Suicide Risk Assessment and Prevention: A Systematic Review Focusing on Veterans

Heidi D. Nelson, M.D., M.P.H., Lauren M. Denneson, Ph.D., Allison R. Low, B.A., Brian W. Bauer, M.S., Maya O’Neil, Ph.D., Devan Kansagara, M.D., M.C.R., Alan R. Teo, M.D., M.S.

Objective: Suicide rates in veteran and military populations in the United States are high. This article reviews studies of the accuracy of methods to identify individuals at increased risk of suicide and the effectiveness and adverse effects of health care interventions relevant to U.S. veteran and military populations in reducing suicide and suicide attempts.

Methods: Trials, observational studies, and systematic reviews relevant to U.S. veterans and military personnel were identified in searches of MEDLINE, PsycINFO, SocINDEX, and Cochrane databases (January 1, 2008, to September 11, 2015), on Web sites, and in reference lists. Investigators extracted and confirmed data and dual-rated risk of bias for included studies.

Results: Nineteen studies evaluated accuracy of risk assessment methods, including models using retrospective electronic records data and clinician- or patient-rated instruments. Most methods demonstrated sensitivity ≥80% or area-under-the-curve values ≥.70 in single studies, including two studies based on electronic records of veterans and military personnel, but specificity varied. Suicide rates were reduced in six of eight observational studies of population-level interventions. Only two of ten trials of individual-level psychotherapy reported statistically significant differences between treatment and usual care.

Conclusions: Risk assessment methods have been shown to be sensitive predictors of suicide and suicide attempts, but the frequency of false positives limits their clinical utility. Research to refine these methods and examine clinical applications is needed. Studies of suicide prevention interventions are inconclusive; trials of population-level interventions and promising therapies are required to support their clinical use.

Psychiatric Services in Advance (doi: 10.1176/appi.ps.201600384)

Despite the large public health impact of suicide in the United States (1), research supporting clinical guidelines for suicide prevention has been limited. Universal evaluation for suicide risk (screening) in general medical practice is currently not recommended (2,3), although a Sentinel Event Alert regarding identification of suicide risk has been issued by The Joint Commission (4). However, primary care clinicians and other health care practitioners are often well positioned to serve as initial contacts in the assessment and referral of high-risk individuals. During the year prior to suicide, an estimated 83% of suicidal individuals used health care services, including 64% who had contact with primary care and 45% who had contact with mental health care clinicians (5).

Efforts to prevent suicide generally include referrals for treatment of underlying conditions, psychotherapy, and management of imminent risk, such as hospitalization and development of safety plans (6,7). In addition to individual-level approaches, an increasing number of suicide prevention initiatives have been implemented at organizational, health system, and community levels. However, their influence on suicide prevention remains unclear because few studies demonstrating effectiveness have been published (7).

Suicide rates in the United States, including rates in veteran and military populations, are high (8,9). Rates of suicide increased among active duty soldiers during the wars in Afghanistan and Iraq (10), and between 2000 and 2010, the suicide rate among veterans surpassed the rate among civilians (11). Female veterans are at especially high risk compared with women who are not veterans (11). These trends have led to new initiatives and additional research funding to address suicide risk detection and prevention in both the U.S. Department of Veterans Affairs (VA) and the Department of Defense. As part of this effort, this systematic review updates evidence of the accuracy of methods to identify individuals at increased risk of suicide and the effectiveness and adverse effects of health care interventions relevant to U.S. veteran and military populations in reducing suicide and suicide attempts. This review adds to existing research (12–16) by focusing on studies applicable to veteran and military populations, including risk assessment studies and trials of interventions, and evaluating the strength of evidence by using established methods.
METHODS

Scope, Key Questions, and Analytic Framework
The scope and key questions were determined during a topic refinement process that included a preliminary review of published peer-reviewed literature; discussion with internal partners, policy makers, and investigators; and consultation with content experts and key stakeholders. This review followed an established systematic review methodology (17), and a protocol was posted to the PROSPERO International Prospective Register of Systematic Reviews Web site (http://www.crd.york.ac.uk/PROSPERO; registration number CRD42015019089) (18). A technical report describes additional methods (19). Institutional review board approval was not required for this systematic review.

We created an analytic framework outlining the key questions, patient populations, interventions, and outcomes of the review. [A figure depicting the framework is included in an online supplement to this article.] Key questions addressed the accuracy and adverse effects of methods to identify individuals at increased risk of suicide and attempts and the effectiveness and adverse effects of suicide prevention interventions in reducing rates of suicide and attempts. Interventions included health care services directed toward populations (for example, educational programs and practice changes) and individuals (for example, psychotherapy and case management).

Outcomes of interventions included reduced suicide and other suicidal self-directed violence. Suicidal self-directed violence is behavior against oneself that deliberately results in injury or the potential for injury with evidence of suicidal intent (that is, nonfatal suicide attempts) (20,21). Suicide is a fatal outcome of suicidal self-directed violence. Suicidal ideation and nonsuicidal self-directed violence, or composite outcomes that include these thoughts or behaviors, were outside the scope of this review. Adverse effects of these interventions included health outcomes as well as other outcomes, including patient anxiety and distress, labeling, and stigma.

A research librarian conducted electronic database searches of MEDLINE by using PubMed, PsycINFO, SocINDEX, Cochrane Central Register of Controlled Trials, and the Cochrane Database of Systematic Reviews between January 1, 2008, and September 11, 2015, to capture studies published since previous systematic reviews on this topic (6,7,22). [A list of search terms is provided in the online supplement.] Additional sources included a search of relevant research and organizations’ Web sites and citations from reference lists, reviews, conference proceedings, and experts.

Study Selection
Prespecified eligibility criteria were used to determine inclusion of abstracts and articles [see online supplement]. English-language full-text articles identified as potentially relevant were independently examined by two of the authors. At least one reviewer for each study was a physician or a doctoral-level investigator experienced in the field; disagreements were resolved through consensus by using a third reviewer.

Consistent with the previous reviews (6,7,22), eligible studies included populations of veterans, military personnel, and demographically comparable nonveteran adults ages 18 and older from the United States, United Kingdom, Canada, New Zealand, and Australia. These countries were chosen because of their similarities to the United States in terms of health care services and their involvement in Operation Iraqi Freedom and Operation Enduring Freedom and conflicts in Iraq and Afghanistan. Studies enrolling participants substantially different from the general veteran population (for example, children and adolescents) or with serious comorbidities (for example, advanced cancer) were outside the scope of this review.

Studies were included that evaluated the diagnostic accuracy or adverse effects of methods to assess risk of suicide and attempts, including instruments, checklists, and other approaches appropriate for clinical settings. Outcomes included sensitivity, specificity, area under the receiver-operator characteristic curve (AUC), and similar measures. Studies were excluded that primarily determined associations between individual risk factors and suicide and attempts, evaluated psychometric characteristics of instruments, or provided only descriptions of methods without reporting measures of accuracy.

Studies of the efficacy or effectiveness of interventions were eligible if the interventions were specifically designed to prevent suicide and attempts, including randomized controlled trials (RCTs), observational studies with comparison groups, and systematic reviews of studies with these study designs. These included primary, secondary, and tertiary prevention intervention studies of health care services clinically relevant to medical practice in the United States. Studies of interventions that were intended to primarily treat coexisting conditions and studies of pharmacotherapy were outside the scope of this review.

Data Extraction and Quality Assessment
Details of the study design, setting, patient population, methodology, and results were abstracted into a customized, prepiloted database by one author and confirmed by a second. Two authors independently assessed the risk of bias of included studies by using prespecified criteria for RCTs (23), observational studies (24,25), and diagnostic-accuracy studies (24,26). Each study was given an overall summary assessment of low, high, or unclear risk of bias. Discrepancies were resolved through consensus by using a third author.

Data Synthesis
Measures, interventions, outcomes, and study participants were too heterogeneous to combine in statistical meta-analysis, and conclusions are based on qualitative synthesis of the
findings that emphasizes the strongest studies (that is, best evidence approach). The overall strength of evidence for studies of interventions was determined by using a method developed for the Agency for Healthcare Research and Quality’s Evidence-Based Practice Centers (27). This method does not provide strength-of-evidence grades for diagnostic-accuracy studies.

RESULTS

Results of the literature search and selection process are summarized in a literature flow diagram [see figure in online supplement]. Of 7,788 citations reviewed, 681 articles were selected for full-text review, and 37 studies met inclusion criteria.

Methods to Identify Suicide Risk

Nineteen studies evaluated the accuracy of methods to identify suicide risk (28–46) [see table in online supplement]. No studies described the adverse effects of risk assessment methods. Five studies evaluated methods that could be offered to general populations (universal or primary prevention) (32,36,37,39,42); nine studies evaluated methods that target subpopulations likely to be at increased risk of suicide, such as patients presenting to emergency departments (selective or secondary prevention) (28–30,33,34,38,40,41,46); and five studies evaluated methods for individuals who have already been identified as having increased risk, such as patients admitted to a hospital for suicidal ideation (tertiary prevention) (31,35,43–45). Studies enrolled participants from the community, online, emergency departments, and psychiatry services or used data from existing patient medical records or administrative databases. Six studies included veterans or military personnel specifically (28–31,36,37). Most studies included high-risk populations, such as adults presenting to emergency departments with suicide attempts (33,38,40,43–45) or who had psychiatric hospitalizations or psychiatric risk factors that indicated higher risk of suicide (28–31,34–36,39,41,46), most commonly depression or previous suicide attempts.

Thirteen studies used one or more clinician-rated or patient–self-reported instruments to assess individual levels of risk (28–30,32–35,38,42–46). Instruments incorporated questions and scales indicating the presence and severity of known or suspected suicide risk factors, and the studies determined how well participants’ responses predicted suicidal behaviors. In one study, participants’ responses to nine separate scales were analyzed in various combinations to determine a model with the best diagnostic accuracy (34).

Six studies used existing data from electronic medical records or administrative databases to identify suicide risk factors in populations of patients and then stratified the patient populations by levels of suicide risk by using regression analysis (31,36,37,39–41). Of these studies, three derived models from development data sets and tested them in validation data sets (37,40,41).

Six of the 19 studies had important methodological limitations resulting in high risk-of-bias ratings (28,29,35,43–45); risk of bias was unclear in ten studies (30–34,38,39,41,42,46), and low in three studies (36,37,40). Major limitations of studies with high or unclear risk of bias included small sample sizes, including six studies with sample sizes less than 200 (28,35,38,43–45), high or unclear loss to follow-up (28,29,33–35,37,38,44,45), and potentially biased participant selection (28,30–32,34,35,38,43–46).

Results of studies indicated estimates of sensitivity ranging from 50% to 100% (with one outlier at 11%) and AUC values from .57 to .97 (Figures 1 and 2) [also see tables in the online supplement]. The majority of methods had estimates of sensitivity ≥80% or AUC ≥ .70, a threshold suggesting fair or better discrimination between patients with and without suicide or suicide attempts. These methods include SAD PERSONS and variations (33), Suicide Opinion Questionnaire (35), ReACT Self-Harm Rule (40), Suicidal Ideation Attributes Scale (42), modification of the Affective Intensity Rating Scale (43), Suicide Trigger Scale (44,45), Schedule for Nonadaptive and Adaptive Personality self-harm subscale (46), suicide potential index subscale of the Personality Assessment Inventory (28), a decision tree with predictors of suicide attempts (31), and prediction models incorporating electronic database records (34,36,37,41).

All three studies with low risk of bias evaluated methods derived from large patient databases (36,37,40). The ReACT Self-Harm Rule is a four-item model developed from data from 18,680 patients in England presenting to emergency departments with self-harm. Items include self-harm in the past year, living alone or homelessness, cutting as a method of harm, and treatment for a current psychiatric disorder. The ReACT Self-Harm Rule demonstrated sensitivity of 88% and specificity of 24% for predicting suicide within six months (40).

A study using data from nearly six million VA patients to create a prediction model to stratify patients according to their risk of suicide within the next year reported an AUC of .761 (95% confidence interval [CI]=.751–.771) (37). Results suggested that the model was a better predictor of suicide than the current clinical practice standard of health care providers “flagging” medical records of individuals believed to be at higher risk of suicide.

The only study of military personnel was based on the Army Study to Assess Risk and Resilience in Service-members (Army STARRS) and included 40,820 active duty U.S. Army soldiers hospitalized with psychiatric diagnoses (36). A risk algorithm to predict suicides within one year of hospitalization was developed from administrative data systems and demonstrated AUCs as high as .89.

Health Care Service Interventions for Suicide Prevention

Interventions directed toward populations. Eight studies of the effectiveness of population-level interventions met inclusion criteria (Table 1) (47–54). One study evaluated a secondary prevention intervention (51), and the others...
targeted primary prevention. These initiatives included components generally categorized as education, awareness, individual health, and individual risk monitoring. One study was designed as a retrospective cohort study (48), six were before-after studies (47,49,51–54), and one was a postintervention series (50) for which risk-of-bias criteria were not applicable.

Suicide rates were lower after interventions in six observational studies, including studies of the Air Force Suicide Prevention Program (49,55), a program for an Army Infantry Division deployed to Iraq (50), and studies of police officers (52), university students (51), and health systems (47,54,56). Suicide rates were not lower in two studies of community programs (48,53). However, interventions differed across studies, and all studies were limited by risk of bias inherent to nonrandomized study designs—particularly, that potential confounders were not considered and groups may not have been comparable. Strength-of-evidence grades were low for suicide outcomes (eight observational studies), insufficient for suicide attempt outcomes (no studies), and insufficient for adverse effects (no studies) [see table in online supplement].

Interventions directed toward individuals. Ten RCTs of the effectiveness of individual-level health care interventions met inclusion criteria (Table 2) (57–67). Two studies evaluated interventions for secondary prevention (59,64), and the others evaluated interventions for tertiary prevention. Most trials compared usual care with various types of psychotherapy, including cognitive-behavioral therapy, dialectical behavior therapy, personal construct psychotherapy, and problem-solving therapy. One trial examined collaborative care (57,64), and another examined day hospital attendance (59). Participants included outpatient military personnel and nonmilitary psychiatric inpatients or patients at acute risk of suicide. No studies specifically evaluated adverse effects of the interventions, and none specifically considered the effects of pharmacotherapy administered outside the study protocol during the study.

Risk of bias was rated high in five trials (58,59,61,66,67) and unclear in five (57,60,62–65). Most studies were underpowered to detect differences between treatment and usual care. Additional limitations included lack of information on randomization (59,66,67), allocation concealment (58–60,63,65–67), and blinding (58,59,64,66,67); lack of outcome reporting (59,61,67); and unclear or lack of specified outcome measures (60,63,66).

Only two of ten trials reported statistically significant differences between treatment and usual care (63,65), although risk of bias was unclear for both trials. In a trial of outpatient active duty soldiers with recent suicide attempts or ideation, those in a brief cognitive-behavioral therapy program were less likely to have made suicide attempts at two-year follow-up than those in usual care (14% versus 40%, p=.02; hazard ratio=.38, 95% CI=16–87) (63). In a trial involving women with borderline personality disorder, those receiving dialectical behavior therapy had fewer suicide attempts than those receiving usual care at one-year follow-up (23% versus 46%, p=.01) (65). Strength-of-evidence grades were low for suicide attempt outcomes (seven trials), insufficient for suicide outcomes (four trials), and insufficient for adverse effects (no studies).
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Population</th>
<th>Intervention</th>
<th>Results</th>
<th>Risk-of-bias rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coffey, 2007 (47)</td>
<td>Before-after</td>
<td>General mental health and substance abuse patients in a U.S. health care system; baseline year 2000, start-up year 2001, follow-up 2002–2005</td>
<td>Henry Ford Health System’s Perfect Depression Care initiative utilized 6 aims and 10 rules from the Institute of Medicine <em>Crossing the Quality Chasm</em> report and focused on improving partnership with patients, clinical care (planned care model), access, and information flow.</td>
<td>Suicide rate (per 100,000): 89 baseline, 77 start-up year, 22 follow-up average (p=.007 compared with baseline)</td>
<td>NA&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Harris, et al., 2008 (48)</td>
<td>Retrospective cohort</td>
<td>7,760 patients ages 15–29, 59.7% male, with psychotic disorders receiving mental health services in Victoria, Australia, 1991–1999</td>
<td>The Early Psychosis Program involved inpatient specialized care at the Early Psychosis Prevention and Intervention Center for up to 24 months. All patients receiving care in the program were compared with patients receiving nonspecialist adult mental health care from other clinics in the area.</td>
<td>Suicide rate across 8.5 years: 3.8% intervention versus 4.2% usual care (p=.84)</td>
<td>Unclear</td>
</tr>
<tr>
<td>Joffe, 2008 (51)</td>
<td>Before-after</td>
<td>University of Illinois students and citizens of Champaign County, 1984–2005</td>
<td>Secondary prevention intervention that mandated students with suicide attempts or threats to receive 4 treatment sessions conducted by mental health professionals. Failure to comply resulted in sanctions.</td>
<td>Suicide rate (per 100,000): 6.91 preintervention versus 3.78 postintervention after 21 years of the program (p&lt;.05)</td>
<td>NA&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Knox et al., 2010 (49); Knox et al., 2003 (55)</td>
<td>Before-after</td>
<td>&gt;5 million service personnel in the U.S. Air Force, 1981–2008</td>
<td>An 11-component initiative was implemented starting in 1997: leadership involvement, suicide prevention education, commander guidelines for use of mental health services, community prevention services, community education and training, investigative interview policy, trauma stress response, integrated delivery system and community action information board, limited-privilege suicide prevention program (increased confidentiality), assessment, and suicide event surveillance.</td>
<td>Mean quarterly suicide rate (per 100,000): 3.033 preintervention versus 2.387 postintervention (p&lt;.01); relative risk of suicide pre-versus postimplementation: .67 (95% CI=.57–.80)</td>
<td>NA&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mishara and Martin, 2012 (52)</td>
<td>Before-after</td>
<td>4,178 members of the Montreal police in Quebec, Canada, 1997–2008</td>
<td>The Together for Life Suicide Prevention Program consisted of education, police resources, training for supervisors and union representatives, and a publicity campaign.</td>
<td>Suicide rate (per 100,000): 30.5 preintervention versus 6.4 postintervention (p=.008)</td>
<td>NA&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Continued
| Study                  | Design                     | Population                                                                 | Intervention                                                                                                                                                                                                 | Results                                                                                                                   | Risk-of-bias rating |
|-----------------------|----------------------------|----------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| Walrath et al., 2015  | Ecological comparison      | Youths and adults in 479 counties across the United States, 2007–2010       | The Garrett Lee Smith youth suicide prevention program consisted of education, gatekeeper training, screening activities, improvement of linkages to services, crisis hotlines, and community partnerships. The study compared counties that implemented gatekeeper training with matched counties without training. | At 1 and 2 years after training, no differences for adults age ≥19                                                   | NA<sup>b</sup>      |
| Warner et al., 2011   | Postintervention series    | 40,283 soldiers in a deployed U.S. Army unit, 15 months in Iraq (March 2007 to May 2008) | Multiple-component intervention for deployed unit included 4 phases: predeployment (suicide risk recognition and response training, early identification, and resiliency training for soldiers and families), deployment (education, suicide prevention review board and suicide risk management teams, unit behavioral health needs assessment, unit behavioral health advocates, incident response, and trend monitoring), redeployment (education, postdeployment health assessment, and risk stratification), and reintegration (complete redeployment tasks, prepare for reuniting with families, and address postdeployment health issues). | Suicide rate (per 100,000): 16.0 for intervention unit during the deployment cycle, 24.0 for service members in theater (19.2 across the U.S. Army) | NA<sup>b</sup>      |
| While et al., 2012    | Before-after               | 12,881 suicide deaths in mental health services in Wales and England, 1997–2006 | Assessed 9 of the 12 key service recommendations from the English Suicide Prevention Strategy: removing ligature points on inpatient wards, assertive outreach, 24-hour crisis team, follow-up after psychiatric discharge within 7 days, written policy on response to patients noncompliant with treatment, written policy on management of patients with co-occurring mental and substance use disorders, criminal justice sharing, multidisciplinary review and sharing information with families after suicide, and clinical training about suicide risk for staff | Suicide rates that declined pre-versus postintervention (per 10,000 per year): 24-hour crisis care, 11.44 versus 9.32 (p<.001); multidisciplinary review, 11.51 versus 11.39 (p<.001); and co-occurring disorders, 10.51 versus 9.61 (p<.001). | NA<sup>b</sup>      |

<sup>a</sup> The outcome evaluated in each study was suicide.

<sup>b</sup> Risk-of-bias criteria not applicable (NA) for this study design.
<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Intervention and comparison</th>
<th>Outcome</th>
<th>Results</th>
<th>Risk-of-bias rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comtois et al., 2011 (58)</td>
<td>32 adults ages 19–62, 62% female, with recent suicide attempt or imminent risk; U.S.</td>
<td>CAMS® (patients identify the causes of suicidal ideation and the reduction in suicidal ideation and behavior as a coping strategy; 4–12 sessions lasting 50–60 minutes with CAMS clinicians versus enhanced usual care (intake with psychiatrist, 1–11 visits with case manager, and medication management as needed))</td>
<td>Suicide attempts and Self-Injury Count score</td>
<td>Suicide attempts and self-inflicted injuries at 12-month follow-up (M±SD): CAMS, 1.2±3.9; enhanced usual care, 3.3±7.6</td>
<td>High</td>
</tr>
<tr>
<td>Gallo et al., 2007 (64); Alexopoulos et al., 2009 (57)</td>
<td>599 adults ages ≥60, 72% female, with score ≥20 on the Centers for Epidemiologic Studies Depression Scale</td>
<td>Intervention (on-site depression care manager working with primary care physicians to provide algorithm-based care) versus usual care (educational sessions for primary care physicians and notification of patients' depression but no specific recommendations for individual patients except for psychiatric emergencies)</td>
<td>Suicide</td>
<td>N of suicides at 2-year follow-up and N per 1,000 person-years: intervention, 1 and .7 (95% CI=.0–4.2); usual care, 0 and .0 (95% CI=.0–3.3); N of suicide attempts at 2-year follow-up: intervention, 2; usual care, 3</td>
<td>Unclear</td>
</tr>
<tr>
<td>Jones et al., 2008 (59)</td>
<td>206 patients admitted to an adult psychiatric ward from 1999 to 2002, 52% female; London</td>
<td>Day hospital (attendance expected 9:30 a.m. to 4:30 p.m., with drop-in service on weekends; emphasis on group activities) versus inpatient (conventional psychiatric care and limited program of daily activities)</td>
<td>Suicide</td>
<td>1 suicide in the day hospital group and 1 in the inpatient group at 12 months postdischarge</td>
<td>High</td>
</tr>
<tr>
<td>Linehan et al., 2006 (65)</td>
<td>111 women ages 18–45 with borderline personality disorder and current and past suicidal behaviors</td>
<td>DBT® (cognitive-behavioral treatment for suicidal women meeting criteria for borderline personality disorder; treatment targets suicidal behavior, behaviors interfering with treatment delivery, and other severe behaviors for 1 year versus community treatment by experts (usual care, with treatment uncontrolled by the research team))</td>
<td>Suicide attempts and Suicide Attempt Self-Injury Interview</td>
<td>Suicide attempts: DBT, 23%; community treatment by experts, 46% (p=.01, hazard ratio [HR]=2.66, p=.005)</td>
<td>Unclear</td>
</tr>
<tr>
<td>Linehan et al., 2015 (60)</td>
<td>99 adult women with borderline personality disorder with ≥2 suicide attempts or nonsuicidal self-injury acts within 5 years from 2004 to 2010; Seattle</td>
<td>Standard DBT (weekly individual therapy and group skills training, a therapist consultation team, and between-session telephone coaching as needed) versus DBT with skills training (provided group skills training, removed individual component and replaced it with case management) versus DBT with individual therapy (eliminated all skills training and added an activity-based support group)</td>
<td>Suicide and suicide attempts</td>
<td>1 suicide in standard DBT group; no differences in suicide attempts</td>
<td>Unclear</td>
</tr>
</tbody>
</table>

**TABLE 2. Randomized controlled trials of individual-level health care service interventions for suicide prevention included in the review**
<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Intervention and comparison</th>
<th>Outcome</th>
<th>Results</th>
<th>Risk-of-bias rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>McAuliffe et al., 2014 (61)</td>
<td>433 psychiatric patients ages 18–64, 65% female, in emergency or inpatient units who reported self-harm within past 3 days; Ireland</td>
<td>Problem-solving skills training (six 2-hour sessions of manualized interpersonal problem-solving skills training) versus usual care (assessment and mental health or crisis services referral)</td>
<td>Suicide</td>
<td>1 suicide in problem-solving skills training versus 2 in usual care at 12-month follow-up</td>
<td>High</td>
</tr>
<tr>
<td>McMain et al., 2012 (62)</td>
<td>180 adults with borderline personality disorder; Toronto</td>
<td>DBT (comprehensive multicomponent intervention for individuals with high suicide risk; contains 4 weekly components, including individual therapy, group skills training, therapist consultation, and as-needed between-session telephone coaches for 1 year) versus general psychiatric management (psychodynamic psychotherapy, case management, and pharmacotherapy for 1 year)</td>
<td>Suicide attempts</td>
<td>At 36-months follow-up, no differences between groups (p=.83).</td>
<td>Unclear</td>
</tr>
<tr>
<td>Rudd et al., 2015 (63)</td>
<td>152 active duty Army personnel, 88% male; Fort Carson, Colorado</td>
<td>Brief outpatient cognitive-behavioral therapy (12 sessions, 1–2 weeks apart; first session 90 minutes and following sessions 60 minutes; 3 phases included assessment, cognitive strategies to reduce beliefs and assumptions that serve suicidal thoughts, and relapse prevention) versus usual care (treatment as usual)</td>
<td>Suicide and Suicide Attempt Self-Injury Interview</td>
<td>After 2 years of follow-up, ≥1 suicide attempt by 8 individuals in therapy versus 18 in usual care (14% versus 40%, p&lt;.02); multivariate Cox regression controlled for baseline risk: HR=.31 (95% CI=13–75).</td>
<td>Unclear</td>
</tr>
<tr>
<td>Stewart et al., 2009 (66)</td>
<td>32 adults ages 20–58, 53% female, receiving inpatient treatment for suicide attempts</td>
<td>Cognitive-behavioral therapy (7 sessions, 1 hour each) versus problem-solving therapy (4 sessions, 1 hour each) versus treatment as usual (usual care provided by local hospital)</td>
<td>Repeated suicide attempts</td>
<td>Average N of suicide attempts (M±SD): cognitive-behavioral therapy, 22±6.4, versus usual care, 22±50 (not significant); problem-solving therapy, 33±63, versus usual care, 22±50 (not significant)</td>
<td>High</td>
</tr>
<tr>
<td>Winter et al., 2007 (67)</td>
<td>64 adults, 53% female, receiving emergency care following self-harm</td>
<td>Personal construct psychotherapy (2–22 sessions [mean 10.38]), therapeutic techniques appropriate to particular personal construct formulations of the patient’s self-harm as set out in a brief manual versus usual care (assessment and possible follow-up appointments with a mental health team)</td>
<td>Suicide</td>
<td>1 suicide for therapy intervention versus 2 for usual care</td>
<td>High</td>
</tr>
</tbody>
</table>

---

<sup>a</sup> CAMS, Collaborative Assessment and Management of Suicidality  
<sup>b</sup> DBT, dialectical behavior therapy  
<sup>c</sup> Two suicide attempts or self-injuries within the past five years, with one in past eight weeks
DISCUSSION

Nineteen studies of risk assessment methods to identify individuals at increased risk of suicide or suicide attempts evaluated 19 different approaches. Results indicated fair or better diagnostic accuracy for most methods. However, methodological limitations in many of the studies threatened the validity and applicability of the findings.

Prediction models derived from large patient databases may provide a more rigorous approach to risk assessment than other methods. Although some models demonstrated fair to good diagnostic accuracy in single studies with low risk of bias, they have not yet been implemented and further tested in clinical settings. Several remaining concerns include their feasibility in clinical practice, impact on workflow, and incremental value in identifying patients for further intervention. For example, although results indicated high predictive accuracy in a study of the Army STARRS risk algorithm, this method incorporated information from 38 administrative databases, including data that may not or cannot be obtained prospectively. Whether key information in the model would be routinely kept up to date and available to inform clinical practice in real time is unclear. A new model based on Army STARRS administrative data from outpatient mental health visits, published after the systematic review searches, may provide a more feasible approach (68). Another approach, using responses to the suicidal ideation question from the nine-item Patient Health Questionnaire (PHQ-9), a measure frequently used in clinical practice, was recently found to be associated with up to a 75% increased risk of suicide in a sample of nearly 400,000 VA patients (69).

Studies of various clinician-rated or patient–self-reported risk assessment instruments and scales indicated accuracy that varied across methods and cutoff points. Some instruments may provide diagnostic value to specific patient subgroups, such as those with previous suicide attempts or co-occurring conditions. Although these methods may be useful, studies evaluating them are currently inconclusive and limited by small sample sizes, methodological limitations, and unclear applicability to clinical practice. These deficiencies also limit comparisons between types of risk assessments, patient subgroups, settings, or other factors. Although these methods appear in general to have lower accuracy than more complex models based on electronic records or administrative data, the accuracy of the two approaches was compared in only one study (41).

One of the biggest challenges in using methods to identify suicide risk in clinical settings is the low incidence rate of suicide, especially in primary care or ambulatory mental health care populations. Our results revealed several methods with reasonably high sensitivity. However, positive predictive values were much lower because, unlike sensitivity, predictive values are affected by the low incidence of suicide. It is unclear at this point what level of risk would be needed for a specific method to be clinically useful. For example, even with a positive predictive value of 30% for suicidal behavior, as observed in included studies of high-risk veteran populations (30,31), 70 of every 100 cases would be false positives. Moreover, the false-positive rate becomes exponentially higher when the outcome is suicide itself, even in high-risk populations. The false-positive rate is even higher when suicide risk assessment is applied broadly across unselected populations. For example, if using the ReACT Self-Harm Rule with a sensitivity of 88% and specificity of 24% for assessing risk among all veterans receiving care in VA facilities (suicide rate of 27.6 per 100,000) (11), the positive predictive value would be only 3%. Given that this review found no studies of adverse effects, such as involuntary hospitalization or other actions that might be triggered by a positive risk assessment, the inadequacy of the current evidence base to offer practical guidance to clinicians becomes even more apparent.

Future research should build on current work. Risk models should be further refined, and those demonstrating fair to good diagnostic accuracy, such as the ReACT Self-Harm Rule, should be prospectively evaluated in large clinical populations, preferably through cluster-randomized trials. Studies should also assess the effectiveness and potential adverse effects of risk prediction tools in clinical care pathways designed to prevent suicide or suicide attempts. To ensure their clinical applicability to VA and military populations, methods should be specifically evaluated in these groups.

Future research should also test novel approaches to suicide risk assessment that build on existing developmental work. Although many studies of risk assessment methods have relied on patient self-reported data, a study of the computer-administered Implicit Association Test is an important exception (38). This test uses individuals’ reaction times when classifying semantic stimuli to predict future suicide attempts. In addition, individuals with histories of suicide attempts exhibit certain patterns of cognitive deficits that can be detected by tasks commonly included in neuro-psychological testing batteries (70,71). Exploratory studies examining biological markers for suicide or suicidal behaviors (72–74) or neuroimaging to identify associations with suicide-related outcomes (75,76) could be further expanded to determine their roles in clinical care and among specific patient subgroups. Ultimately, risk assessment approaches should be brief and easily administered for clinical applications across multiple settings.

Eight studies of the effectiveness of population-level interventions and ten RCTs of individual-level health care interventions met inclusion criteria for this review. However, results were inconclusive because interventions varied across studies, risk of bias could not be assessed for many studies because of study design, comparability of groups was not established, and most studies were underpowered to detect differences between intervention and comparison groups. In addition, high-risk patients may have been excluded from studies because of their need for acute care, and adverse effects were not reported.
Studies of population-level health care service interventions for suicide prevention should be conducted in additional populations to further validate results of the initial studies and to demonstrate the programs’ applicability to general clinical practice. Many of the intervention studies compared outcomes before and after the interventions were implemented, raising concerns regarding regression to the mean and cohort effects. Although RCTs may be impractical on a population level, more robust observational study designs, such as interrupted time-series analysis, would strengthen the evidence base. Additional details about how the program components were implemented and maintained in practice are also necessary to establish portable service packages and translate this work to other settings. Restricting access to lethal means is an effective method of suicide prevention (7,77), although no recently published studies of health care services using this approach met inclusion criteria for this review. Examples of current efforts in the VA to reduce access to lethal means include an ongoing study of blister packaging for medications (78) and distribution of gun locks. Studies examining efforts to counsel veterans on firearms safety and delaying or restricting access to firearms have been published (79,80), but more work is needed to help guide clinician discussions with veterans and to establish the effectiveness of such strategies for VA health care settings.

Studies of individual-level interventions generally targeted individuals identified as at high risk of suicide because of recent suicide attempts or self-harm or existence of psychiatric conditions. Although these approaches are important, no studies tested interventions to increase protective factors, such as social integration (81,82). Improvement of this evidence base will require larger, more rigorous RCTs of existing interventions as well as innovative approaches, such as novel uses of technology to support or enhance care for individuals at risk of suicide (83,84), safety planning in health settings (85–87), and peer support.

This review was limited by its focus on veteran and military populations in the United States, which may have excluded important studies. However, this focus improved its clinical relevance to this population, as well as other adults with similar demographic characteristics. Of the 37 studies included in this review, six studies of risk assessment and three studies of interventions specifically included veterans or military personnel. Also, this review included only English-language articles, which could result in language bias, although studies not published in English may be less relevant to the target population. Studies were not available for some key questions, including comparisons of the accuracy of risk methods by settings, delivery modes, and targeted populations. One of the useful measures of test performance—positive predictive value—could not be summarized and compared because it was often not included in studies. Data on adverse effects of risk assessment and prevention interventions were also absent across studies.

**CONCLUSIONS**

Risk assessment methods have been shown to be sensitive predictors of subsequent suicide and suicide attempts, but the frequency of false positives limits their clinical utility. Future research should continue to refine these methods and examine clinical applications. Studies of suicide prevention interventions provide inconclusive evidence to support their use, and additional RCTs of promising individual therapies and site-randomized population-level interventions are needed.

**AUTHOR AND ARTICLE INFORMATION**

Dr. Nelson and Dr. Kansagara are with the Department of Medical Informatics and Clinical Epidemiology and Dr. Denneson, Dr. O’Neil, and Dr. Teo are with the Department of Psychiatry, all in the School of Medicine at Oregon Health and Science University, Portland. Dr. Denneson, Dr. O’Neil, Dr. Kansagara, and Dr. Teo are also with the Center to Improve Veteran Involvement in Care, Portland Department of Veterans Affairs Medical Center, Portland, Oregon, where Ms. Low is affiliated. Mr. Bauer is with the Department of Psychology, University of Southern Mississippi, Hattiesburg. Send correspondence to Dr. Teo (e-mail: teoa@ohsu.edu).

This work was funded by the Quality Enhancement Research Initiative, Office of Research and Development, Veterans Health Administration (VHA), U.S. Department of Veterans Affairs. Dr. O’Neil’s work was supported in part by PCOR K12 award 1 K12 HS019456 01 from the Agency for Healthcare Research and Quality. Dr. Teo’s work was supported in part by career development award CDA 14-428 from the VHA’s Health Service Research and Development. The U.S. Department of Veterans Affairs had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication. The findings and conclusions reported here are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the U.S. Department of Veterans Affairs or the U.S. government. No statement in this article should be construed as an official position of the U.S. Department of Veterans Affairs.

The authors report no financial relationships with commercial interests. Received August 19, 2016; revisions received December 14, 2016, and January 31, 2017; accepted March 10, 2017; published online June 15, 2017.

**REFERENCES**


51. Joffe P: An empirically supported program to prevent suicide in a college student population. Suicide and Life-Threatening Behavior 38:87–103, 2008


