollment rate and rigorous fidelity to the instrument. Additionally, further study must be performed to determine if two of the rulers used together may have a potentiating effect and whether this affects the sensitivity and specificity of the screen. Finally, the feasibility and ease of implementing and interpreting the ruler as a screening tool should be assessed with providers and patients. An important additional consideration would be whether or not practitioners are content to use a positive screen on the rulers to justify a mandatory psychiatric evaluation: empirical support of the tools' predictive utility would help to inform this issue.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.genhosppsych.2018.08.004>.

References

- [1] Skinner H, Blanchard J, Elixhauser A. Trends in Emergency Department Visits, 2006–2011. HCUP Statistical Brief #179 Agency for Healthcare Research and

- Quality; 2014.
- [2] Claassen C, Larkin G. Occult suicidality in an emergency department population. *Br J Psychiatry* 2005;186:352-3.
- [3] Allen, et al. Screening for suicidal ideation and attempts among emergency department medical patients: instrument and results from the psychiatric emergency research collaboration. *Suicide Life Threat Behav* 2013;43(3):313-23.
- [4] Ballard, et al. Recent medical service utilization and health conditions associated with a history of suicide attempts. *Gen Hosp Psychiatry* 2014;36(4):437-41.
- [5] Ilgen, et al. Recent suicidal ideation among patients in an inner city emergency department. *Suicide Life Threat Behav* 2009;39(5):508-17.
- [6] Ahmedani, et al. Health care contacts in the year before suicide death. *J Gen Intern Med* 2014;29(6):870-7.
- [7] Boudreaux, et al. Improving suicide risk screening and detection in the emergency department. *Am J Prev Med* 2016;50(4):445-53.
- [8] Roaten, et al. Development and implementation of a universal suicide risk screening program in a safety-net hospital system. *Jt Comm J Qual Patient Saf* 2018;44(1):4-11.
- [9] The Joint Commission. Sentinel event alert 56: Detecting and treating suicide ideation in all settings. Joint Commission Resources; 2016.
- [10] Nugus P, Holdgate A, Fry M, Forero R, McCarthy S, Braithwaite J. Work pressure and patient flow management in the emergency department: findings from an ethnographic study. *J Acad Emerg Med* 2011;18(10):1045-52.
- [11] Boudreaux E, Horowitz L. Suicide risk screening and assessment: designing instruments with dissemination in mind. *Am J Prev Med* 2014;47(3 Suppl 2):163-9.
- [12] Beck, Kovacs M, Weissman A. Assessment of suicidal intention: the scale for suicide ideation. *J Consult Clin Psychol* 1979;47(2):343-52.
- [13] Boudreaux, et al. The patient safety screener: validation of a brief suicide risk screener for emergency department settings. *Arch Suicide Res* 2015;19(2):151-60.
- [14] Horowitz, et al. Ask suicide-screening questions (ASQ): a brief instrument for the pediatric emergency department. *Arch Pediatr Adolesc Med* 2012;166(12):1170-6.
- [15] Bijur P, Latimer C, Gallagher E. Validation of a verbally administered numerical rating scale of acute pain for use in the emergency department. *Acad Emerg Med* 2003;10(4):390-2.
- [16] Mohan H, Ryan J, Whelan B, Wakai A. The end of the line? The visual analogue scale and verbal numerical rating scale as pain assessment tools in the emergency department. *Emerg Med J* 2010;27(5):372-5.
- [17] Martin A, Rief W, Klaiberg A, Brachler E. Validity of the brief patient health questionnaire mood scale (PHQ-9) in the general population. *Gen Hosp Psychiatry* 2006;28(1):71-7.
- [18] Lillian M, Range E, Knott C. Twenty suicide assessment instruments: evaluation and recommendations. *Death Stud* 2010;21(1):25-58.
- [19] Beck A, Brown G, Steet R, Dahlsgaard K, Grisham J. Suicide ideation at its worst point: a predictor of eventual suicide in psychiatric outpatients. *Suicide Life Threat Behav* 1999;29:1-9.
- [20] Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med* 2013;4(2):627-35.
- [21] Schneider S. Is universal suicide screening in the emergency department saving lives or wasting time? *ACEP Now* 2018;37(1). [Epub].
- [22] Steeg, et al. The development of a population-level clinical screening tool for self-harm repetition and suicide: the ReACT self-harm rule. *Psychol Med* 2012;42(11):2383-94.
- [23] Randall J, Colman I, Rowe B. A systematic review of psychometric assessment of self-harm risk in the emergency department. *J Affect Disord* 2011;134(1-3):348-55.
- [24] McCabe R, Sterno I, Priebe S, Barnes R, Byng R. How do healthcare professionals interview patients to assess suicide risk? *BMC Psychiatry* 2017;17(1):122.
- [25] Posner, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry* 2011;168(12):1266-77.