Suicide Risk Screening and Assessment

Designing Instruments with Dissemination in Mind

Edwin D. Boudreaux, PhD, Lisa M. Horowitz, PhD, MPH

This paper summarizes recommendations made regarding the National Action Alliance for Suicide Prevention Research Prioritization Task Force's Aspirational Goal 2, to "determine the degree of suicide risk (e.g., imminent, near-term, long-term) among individuals in diverse populations and in diverse settings through feasible and effective screening and assessment approaches." We recommend that researchers shift to using "design for dissemination" principles to maximize both the goodness of fit and validity of screening and assessment measures for a given setting. Three specific recommendations to guide research efforts are made to achieve this shift: (1) the parameters related to each setting, including the logistics, scope of practice, infrastructure, and decision making required, should be identified and used to choose or design screening and assessment instruments that have a good fit; (2) to the greatest feasible extent, technology should be used to support screening and assessment; and (3) researchers should study the best methods for translating validated instruments into routine clinical practice. We discuss the potential barriers to implementing these recommendations and illustrate the paradigm shift within the emergency department setting.

(Am J Prev Med 2014;47(3S2):S163–S169) © 2014 American Journal of Preventive Medicine. All rights reserved.

Introduction

The National Action Alliance for Suicide Prevention (Action Alliance) Research Prioritization Task Force's (RPTF's) Aspirational Goal 2 (AG2) seeks to outline a research pathway that will lead to the development of validated procedures that can "determine the degree of suicide risk (e.g., imminent, near-term, long-term) among individuals in diverse populations and in diverse settings through feasible and effective screening and assessment approaches" (p. 24).[1]

This paper reviews the specific assertions that underlie AG2, proposes a paradigm shift for screener and assessment development and research, and outlines three specific recommendations to actualize the new paradigm. AG2 is meant to apply to all settings, and although our recommendations can apply to schools, detention settings, and the armed forces, this paper focuses on adult patients in healthcare settings.

From the Department of Emergency Medicine, Department of Psychiatry, and Department of Quantitative Health Sciences (Boudreaux), University of Massachusetts Medical School, Worcester, Massachusetts; and the Intramural Research Program (Horowitz), National institute of Mental Health, NIH, Bethesda, Maryland

Address correspondence to: Edwin D. Boudreaux, PhD, Department of Emergency Medicine, LA-189, 55 Lake Avenue North, Worcester MA 01655. E-mail: edwin.boudreaux@umassmed.edu.

0749-3797/$36.00

http://dx.doi.org/10.1016/j.amepre.2014.06.005

Aspirational Goal 2 Assertions

AG2 can be broken down into several key assertions briefly reviewed below, which will be expanded upon in relation to our specific recommendations in the sections that follow. In addition, some components, such as the definition and measurement of imminent risk, are discussed in detail by other papers appearing in this supplement.

Suicide risk can be operationalized along a timeline of imminent, near-term, and long-term risk. Although a coherent system that operationalizes these terms does not yet exist, and is a key component of the pathway that the Action Alliance is attempting to elucidate, most suicidologists acknowledge that suicide risk is a varying trait that is not stable over time. As embodied in the American Psychiatric Association's suicide assessment and treatment guidelines,[2] suicide risk can be formulated as an interaction between relatively stable risk factors or predisposing characteristics, protective factors, and acute precipitants.

This conceptualization of suicide risk promotes the logical conclusion that an individual should be screened and assessed in reference to a specific risk horizon. All or most settings must attend to imminent risk, because it is critical for deploying suicide prevention efforts that require immediate action. However, some settings also have the capacity to focus on assessing and managing long-term risk, which will strongly influence the

© 2014 American Journal of Preventive Medicine. All rights reserved.

screening and assessment instruments' composition, as well as clinical decision making guided by these instruments and the research methods used to study them.

Individuals in diverse populations and settings may require different approaches. Approaches that work for adults in primary care may not be the same as approaches that work for children in educational settings. Each population and setting must be evaluated individually and a good fit approach should be built with stakeholder involvement.

Screening and assessment are different. Screening is performed to detect whether any actionable risk is present, or put differently, to screen out those with negligible risk. As such, it requires easy administration by front-line staff, should be highly sensitive, and should have a strong ability to confidently rule out patients with no appreciable risk (i.e., low false negatives).[3,4] Assessment, in contrast, is a more in-depth evaluation performed to further quantify the severity of risk to guide further clinical action. Assessments should have strong specificity and be able to identify individuals who are at true risk and need immediate or increased resources and support. Ideally, screening and assessment should work in a coordinated fashion, with screening sensitively detecting any clinically actionable risk and assessment specifying risk on a severity continuum.

The approaches used to screen and assess suicide risk must balance feasibility and effectiveness. In many settings, tension exists between feasibility—or what is most efficiently performed by providers—and effectiveness—or what is most valid in identifying and quantifying risk. For example, the most feasible path may be to screen only patients with frank psychiatric symptoms, which contrasts with what may be the most valid and effective path for identifying risk among the population, such as universal screening.

A Paradigm Shift

The prevailing paradigm guiding suicide risk research has been characterized by mental health specialists creating multi-item instruments and testing them under research conditions to determine if they predict future suicidal behavior. These instruments are often created independent of the specific clinical decisions they will be guiding and without full consideration of the parameters that would be relevant to whether the instrument could be applied under routine "real-world" conditions. Often, the AG2 assertions are not sufficiently considered.

A classic example is the Beck Scale for Suicide Ideation (BSSI).[5] Although this scale is a good fit for mental health settings, where multi-item self-reported paper-and-pencil instruments are commonplace, it can be a poor fit for other settings. For example, most clinicians in adult general medical settings, such as primary care practices, are ill equipped to administer, score, and interpret such instruments. In contrast, as a general rule of thumb, the behavioral health screeners that have fared best in general medical settings are ultra-short and easily memorized.

This efficiency principle has recently been acknowledged by an expert panel convened by the NIH.[6] The ten behavioral health screeners they recommend being integrated into electronic health record systems (EHRs), including those assessing tobacco, alcohol, depression, and other behavioral health domains, consist of no more than three items for each domain. The end result of the prevailing paradigm of suicide risk research is that we are no closer to having a validated, practical screening and assessment approach across most settings, and suicidal individuals continue to be undetected by "front-line" personnel, such as physicians, nurses, teachers, counselors, and detention facility staff.

For example, studies across a range of settings, from schools to emergency departments (EDs) to primary care, indicate that suicide risk screening is simply not being done in any systematic, universal fashion.[3] In addition, because there is a dearth of published, validated screeners, it is likely that suicide screening, when it does occur, is performed in an idiosyncratic, non-standardized manner using questions with unproven reliability or validity.

We propose a new paradigm to guide suicide risk screening and assessment research across diverse settings and populations. Namely, screening and assessment approaches should be selected or designed with dissemination in mind. This means that the screener and assessment should be selected or developed from the ground floor to be tailored to the individual needs of the setting or population with which they are to be used. This shift parallels the work of others who have emphasized that target population characteristics, provider characteristics, and setting demands are important when designing and deploying interventions.[7]

Researchers should actively consider key design parameters inherent to each setting and population at every step of development, from item construction to prospective validation to studying translation into routine use. This paradigm shift should help the field to break out of answering the question "Does instrument 'A' predict suicide attempts at some point in the future?" and reconnect with complex, real-world decision making that is more nuanced than this bivariate perspective.

The goals of standardizing the screening process and developing screening approaches with ecological validity may appear at odds with one another; however, both goals are achievable. Once an ecologically valid approach is developed and validated for a given setting, it can be translated to individual locations where it becomes a standardized protocol. This translation may require local adaptations aimed at improving standardization of the process by addressing local barriers.

This entire translation cycle of developing approaches that are ecologically valid, adapting them to standardized local protocols, and studying the impact of these adaptations can inform recommendations for blending standardization and adaptation at the local level. Below, we review three specific recommendations that will help to operationalize this paradigm shift and guide future research endeavors.

Recommendations

The parameters related to each individual setting, including the logistics, scope of practice, infrastructure, and decision making required, should be identified and used to choose or design screening and assessment instruments that have a good fit. The ultimate purpose of screening for and assessing suicide risk in any setting is to detect when people are at any non-zero risk and then gauge the degree of risk present to guide decision making; however, settings differ dramatically in the kinds of decisions that must be guided by this process, the resources available to manage positive screens, and the protocols that must be followed. Researchers need to carefully consider these factors when designing suicide risk instruments for each setting.

As mentioned earlier, screening and assessment are not the same, and the instruments and protocols employed should be considered as separate but interrelated processes. To use an analogy from the depression literature, the Patient Health Questionnaire (PHQ)-2,[8] a two-item screener for depression, has a weak 38% positive predictive value for diagnosing major depressive disorder; nevertheless, it remains one of the most widely used quick screening instruments for depression in medical settings because it is useful for identifying when further action, such as additional assessment, is warranted.

As introduced earlier, a very important consideration is the risk timeline most pertinent to the setting under consideration. Although it is optimal from a public health perspective to detect and manage lifetime risk, many settings will be focused on detecting and managing imminent or short-term risk. For example, the primary care setting is generally focused on long-term care; the screener and assessment in this setting should not only identify those at imminent risk who need urgent intervention, such as transport to an ED, but also identify those at long-term risk who need more chronic interventions, such as continued monitoring, more frequent visits, and psychotropic medication.

Critical features of the screening and assessment measures will be impacted by such a tailored risk horizon approach, including the nature of the questions asked, the length of the screening or assessment, who is responsible for the screening and assessment, how often the questions are asked, and the kinds of actions that will be taken if the individual screens positive.

Each setting and population will have practical logistics that are important in determining good fit. Researchers should consider these early in the design process. Instruments should be designed for the setting instead of expecting the setting to adapt to the instrument. For example, in many medical settings, providers have limited time to interact with an individual patient and do not have "props" like paper-based forms to help them remember the questions. Consequently, the questions used for primary suicide risk screening in these settings will have to be simple, quickly administered, easily memorized, and use a yes/no response format if they are to be applied with fidelity.

In addition, some frontline personnel may be reluctant to screen for suicide risk because they are uncertain of how to handle positive screens. This "Pandora's box" phenomenon is exacerbated by the lack of evidence-based interventions readily available for most settings that do not specialize in mental health treatment. This reluctance can be minimized by establishing clear protocols for further assessing and managing suicide risk for the specific setting, making sure all staff members are trained, and providing supports or props to help navigate the next actions to take once actionable suicide risk is detected.

Recently, researchers have decried the lack of suicide risk instrument validation across settings and populations.[9,10] Newly designed instruments will need to be rigorously validated. The shift in focus on aligning screening and assessment with decision making appropriate for the setting has important implications for the validation process. When establishing the operating characteristics of an instrument, the criterion reference against which a screener will be validated should be different from the reference against which a full risk assessment is validated.

More specifically, the criterion reference for screening does not need to be suicidal behavior; rather, it can be clinical judgment that actionable risk is present, that is, enough risk is present that some minimal clinical action is necessary, such as additional assessment or referral to mental health care. In contrast, the criterion reference for the risk assessment should be suicidal behavior or other

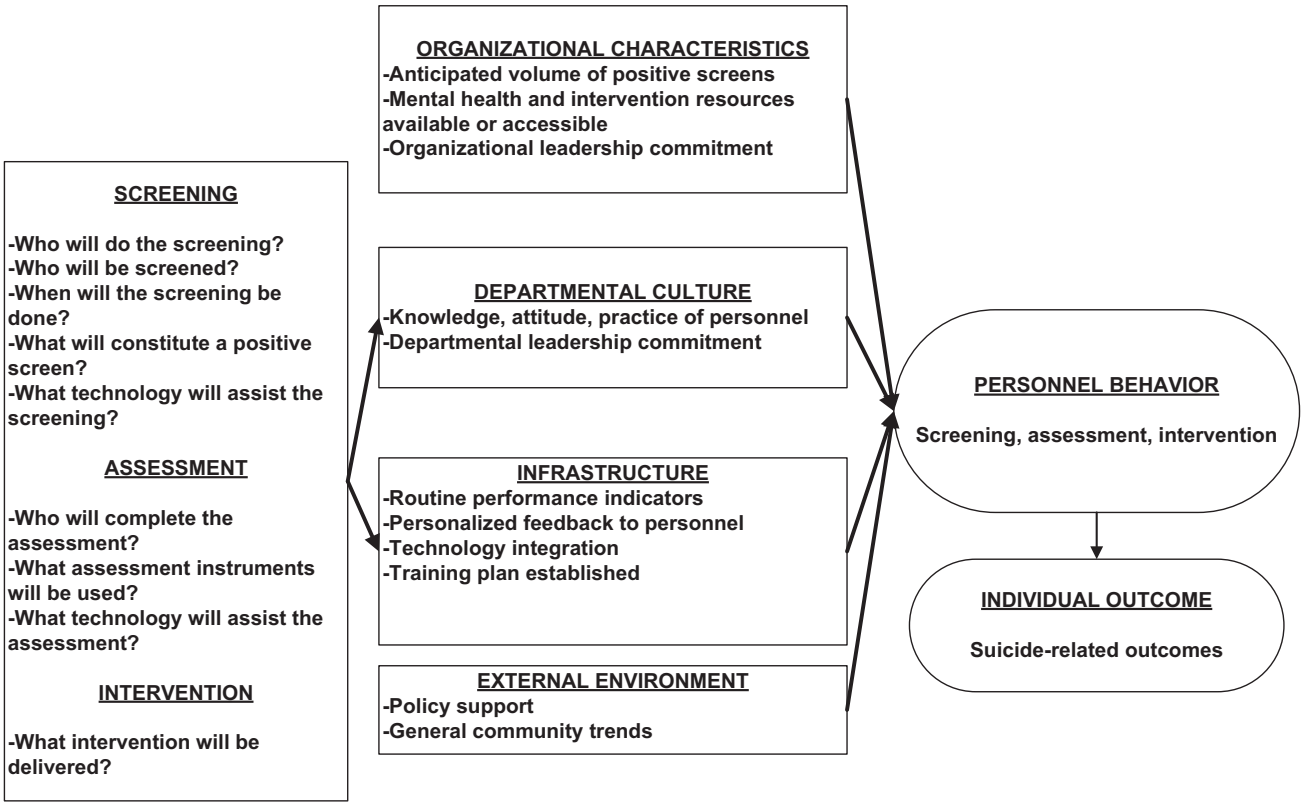

Figure 1. PRISM model template for screening, assessment, and intervention
PRISM, Practical, Robust Implementation and Sustainability Model

important outcomes, such as all-cause death, inpatient admissions, ED visits, or clinical worsening.

To the greatest extent feasible, technology should be used to support screening and assessment. Technology is rapidly revolutionizing health care, and it could be used to help foster the implementation of suicide screening and assessment. Current efforts to design and study instruments should consider the downstream changes that will enable screening and assessment strategies that are simply infeasible now to be disseminated once the technology becomes more readily available. Below, we briefly review several avenues in which technology has the ability to improve suicide risk screening and assessment, and should be the focus of future studies.

EHRs have been publicized as an important tool in improving screening, patient safety, and adherence to clinical guidelines. The existing literature on whether EHRs can accomplish these goals is admittedly mixed; however, the field remains in its early development. We have just begun to establish principles for effective use of EHRs to improve care and establish methods for studying their short- and long-term impacts.

EHRs can be programmed to prompt suicide risk screening, provide guidance for further risk assessment, and facilitate clinical interventions such as discharging patients with outpatient suicide prevention resources. Also, EHRs can be designed to "pre-screen" and alert providers when particularly high-risk individuals are seen, like those with a documented psychiatric disorder or a history of past attempts.

For example, an automated system has recently been developed to predict the development of post-traumatic stress disorder (PTSD) among hospitalized injury survivors using a ten-item algorithm embedded in an EHR.[11] The screener included items reflecting a variety of ICD-9 psychiatric diagnoses, clinical factors such as positive blood alcohol screening, and demographics, and it achieved a sensitivity of 0.71 and specificity of 0.66.

As described previously, the NIH's Patient Reported Outcomes expert panel has made recommendations on a battery of screeners assessing behavioral and mental health outcomes to be integrated into EHRs.[6] Suicide was not included in this because of the lack of evidence-based screeners; however, should such screeners be validated, they could be added to this battery. In addition to helping improve patient care, this would promote standardization in assessment and improve our ability to harmonize data on suicide from diverse healthcare settings.

Patient-facing computing, or readily accessible computer hardware, is not currently commonplace in most

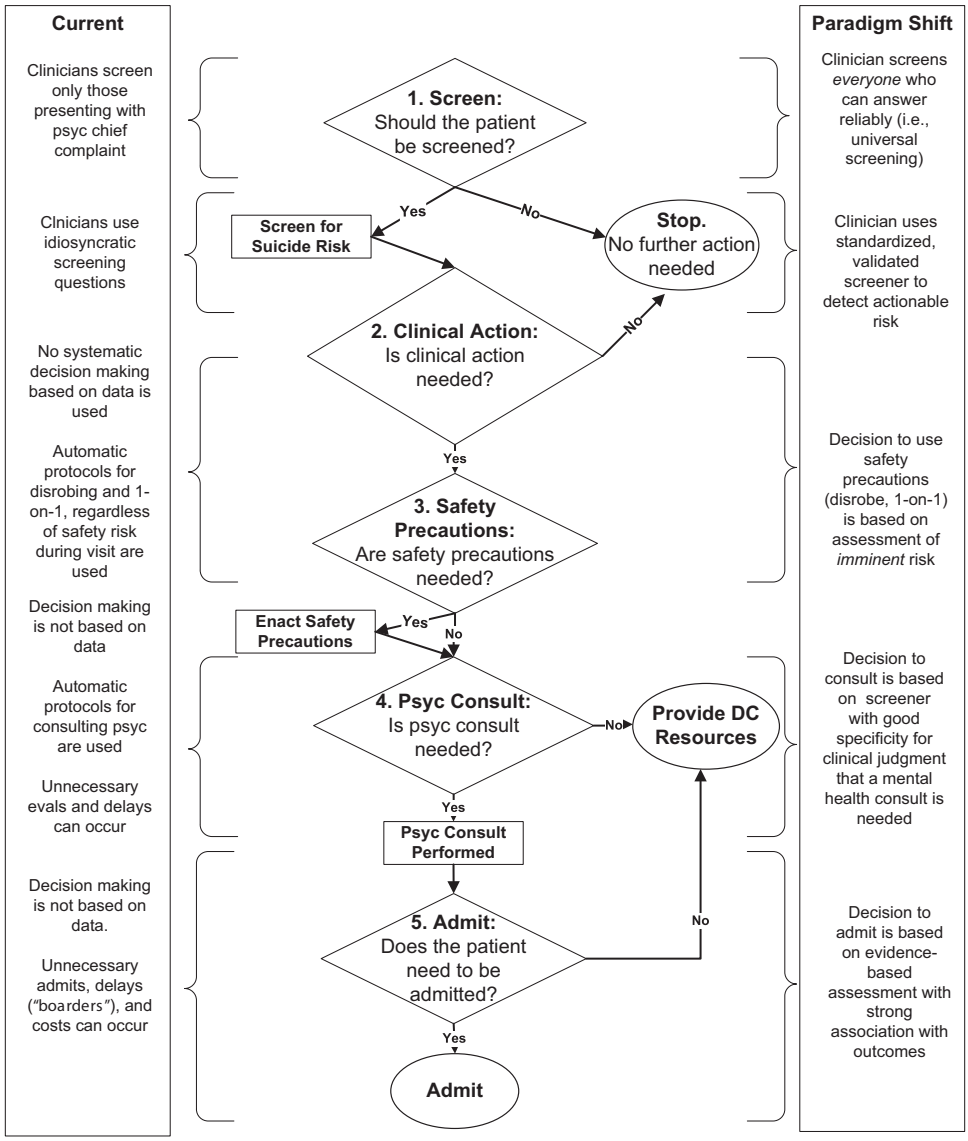

Figure 2. Clinical decision making and suicide risk in the emergency department
DC, Discharge; Psyc, Psychiatry

medical settings, but it is gaining traction. As technology transforms all settings, the probability that medical settings will increase patient-facing computing is highly likely. Computerizing screening and assessment for suicide risk may improve standardization, efficiency, reliability, and validity. In particular, computer adaptive testing and modern psychometric approaches can lead to greater accuracy with the fewest questions necessary, thus maximizing efficiency of both screening and assessment.

Finally, the field of mobile health, in which patients can provide clinical information and receive interventions through mobile phone platforms, is rapidly expanding. This technology may allow for the monitoring of suicidal ideation in an ongoing, longitudinal fashion, rather than simply at discrete points of contact with a healthcare provider. This may be particularly useful for patients in behavioral health settings, those at particularly high risk, and adolescents and young adults who have readily adopted these technologies.

Researchers should study the best methods for translating validated instruments into routine clinical practice. Although following the first two recommendations should help researchers to build validated instruments with a good fit for the setting, it is critical to study how to best translate them in routine practice. Fidelity, or the degree to which an individual adheres to the risk screening and assessment protocols, can be quite different when comparing routine integration into real-world settings against highly standardized research protocols. Consequently, implementation studies are needed to examine the optimal methods for translating these

Patient Safety Screener-5[14] (adults)
Introductory script (sample; modify to fit setting): Now I'm going to ask you some questions that we ask everyone. It is part of our policy and it helps us to make sure we are not missing anything important.
Over the past 2 weeks . . .
1. . . . have you felt down, depressed, or hopeless? ☐ Yes ☐ No
2. . . . have you felt little interest or pleasure in doing things? ☐ Yes ☐ No
3. . . . have you wished you were dead or wished you could go to sleep and not wake up? ☐ Yes ☐ No
4. . . . have you had thoughts of killing yourself? ☐ Yes ☐ No
In your lifetime. . .
5. . . . have you ever attempted to kill yourself? ☐ Yes ☐ No
6. . . . When did this happen? ☐ Today ☐ Within the last 30 days (but not today) ☐ Between 1 and 6 months ago ☐ More than 6 months ago
*Positive screen: yes on #4 or #5.
Ask Suicide-Screening Questionnaire (ASQ)[15] (children)
1. In the past few weeks, have you wished you were dead? ☐ Yes ☐ No ☐ No response
2. In the past few weeks, have you felt that you or your family would be better off if you were dead? ☐ Yes ☐ No ☐ No response
3. In the past week, have you been having thoughts about killing yourself? ☐ Yes ☐ No ☐ No response
4. Have you ever tried to kill yourself? ☐ Yes ☐ No ☐ No response
5. If yes, how? When?
* Positive screen: a positive response to questions 1, 2, 3, or 4.

Figure 3. Examples of frontline screeners

instruments into regular community use in a manner that maintains strong fidelity.

There are many implementation science theories to help guide these examinations.[12] A recent model that has been proposed is the Practical, Robust Implementation and Sustainability Model (PRISM).[13] It integrates several implementation science theories to more fully address the components and shareholders involved in program implementation. The model has four major domains: the intervention or program of focus; the recipients of the intervention or program (usually organizations, clinicians or frontline staff, and patients or students); related infrastructure; and the external environment. Figure 1 depicts a generic PRISM model applied to suicide risk screening, assessment, and intervention.

Illustrative Example: the Emergency Department

In this section, the ED setting is used to illustrate how the paradigm shift and associated recommendations can be put into action. In this setting, a screener should foster the early clinical decisions outlined in Figure 2 that center around detecting and managing imminent risk. A good ED screener should (1) detect when clinically actionable risk is present; (2) identify when an individual requires immediate safety precautions; and (3) identify when a mental health consult is required. Following screening, the risk assessment completed by a mental health professional should then guide the decision to admit the patient to the hospital or provide other services. In this manner, the screening and assessment work hand in hand with clearly defined goals keyed to the decision making each is designed to support.

Instruments such as the BSSI are too complicated or detailed to use as primary screeners without props, so they would be inappropriate for most ED settings. However, other efforts have been closer to the mark, like the Patient Safety Screener–5[14] for adults and the Ask Suicide-Screening Questions[15] for youth, which were developed specifically for use in the ED setting and optimized to be as brief and simple as possible (Figure 3).

The screening process in the ED setting could be enhanced through the use of computerized screeners. Most EDs do not currently have the capacity for patient-facing technology, such as a touch-screen computer that can be used in the treatment area and complies with infection control standards. However, considering the growing technologic revolution that is overcoming health care, it will likely happen within the next 10–20 years, and one can imagine having patients complete computerized assessments while they wait for care. In such an application, safeguards would have to be initiated to ensure patient safety and that patients who screen positive for suicide on the computerized assessment are not accidentally discharged without further evaluation.

Conclusions

The road forward for research on screening and assessing suicide risk within diverse settings will need to navigate a path between two important considerations that are often at odds with one another: The field has to build an evidence base to support clinical decision making based on suicide risk screening and assessments, while developing a better understanding of the practical considerations that influence clinical practice (i.e., feasibility). For research to progress, we must promote the creation and adoption of an evidence-based risk assessment practice with adequate considerations to practical implications important to clinicians, patients, families, and healthcare administrators.

Publication of this article was supported by the Centers for Disease Control and Prevention, the National Institutes of Health Office of Behavioral and Social Sciences, and the National Institutes of Health Office of Disease Prevention. This support was provided as part of the National Institute of Mental Health-staffed Research Prioritization Task Force of the National Action Alliance for Suicide Prevention.

No financial disclosures were reported by the authors of this paper.

References

1. National Action Alliance for Suicide Prevention: Research Prioritization Task Force. A prioritized research agenda for suicide prevention: an action plan to save lives. Rockville MD: National Institute of Mental Health and Research Prioritization Task Force, 2014.
2. Jacobs DG, Baldessarini RJ, Conwell Y, et al. Practice guideline for the assessment and treatment of patients with suicidal behaviors. psychia tryonline.org/content.aspx?bookid=28&sectionid=1673332.
3. Horowitz L, Ballard E, Paoa M. Suicide screening in schools, primary care and emergency departments. Curr Opin Pediatr 2009;21(5):62–7.
4. Substance Abuse and Mental Health Services Administration. Preventing suicide: a toolkit for high schools. Rockville MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration, 2012. HHS Publication No. SMA-12-4669.
5. Beck AT, Steer RA. Manual for the Beck Scale for Suicide Ideation. San Antonio TX: Pearson, 1991.
6. Glasgow RE, Kaplan RM, Ockene JK, Fisher EB, Emmons KM. Patient-reported measures of psychosocial issues and health behavior should be added to electronic health records. Health Aff 2012;31(3):497–504.
7. Schoenwald SK, Hoagwood K. Effectiveness, transportability, and dissemination of interventions: what matters when? Psychiatr Serv 2001;52(9):1190–7.
8. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire–2: validity of a two-item depression screener. Med Care 2003; 41(11):1284–92.
9. Haney EM, O'Neil ME, Carson S, et al. Suicide risk factors and risk assessment tools: a systematic review. VA-ESP Project #05-255. Portland OR: Veterans Health Administration, Department of Veterans Affairs, 2012.
10. National Institute for Health and Clinical Excellence. Self harm: longer term management. NICE clinical guideline—CG133. London. National Institute for Health and Clinical Excellence, 2011.
11. Russo J, Katon W, Zatzick D. The development of a population-based automated screening procedure for PTSD in acutely injured hospitalized trauma survivors. Gen Hosp Psychiatry 2013;35(5):485–91.
12. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. Am J Prev Med 2012;43(3):337–50.
13. Feldstein AC, Glasgow RE. A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. Jt Comm J Qual Patient Saf 2008;34(4):228–43.
14. Allen MH, Abar BW, McCormick M, et al. Screening for suicidal ideation and attempts among emergency department medical patients: instrument and results from the Psychiatric Emergency Research Collaboration (PERC-02). Suicide Life Threat Behav 2013;43(3):313–23.
15. Horowitz LM, Bridge JA, Teach SJ, et al. Ask Suicide-Screening Questions (ASQ): A brief instrument for the pediatric emergency department. Arch Pediatr Adolesc Med 2012;166(12):1170–6.