Archives of Suicide Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/usui20

Suicide and Self-Harm Risk Assessment: A Systematic Review of Prospective Research

Mohamad M. Saab, Margaret Murphy, Elaine Meehan, Christina B. Dillon, Selena O'Connell, Josephine Hegarty, Sinead Heffernan, Sonya Greaney, Caroline Kilty, John Goodwin, Irene Hartigan, Maidy O'Brien, Derek Chambers, Una Twomey & Aine O'Donovan

To cite this article: Mohamad M. Saab, Margaret Murphy, Elaine Meehan, Christina B. Dillon, Selena O'Connell, Josephine Hegarty, Sinead Heffernan, Sonya Greaney, Caroline Kilty, John Goodwin, Irene Hartigan, Maidy O'Brien, Derek Chambers, Una Twomey & Aine O'Donovan (2021): Suicide and Self-Harm Risk Assessment: A Systematic Review of Prospective Research, Archives of Suicide Research, DOI: 10.1080/13811118.2021.1938321

To link to this article: https://doi.org/10.1080/13811118.2021.1938321

ARCHIVES OF SUICIDE RESEARCH
https://doi.org/10.1080/13811118.2021.1938321

Routledge
Taylor & Francis Group

Suicide and Self-Harm Risk Assessment: A Systematic Review of Prospective Research

Mohamad M. Saab, Margaret Murphy, Elaine Meehan, Christina B. Dillon, Selena O'Connell, Josephine Hegarty, Sinead Heffernan, Sonya Greaney, Caroline Kilty, John Goodwin, Irene Hartigan, Maidy O'Brien, Derek Chambers, Una Twomey, and Aine O'Donovan

ABSTRACT

Objective: Suicide and self-harm are widespread yet underreported. Risk assessment is key to effective self-harm and suicide prevention and management. There is contradicting evidence regarding the effectiveness of risk assessment tools in predicting self-harm and suicide risk. This systematic review examines the effect of risk assessment strategies on predicting suicide and self-harm outcomes among adult healthcare service users.

Method: Electronic and gray literature databases were searched for prospective research. Studies were screened and selected by independent reviewers. Quality and level of evidence assessments were conducted. Due to study heterogeneity, we present a narrative synthesis under three categories: (1) suicide- and self-harm-related outcomes; (2) clinician assessment of suicide and self-harm risk; and (3) healthcare utilization due to self-harm or suicide.

Results: Twenty-one studies were included in this review. The SAD PERSONS Scale was the most used tool. It outperformed the Beck Scale for Suicide Ideation in predicting hospital admissions and stay following suicide and self-harm, yet it failed to predict repeat suicide and self-harm and was not recommended for routine use. There were mixed findings relating to clinician risk assessment, with some studies recommending clinician assessment over structured tools, whilst others found that clinician assessment failed to predict future attempts and deaths.

Conclusions: There is insufficient evidence to support the use of any one tool, inclusive of clinician assessment of risk, for self-harm and suicidality. The discourse around risk assessment needs to move toward a broader discussion on the safety of patients who are at risk for self-harm and/or suicide.

KEYWORDS

Risk assessment; self-harm; suicidal ideation; suicide; systematic review

HIGHLIGHTS

- There is insufficient evidence to support using standalone risk assessment tools.
- There are mixed findings relating to clinician assessment of risk.
- Structured professional judgment is widely accepted for risk assessment.

Supplemental data for this article can be accessed at publisher's website.

INTRODUCTION

Suicide and self-harm tend to be under-reported, underappreciated, and affect every country and society worldwide (Oyesanya, Lopez-Morinigo, & Dutta, 2015; Pritchard & Hansen, 2015). It is estimated that 800,000 individuals die by suicide each year and many more utilize healthcare services for self-harm (World Health Organization, 2019). These figures may be underestimated due to legal, societal, and cultural taboos surrounding suicide and self-harm (Centers for Disease Control and Prevention, 2010). For instance, in the United States of America (USA), self-harm data are not collated centrally; however, the Centers for Disease Control and Prevention collect survey data, as well as hospital data on non-fatal injuries from self-harm. In 2015—the most recent year for which data are available—approximately 575,000 people attended a hospital for injuries due to self-harm (American Foundation for Suicide Prevention, 2020; Centers for Disease Control a Prevention, 2010). In risk assessment, utilizing near-miss information is key in preventing seminal or serious adverse events (Jeffs, Berta, Lingard, & Baker, 2012).

Risk screening and risk assessment have been identified as important components of effective self-harm and suicide management (Boudreaux et al., 2016; Jobes, 2012). Risk screening refers to the use of standardized instruments to identify at-risk individuals, whereas risk assessment refers to a more comprehensive evaluation to confirm suspected suicide and self-harm risk, estimate the immediate danger to the individual, and decide on risk management strategies (Suicide Prevention Resource Center, 2014). One study found that greater risk screening in emergency departments was associated with a significant increase in the detection of suicidal ideation (Boudreaux et al., 2016). Studies indicate that people who die by suicide have had contact with primary care services, emergency services and, to a lesser extent, mental health services in the month prior to their death (King, Horwitz, Czyz, & Lindsay, 2017; Luoma, Martin, & Pearson, 2002; Vasiliadis, Ngamini-Ngui, & Lesage, 2015). Therefore, universal self-harm and suicide risk screening and assessment were recommended across various healthcare settings, including primary care, specialty medical care, and emergency departments (King et al., 2017). Notwithstanding, there is no gold standard for suicide and self-harm risk assessment which tend to vary globally (Vasiliadis et al., 2015). For instance, in the USA, the US Preventive Services Task Force (2014) concluded that "current evidence is insufficient to assess the balance of benefits and harms of screening for suicide risk in adolescents, adults, and older adults in primary care." However, tools like Suicide Risk Screen, the Patient Health Questionnaire (PHQ), the SAFE-T tool, and the Columbia-Suicide Severity Rating Scale (C-SSRS) remain widely used in various healthcare settings in the USA (O'Rourke, Jamil, & Siddiqui, 2021).

In the international literature, a number of risk assessment tools have been used to measure self-harm and suicide risk such as the SAD PERSONS (SPS) and modified SPS (Chang & Tan, 2015); the Beck Suicide Intent Scale (SIS) (Jordan & McNiel, 2018); the Beck Scale for Suicide Ideation (SSI) (de Beurs et al., 2015); Manchester Self-Harm Rule (Quinlivan et al., 2017); the ReACT Self-Harm Rule (Quinlivan et al., 2017); among others. Previous literature reviews concluded that the available assessment tools did not reliably predict future risk of suicide (Runeson et al., 2017), repeat self-harm (Quinlivan

et al., 2017), or suicide following self-harm (Chan et al., 2016). Tools often performed well in terms of sensitivity or specificity but seldom both (Quinlivan et al., 2017; Runeson et al., 2017). To put this in context, for example, an assessment tool with a sensitivity of 85% will detect 85 out of every 100 individuals with the outcome, whereas 15 will be missed (i.e., false negatives). Similarly, an assessment tool with a specificity of 70% indicates that for every 100 individuals without the outcome, 30 will be wrongly categorized as having a risk for the outcome (i.e., false positives) (Bossuyt et al., 2008).

Risk assessment tools could potentially incorrectly identify people as having high risk, impacting resource usage, or conversely, may fail to identify individuals who are at high risk, compromising patient safety (Chan et al., 2016; Quinlivan et al., 2017; Runeson et al., 2017). The previous generation approach to risk assessment, including unstructured clinician risk assessment, has been recently evaluated in terms of predicting future risk of self-harm and was also found to be potentially inaccurate for clinical use (Woodford et al., 2019).

In the past, six systematic reviews (Chan et al., 2016; O'Shea & Dickens, 2014; Quinlivan et al., 2016; Runeson et al., 2017; Warden, Spiwak, Sareen, & Bolton, 2014; Woodford et al., 2019) and one narrative review (Thom, Hogan, & Hazen, 2020) evaluated how well multiple risk assessment tools predicted future suicide or self-harm in clinical practice. These reviews concluded that no single risk assessment tool was found to have enough evidence to support its routine use in clinical practice.

Some of the past reviews were limited by their focus on single instruments such as SPS (Warden et al., 2014), Short Term Assessment of Risk and Treatability (START) (O'Shea & Dickens, 2014), and unstructured clinician risk assessments (Woodford et al., 2019). Previous reviews also focused either on self-harm alone (Chan et al., 2016; Quinlivan et al., 2016; Woodford et al., 2019), suicide alone (Runeson et al., 2017; Warden et al., 2014; Thom et al., 2020), but seldom both (O'Shea & Dickens, 2014). From a methodological perspective, a number of past literature reviews did not provide a structured approach to searching the gray literature (Chan et al., 2016; Runeson et al., 2017; Thom et al., 2020; Warden et al., 2014), included studies published up until early 2014 (Chan et al., 2016; O'Shea & Dickens, 2014; Warden et al., 2014), and failed to address the methodological quality, level of evidence, and/or risk of bias within the reviewed studies (Thom et al., 2020).

For all the above reasons, a more up-to-date review of the empirical and gray literature would provide information on effective methods of suicide as well as self-harm risk assessment to identify those at risk of suicide and self-harm and ultimately offer appropriate support. Therefore, the aim of this systematic review was to examine the effect of risk assessment strategies on predicting suicide and self-harm outcomes among adult healthcare service users, with a focus on (i) suicide and self-harm related outcomes; (ii) clinician assessment of risk outcomes; and (iii) healthcare utilization outcomes.

METHODS

This systematic review was guided by the principles of conducting systematic reviews (Centre for Reviews and Dissemination, 2009), and reported according to the Preferred

Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (Moher, Liberati, Tetzlaff, Altman, & Prisma, 2009).

Eligibility Criteria

Review eligibility criteria were pre-determined using the PEO (Population, Exposure[s], and Outcome[s]) framework (Moola et al., 2015). Studies eligible for inclusion met the following criteria: Population: included adult (≥18 years of age) patients or service users who have a history of suicide or self-harm within any healthcare setting, including those with a history of any psychiatric and/or physical disorders which put them at risk for suicide and/or self-harm or repeat suicide and/or self-harm. Of note, in the context of the current review, repeat self-harm refers to individuals who have self-harmed in the past and present again with another episode of self-harm (Quinlivan et al., 2017); Exposure: involved the use of one or more instrument(s) to assess the risk of suicide or self-harm; and Outcome: followed service users up for varying lengths of time in order to evaluate the ability of risk assessment instruments to predict suicidal or self-harming ideations, suicide or self-harm attempts/behaviors, and death by suicide or self-harm. Notably, suicide and self-harm are related but not synonymous. Self-harm, also referred to as self-injury is defined as direct and deliberate harm to one's body often without intent to die. On the other hand, suicidal attempts and behaviors are often linked to an intention to cause death (Cipriano, Cella, & Cotrufo, 2017). Suicidality and self-harm have different prevalence rates, functions, clinical correlates, and outcomes yet they are often measured using the same instruments (Klonsky, May, & Saffer, 2016). Therefore, this review will explore and capture the risk assessment for both, suicide and self-harm intentions and behaviors.

Studies conducted among pediatric patients (<18 years of age), in non-healthcare settings, and focusing on interventions for self-harm or suicide prevention or management were excluded. Literature reviews, surveys, qualitative studies, policy documents, dissertations, conference proceedings, commentaries, and editorials were also excluded.

Information Sources and Search

The following electronic databases were searched: CINAHL; MEDLINE; APA PsycINFO; APA PsycARTICLES; Psychology and Behavioral Science Collection; ERIC; SocINDEX; and The Cochrane Library. Subject headings were used where appropriate and combined using Boolean operators "OR" and "AND," the proximity indicator for EBSCO "N," and truncation "*." The search was conducted on title or abstract as follows: Self-harm* OR "self harm*" OR self-poison* OR "self poison*" OR self-injur* OR "self injur*" OR self-mutilat* OR "self mutilat*" OR parasuicid* OR suicid* OR "suicid* idea*" OR DSH AND (risk N5 assess*) OR (risk N5 manag*).

A focused gray literature search was carried out and included customized Google and targeted website searches. This search was designed to source records from Australia, Canada, Ireland, New Zealand, the United Kingdom (UK), and USA. These countries were selected since they have similar health systems and infrastructure (Hegarty et al., 2020; United Nations Development Programme, 2019). Six separate Google searches

were conducted within these countries using the terms "suicide," "self-harm," and "risk" and the domains of the selected countries. The first ten pages, or 100 hits, were reviewed to capture the most relevant hits (Godin, Stapleton, Kirkpatrick, Hanning, & Leatherdale, 2015). Targeted websites included ministries of health and national organizations involved in suicide prevention in each of the selected countries (see Table S1 in supplemental file for the full list of websites). Electronic database and gray literature searches were last conducted in August 2019. All the searches were limited to records published in English between January 2014 and August 2019.

Study Selection

Records from electronic database and gray literature searches were exported to a reference management software (EndNote 7) and duplicates were deleted. Records were then transferred to Covidence, an online software package recommended by Cochrane to produce systematic reviews (Cochrane Community, 2020). Records were initially screened on title and abstract for relevance. The full texts of potentially eligible records were subsequently obtained and reviewed. Title, abstract, and full-text screenings were conducted independently by members of the review team. Screening conflicts were resolved by a third independent reviewer.

Data Extraction and Synthesis

The following were extracted for each study using a standardized data extraction table: Reference; country; design; sample; setting; instrument; follow-up; outcome; and findings (see Table S2 in supplemental file for the full data extraction table). Two reviewers conducted data extraction and each extracted study was cross-checked by a third reviewer for accuracy. Studies were synthesized to address the review aims and outcomes. A meta-analysis was not completed due to the use of several tools in single studies; adapted/shortened versions of tools; different cutoff scores to predict the risk of suicide/self-harm across different studies/groups; and various methodological approaches to measuring risk.

Quality and Level of Evidence Assessment

The Joanna Briggs Institute's (2017) critical appraisal checklist for cohort studies was used to determine whether individual studies have addressed potential biases in design, conduct, and analysis. The Scottish Intercollegiate Guidelines Network (SIGN) grading system was then used to assess the level of evidence for each of the included studies based on its design and quality (Healthcare Improvement Scotland, 2019). The eight levels of evidence range between 1++, 1+, 1−, 2++, 2+, 2−, 3, and 4. A score of 1++ corresponds to high quality meta-analyses, systematic reviews of randomized controlled trials, or randomized controlled trials with a very low risk of bias, whereas a score of 4 is assigned to expert opinions. Studies were included regardless of quality and level of evidence to minimize the risk of reporting bias.

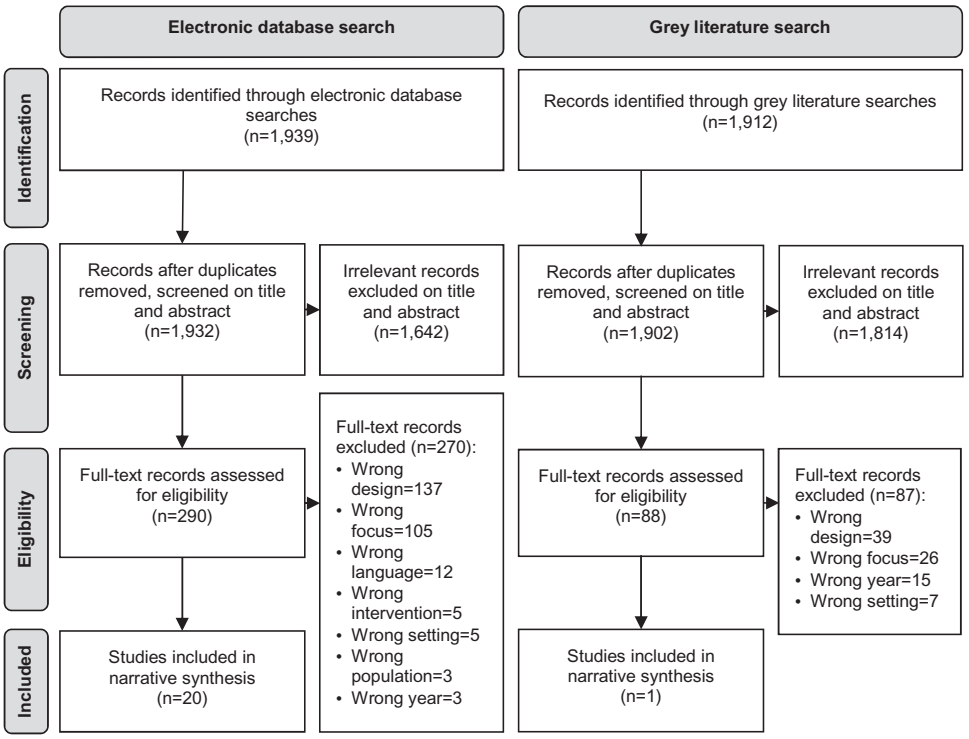

FIGURE 1. Study identification, screening, and selection process.

RESULTS

Study Selection

Electronic database searching yielded 1,939 records. Following deletion of duplicates, 1,932 records were screened based on title and abstract and 1,642 irrelevant records were excluded. The full texts of 290 records were reviewed and 270 records were excluded, resulting in 20 studies that were included from electronic databases. A total of 1,912 records were identified from the gray literature search. Titles and abstracts of 1,902 records were screened and 1,814 irrelevant records were excluded. Of the full texts screened ($n = 88$), only one study was eligible for inclusion. Therefore, a total of 21 studies were included in this review. See Figure 1 for the study identification, screening, and selection process.

Study Characteristics

Study characteristics are summarized in Table 1. Most of the studies were conducted in the USA ($n = 9$) and the UK ($n = 6$) using a prospective cohort design ($n = 17$). Sample sizes ranged between 50 (Chang & Tan, 2015) and 5,462 (Katz et al., 2017) participants. More than half of the reviewed studies were conducted in emergency departments ($n = 7$) and acute care settings ($n = 4$). Several instruments were used to assess suicide, with 13 studies using more than one instrument. The most frequently used instruments

TABLE 1. Study characteristics ($n = 21$).

Country	USA ($n = 9$)
	UK ($n = 6$)
	Canada ($n = 3$)
	Australia ($n = 1$)
	Sweden ($n = 1$)
	Taiwan ($n = 1$)
Study design	Prospective (cohort) study ($n = 17$)
	Pseudo-prospective cohort study ($n = 2$)
	Longitudinal prospective study ($n = 1$)
	Psychometric evaluation with follow-up ($n = 1$)
Sample size (min-max)	50–5,462
Settings	Emergency department ($n = 7$)
	Acute care/hospital/speciality not specified ($n = 4$)
	Forensic inpatient ($n = 2$)
	Inpatient mental health care ($n = 2$)
	Emergency department and general medicine ($n = 1$)
	Inpatient and community ($n = 1$)
	Liaison psychiatry service ($n = 1$)
	Maximum security facility ($n = 1$)
	Psychiatric outpatient ($n = 1$)
	Psychiatry services ($n = 1$)
Instruments[a]	SAD PERSONS (and modified/Chinese) Scale ($n = 6$)
	Beck Scale for Suicide Ideation ($n = 4$)
	Beck Hopelessness Scale ($n = 3$)
	Columbia Suicide Severity Rating Scale ($n = 2$)
	Concise Health Risk Tracking ($n = 2$)
	Historical, Clinical and Risk Management Scales ($n = 2$)
	Patient Health Questionnaire 9 ($n = 2$)
	Others ($n = 21$)
Follow-up (min-max)	2 weeks–20 years

[a]Several studies used more than one instrument; n corresponds to the number of times an instrument was used.

included SPS and modified SPS ($n = 6$), the Beck SSI ($n = 4$), and the Beck Hopelessness Scale (BHS) ($n = 3$). Follow-up times varied between 2 weeks (Chang & Tan, 2015) and 20 years (Green et al., 2015; Stefansson, Nordström, Runeson, Åsberg, & Jokinen, 2015) with almost half of the studies ($n = 10$) reporting a 6-month follow-up.

Quality and Level of Evidence Assessment

All 21 studies used valid exposure measures and reliable outcome measures, but 10 failed to adequately identify or address potential confounders. All the studies were observational and all, but one rated as level 2+ on the SIGN level of evidence criteria, indicating well-conducted cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal. Quality and level of evidence assessment are outlined in Table 2.

Synthesis of Results

Most studies used estimates of sensitivity and specificity or areas under the curve (AUC) to indicate the predictive validity of the risk assessment tools. Outcomes measured in the 21 studies are divided into three categories: (i) suicide and self-harm-related outcomes; (ii) clinician assessment of suicide and self-harm risk; and (iii) outcomes

TABLE 2. Quality appraisal and level of evidence assessment ($n = 21$).

Reference	Similar groups from same population	Similar exposure for study groups	Valid and reliable exposure measure	Confounders identified	Confounders addressed	Groups free of outcome at start	Valid and reliable outcomes measure	Follow-up time reported	Follow-up complete	Strategies for incomplete follow-up	Suitable statistical analysis	Level of evidence
Brucker et al. (2019)	NA	NA	Y	U	U	Y	Y	Y	Y	NA	Y	2+
Campbell and Beech (2018)	NA	NA	Y	Y	Y	Y	Y	Y	Y	NA	Y	2+
Chang and Tan (2015)	NA	NA	Y	U	U	Y	Y	Y	Y	NA	Y	2+
de Beurs et al. (2015)	NA	NA	Y	Y	Y	Y	Y	Y	Y	NA	Y	2+
Dickens and O'Shea (2015)	NA	NA	Y	Y	Y	Y	Y	Y	Y	NA	Y	2+
Glassmire et al. (2016)	NA	NA	Y	Y	Y	Y	Y	Y	Y	NA	Y	2+
Green et al. (2015)	NA	NA	Y	Y	Y	Y	Y	Y	Y	NA	Y	2+
Harrison et al. (2018)	NA	NA	Y	U	U	Y	Y	Y	Y	NA	Y	2+
Hawes et al. (2017)	NA	NA	Y	U	U	Y	Y	Y	Y	U	Y	2+
Jordan and McNiel (2018)	NA	NA	Y	Y	Y	Y	Y	Y	Y	NA	Y	2+
Katz et al. (2017)	NA	NA	Y	U	U	N	Y	Y	Y	Y	Y	2+
Madan et al. (2016)	NA	NA	Y	U	U	U	Y	Y	N	U	Y	2-
O'Shea et al. (2014)	NA	NA	Y	Y	Y	Y	Y	Y	Y	NA	Y	2+
Quinlivan et al. (2017)	NA	NA	Y	U	U	U	Y	Y	Y	U	Y	2+
Randall et al. (2019)	NA	NA	Y	Y	Y	Y	Y	Y	Y	U	Y	2+
Reilly-Harrington et al. (2016)	NA	NA	Y	U	U	U	Y	Y	U	U	Y	2+
Saunders et al. (2014)	NA	NA	Y	U	U	U	Y	Y	Y	U	Y	2+
Stefansson et al. (2015)	NA	NA	Y	Y	Y	Y	Y	Y	Y	NA	Y	2+
Villegas et al. (2018)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	2+
Wang et al. (2016)	NA	NA	Y	U	U	U	Y	Y	Y	Y	Y	2+
Wu et al. (2014)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	2+

N: no; NA: not applicable; U: unclear; Y: yes.

TABLE 3. Summary of findings from individual studies ($n = 21$).

Reference	Country	Design	Sample	Setting	Instrument	Follow-up	Outcome	Findings
Brucker et al. (2019)	USA	Prospective	338	ED	SCS, Physician gestalt VAS, CFI-S	6 mon	Suicidality spectrum	SCS: modest diagnostic accuracy sensitivity 14/32 = 44%, (95%CI 26–62%) and specificity 275/306 = 90%, (95%CI 86–93%); VAS: moderate diagnostic accuracy (AUC = 0.75, 95%CI 0.66–0.85); CFI-S: Best for diagnostic accuracy (AUC = 0.81, 95%CI 0.76–0.87).
Campbell and Beech (2018)	UK	Prospective	89	Forensic inpatient	HCR-20, FAM	1 yr	Self-harm	Positive association between mean scores on HCR-20 and frequency of self-harm, with ($p < 0.001$) and without FAM ($p < 0.001$).
Chang and Tan (2015)	USA	Prospective	50	ED	PHQ-9, BSSI, C-SSRS, SPS	2 wk	Admission Length of stay Adverse events	19 (38%) $\geq$ 1 adverse event. BSSI, PHQ-9, C-SSRS poor at predicting outcomes (AUC $\leq$ 0.5). SPS better at predicting admission ($p = 0.009$) and stay ($p = 0.006$) but not ED adverse events. Positive SPS score: 8.18 times the odds of admission.
de Beurs et al. (2015)	UK	Prospective	366	Liaison psychiatry	BSSI	15 mon	Repeat suicide	94 (26%) treated after repeat suicide. Depression (OR = 1.04, 95%CI 1.02–1.06, $p < 0.001$), hopelessness (OR = 1.07, 95%CI 1.02–1.13, $p < 0.006$), suicidal ideations higher among attempters vs non-attempters (M = 23 ± 8.4 vs M = 18 ± 10.6, $p < 0.001$).
Dickens and O'Shea (2015)	UK	Pseudo-prospective	217	Secure inpatient	START	3 mon	Self-harm Suicide	Survival rates differed between groups rated as low and moderate risk ($\chi^2 = 24.94$, df = 1, $p < 0.001$), and between low- and high-risk groups ($\chi^2 = 33.29$, df = 1, $p < 0.001$), but did not between moderate- and high-risk groups ($\chi^2 = 1.591$, df = 1, $p = 0.207$).
Glassmire et al. (2016)	USA	Prospective	229	Maximum security facility	MMPI-2-RF SUI items	1 yr	Suicidal ideation and behavior	80 (34.9%) endorsed $\geq$1 item. Endorsing 2,3,4–5 items:10,12,14 times greater risk. Past attempt and ideation, and current ideation: 5,6,10 times greater risk. Endorsing $\geq$2 items: 3–9 times greater risk. Having suicidal ideation and endorsing $\geq$1 item: 15–22 times greater risk.
Green et al. (2015)	USA	Prospective	S1: 5,200 S2: 119	Psychiatric outpatient	Diagnostic interviews, BDI, BSSI	S1: 20 yr S2: 18 mon	Repeat suicide Death by suicide	BDI predicted suicide in S1 (Wald $X^2_1 = 35.67$, $p < 0.001$, HR = 2.79, 95%CI 2–3.91) and S2 (Wald $X^2_1 = 8.82$, $p < 0.01$, HR = 1.69, 95%CI 1.19–2.38) BDI predicted death by suicide ($p < 0.01$) in both samples. AUC for BDI suicide = 0.63 and for BSSI = 0.67.
Harrison et al. (2018)	Australia	Prospective	128	ED	d/s-IAT	3 mon 6 mon	Suicidal ideations and attempts	Clinician and patient prediction accounted for 4.9% and 7.2% of variance in ideation ($p = 0.049$ and $p = 0.011$) at 3 and 6 mon, respectively. Clinician prediction as the only significant

(*continued*)

TABLE 3. Continued.

Reference	Country	Design	Sample	Setting	Instrument	Follow-up	Outcome	Findings
								predictor of ideation, unaffected by adding d/s-IAT. Clinician and patient prediction and d/s-IAR did not significantly predict attempts.
Hawes et al. (2017)	USA	Prospective	59	Inpatient	MARIS	4–8 wk	Suicidal ideations and attempts	Correlation between MARIS score, lifetime attempts (rho = 0.30, $p = 0.005$), depression (rho = 0.46, $p < 0.001$), lifetime ideations (rho = 0.25, $p = 0.023$), and ideations in past month (rho = 0.35, $p = 0.001$). Those with behaviors had higher MARIS scores than those who did not (MD = 15.69,2.96; Cohen's $d = 1.54,0.77$; $U = 33,119.5$; $p = 0.001,0.036$, respectively).
Jordan and McNiel (2018)	USA	Prospective	218	Inpatient	SIS	1 yr	Repeat suicide attempt	69 (31.7%) attempted suicide. Cross-validated iterative receiver operator curve analysis of SIS data predicted subsequent attempts (AUC = 0.43, 95%CI 0.34–0.58, sensitivity = 61.54%, specificity = 56.91%, PPV = 37.65%, NPV = 77.78%).
Katz et al. (2017)	Canada	Longitudinal prospective	5,462	ED	SPS, MSPS	6 mon 1 yr 2 yr	Death by suicide	77 (1.4%) died by suicide. Of those, half were assessed as low risk at 6 and 12 mon. AUC for SPS and MSPS was 0.56 (95%CI 0.45–0.67) and 0.59 (95%CI 0.49–0.69). Scaled Brier score for SPS and MSPS indicated poor performance.
Madan et al. (2016)	USA	Prospective	1,055	Specialised hospital	C-SSRS, PHQ-9, BHS, SCogS	6 mon	Suicide-related behavior	C-SSRS correctly classified suicide-related behavior within 6 mon of discharge (AUC = 0.757, $p < 0.001$) with total and summary score from the ideation/behavior factor providing balance between sensitivity (0.694) and specificity (0.652–0.674).
O'Shea et al. (2014)	UK	Pseudo-prospective	504	Forensic inpatient	HCR-20	3 mon	Violence risk self-harm	HCR-20 scores higher in moderate- (M = 28.90 ± 4.78, $p < 0.001$) and high-risk groups (M = 29.52 ± 5.45, $p < 0.001$) compared to low-risk group (M = 24.66 ± 6.30). HCR-20 predicted self-harm but effect sizes not large (0.345–0.749).
Quinlivan et al. (2017)	UK	Prospective	464	Psychiatry services	MSHR, ReACT SHR, SPS, MSPS, BIS	6 mon	Repeat self-harm	145 (30%) repeat self-harm. Sensitivity 1% (95%CI 0–5) for SPS and 97% (95%CI 93–99) for MSHR. PPV 13% (95% CI 2–47) for MSPS and 47% (95%CI 41–53) for clinician assessment. Remaining scales performed worse than clinician and patient estimates.
Randall et al. (2019)	Canada	Prospective	2,643	ED	SLS, SALS, C-CASA	6 mon	Suicide attempt Death by suicide	56 (2.1%) attempts, 13 (0.5%) deaths at 6 mon and 85 (3.2%) attempts, 20 (0.8%) deaths at 12 mon. Clinician assessment not effective at predicting deaths (AUC = 0.546, 95%CI

								0.36–0.73) and moderately accurate for attempts (AUC = 0.728, 95%CI 0.66–0.79). C-CASA moderately accurate at predicting attempts and deaths (AUC = 0.666 and 0.678).
Reilly-Harrington et al. (2016)	USA	Prospective	482	Acute care	CHRT	6 mon	Suicide-related serious adverse event	26 (5%) experiencing ≥ 1 adverse events had a sig higher CHRT score than those without adverse events (M = 29.9 ± 11.7 vs M = 23.7 ± 9.6, $p = 0.003$) this trend remained at follow-up (HR = 1.76, 95%CI 1.01–3.05, $p = 0.04$). CHRT highly correlated with clinician ratings of depression, anxiety, and overall functioning.
Saunders et al. (2014)	UK	Prospective	126	Hospital	SPS	6 mon	Repeat self-harm Admission Aftercare	5 (4%) psychiatric hospital admissions, 70 (55.5%) community aftercare, and 31 (24.6%) repeated self-harm. Specificity of SPS > 90% for all outcomes, yet only 2% sensitivity for admission, 5.8% for community aftercare, and 6.6% for repetition of self-harm.
Stefansson et al. (2015)	Sweden	Prospective	81	University hospital	SIS, KIVS	15–20 yr	Suicide	SIS alone: 52% specificity, 17% PPV; SIS and KIVS combined assessment: 63% specificity, 19% PPV; SIS and KIVS together: 83% sensitivity, 80% specificity, 26% PPV
Villegas et al. (2018)	USA	Prospective	109	Inpatient psychiatry	CHRT-SR12, CHRT-SR7, BHS, QIDS-SR, LOT-R, PANAS	6 wk 12 wk	Self-injury Suicide attempt	Decreased suicide risk using CHRT-SR12 ($b = 0.24$, $p = 0.009$) and CHRT-SR7 ($b = 0.16$, $p = 0.004$). Decreased hopelessness and depression using BHS ($b = 0.21$, $p \leq 0.001$) and QIDS ($b = 0.33$, $p \leq 0.001$)
Wang et al. (2016)	Canada	Prospective	2,792	ED	Clinician prediction scale, SPS	6 mon	Suicide attempts	136 (4.9%) attempted suicide within 6 mon. Clinicians were able to predict future attempts with sig greater accuracy (AUC = 0.73, 95%CI 0.68–0.77, $p < 0.001$) in comparison to SPS ($z = 3.79$, $p < 0.001$). Both showed PPV <7%
Wu et al. (2014)	Taiwan	Psychometric evaluation	147 cases 284 controls	ED Medicine	Chinese SPS, PSIS, HAMD, BSSI, BHS	6 mon	Repeat self-harm	SPS: AUC = 0.66, $p = 0.02$, 95%CI 0.53–0.79; PSIS: AUC = 0.49, $p = 0.88$, 95%CI 0.36–0.61; HAMD: AUC = 0.56, $p = 0.40$, 95%CI 0.42–0.70; BSSI: AUC = 0.59, $p = 0.18$, 95%CI 0.45–0.74; BHS: AUC = 0.55, $p = 0.47$, 95%CI 0.42–0.70. SPS cutoff point of 4/5 determined to predict repeat self-harm with 65.4% sensitivity and 58.1% specificity

AUC: Area under the curve; b: beta; BDI: Beck Depression Inventory; BHS: Beck Hopelessness Scale; BIS: Barratt Impulsiveness Scale; BSSI: Beck Scale for Suicide Ideation; C-CASA: Columbia Classification Algorithm of Suicide Assessment; CFI-S: Convergent Functional Information for Suicidality; CHRT: Concise Health Risk Tracking; CI: confidence interval; C-SSRS: Columbia Suicide Severity Rating Scale; d/s-IAT: Death/Suicide Implicit Association Test; ED: emergency department; FAM: Female Additional Manual; HAMD: Hamilton Rating Scale for Depression; HCR-20: Historical: Clinical and Risk Management Scales; HR: hazard ratio; IQR: interquartile range; KIVS: Karolinska Interpersonal Violence Scale; LOT-R: Life Orientation Test-Revised; M: mean; MARIS: Modular Assessment of Risk for Imminent Suicide; MMPI-2-RF: Minnesota Multiphasic Personality Inventory–2 Restructured Form; MSHR: Manchester Self-Harm Rule; MSPS: Modified SAD PERSONS Scale; NPV: negative predictive value; OR: odds ratio; PANAS: Positive and Negative Affect Schedule; PHQ-9: Patient Health Questionnaire 9; PPV: positive predictive value PSIS: Pierce Suicide Intent Scale; QIDS-SR: Depressive Symptomatology Self-Report; S1: sample 1; S2: sample 2; SALS: Suicide Attempt Likelihood Scale; SCogS: Suicide Cognitions Scale; SCS: Suicide Clinical Screening; SHR: Self-Harm Rule; SIS: Suicide Intent Scale; SLS: Suicide Likelihood Scale; SPS: SAD PERSONS Scale; START: Short Term Assessment of Risk and Treatability; SUI: Suicidal/Death Ideation; UK: United Kingdom; USA: United States of America; VAS: Physician gestalt Visual Analog Scale; χ^2: chi square; z: standard score.

related to the number or frequency of episodes of healthcare utilization due to self-harm or suicide. A summary of findings from individual studies is presented in Table 3.

Suicide and Self-Harm Related Outcomes

Across the six studies that evaluated SPS, or modified SPS, sensitivity for repeat self-harm ranged widely from 1% (Quinlivan et al., 2017) to 65% (Wu et al., 2014), while specificity for repeat self-harm ranged from 7% (Saunders, Brand, Lascelles, & Hawton, 2014) to 58% (Wu et al., 2014). Wu et al. (2014) found the Chinese SPS useful in identifying high-risk individuals. However, five other studies did not support the use of SPS to predict suicide or repeat self-harm and recommended against using SPS to screen patients presenting to hospitals with self-harm (de Beurs et al., 2015; Katz et al., 2017; King et al., 2017; Saunders et al., 2014; Wang et al., 2016).

While de Beurs et al. (2015) found that most items on the Beck SSI were significant predictors of a repeat suicide attempt within 15 months ($p < 0.05$), Wu et al. (2014), using AUCs, found that the Beck SSI performed significantly poorer than the Chinese SPS in predicting repeat self-harm within 6 months (Chinese SPS: AUC = 0.66, $p = 0.02$; Beck SSI: AUC = 0.59, $p = 0.18$). Green et al. (2015) reported that the Beck Depression Inventory (BDI) performed better than the Beck SSI at predicting suicide and repeat attempts (sensitivity of 81% vs 53% respectively); however, in this study, the Beck SSI was found to be better than the BDI in correctly identifying true negatives (specificity of 83% vs 54%, respectively). Given that the BDI suicide item was associated with the risk of repeat suicide attempts and death by suicide, this tool was recommended for use in routine clinical care, coupled with comprehensive clinician suicide risk assessment for a positive screen (Green et al., 2015).

The Beck SIS was used in two studies either alone (Jordan & McNiel, 2018), or with the Karolinska Interpersonal Violence Scale (Stefansson et al., 2015). When used alone, the Beck SIS predicted subsequent suicide attempts with 61.54% sensitivity, 56.91% specificity, 37.65% positive predictive value, and 77.78% negative predictive value (AUC = 0.43, 95%CI 0.34–0.58) (Jordan & McNiel, 2018). Another study found that Beck SIS alone had 52% specificity and 17% positive predictive value; however, when used together with the Karolinska Interpersonal Violence Scale, sensitivity was increased at 83%, specificity at 80%, and positive predictive value at 26% (Stefansson et al., 2015).

The Historical, Clinical and Risk (HCR-20) Management scale was evaluated in two studies (Campbell & Beech, 2018; O'Shea, Picchioni, Mason, Sugarman, & Dickens, 2014). It was found that higher mean total scores on HCR-20 were associated with more frequent self-harm ($p < 0.001$) (Campbell & Beech, 2018); however, effect sizes were not large enough (0.345–0.749) to support the use of HCR-20 in practice (O'Shea et al., 2014).

Madan et al. (2016) reported findings that provide some support for the reliability and validity of the C-SSRS related to its potential to correctly predict suicide-related behavior ($p < 0.01$). The authors recommended using the total C-SSRS score and the summary score from the ideation/behavior factor together in order to find the best balance between sensitivity (69%) and specificity (65–67%) (Madan et al., 2016). The Columbia Classification Algorithm for Suicide Assessment scale (C-CASA) was found

in one study to be moderately accurate at predicting suicide attempts (AUC = 0.666) and deaths from suicide (AUC = 0.678) (Randall, Sareen, Chateau, & Bolton, 2019).

In two studies, the self-report versions of Concise Health Risk Tracking (CHRT) showed good internal consistency and were strongly correlated with subsequent suicide risk (Reilly-Harrington et al., 2016; Villegas et al., 2018). In one study, the likelihood of a suicide-related event increased by 76% for every 10-point increase in baseline self-report CHRT scores (Reilly-Harrington et al., 2016). CHRT scores were also shown to be highly correlated with clinician ratings of depression, anxiety, and overall functioning. Therefore, the CHRT was recommended as a quick and robust self-report tool for assessing suicide risk. Similarly, Hawes, Yaseen, Briggs, and Galynker (2017) found a significant correlation between Modular Assessment of Risk for Imminent Suicide (clinician- and self-report tool) score and lifetime suicide attempts (rho = 0.30, $p = 0.005$), depression (rho = 0.46, $p < 0.001$), lifetime suicidal ideations (rho = 0.25, $p = 0.023$), and suicidal ideations in the past month (rho = 0.35, $p = 0.001$). In addition, those who attempted suicide were found to have higher scores than those who did not (Mean difference = 15.69,2.96; Cohen's $d = 1.54,0.77$; $U = 33,119.5$; $p = 0.001,0.036$, respectively).

Using the 5 Minnesota Multiphasic Personality Inventory–2–Restructured Form, Suicidal/Death Ideation (SUI) items, it was found that the SUI scale demonstrated statistically significant associations ($p < 0.05$), with interview-reported history of suicide attempts ($r = 0.35$) and the total number of suicidal behaviors within one year of testing ($r = 0.28$) (Glassmire, Tarescavage, Burchett, Martinez, & Gomez, 2016). Moreover, Glassmire et al. (2016) found that endorsing SUI items was significantly associated with greater risk for suicide. This supports the use of SUI-item endorsement and interview-reported risk information as predictors for future suicide.

In terms of survival rate post-self-harm among different risk groups, the use of START yielded survival rates that differed significantly between groups rated as low- and moderate-risk ($p < 0.001$), and between low- and high-risk groups ($p < 0.001$) but did not between moderate- and high-risk groups ($p = 0.207$) (Dickens & O'Shea, 2015).

Clinician Assessment of Risk

There were mixed findings relating to clinician assessment of risk. Quinlivan et al. (2017) evaluated the performance of multiple tools (Manchester Self-Harm Rule, ReACT Self-Harm Rule, SPS, modified SPS, and Barratt Impulsiveness Scale) in comparison to clinician estimates of risk following self-harm. AUCs ranged from 0.55 (95%CI 0.50–0.61) for SPS to 0.74 (95%CI 0.69–0.79) for the clinician global estimation of risk scale, indicating that this scale performed better than the SPS in estimating risk for repeat self-harm. The remaining scales performed significantly worse, in comparison to clinician estimates. Similarly, Wang et al. (2016) found that clinicians were able to predict future attempts with significantly greater accuracy in comparison to SPS ($p < 0.001$).

In contrast, Harrison, Stritzke, Fay, and Hudaib (2018) reported that clinician prediction did not significantly predict future attempts at three- and six-month follow-up ($p = 0.16$ and $p = 0.30$, respectively), despite significantly predicting suicidal ideations at both timepoints ($p = 0.049$ and $p = 0.011$, respectively). Another study found that, while clinician assessment of risk was moderately accurate at predicting future suicide

attempts (AUC = 0.728, 95%CI 0.66–0.79), it was not effective at predicting deaths from suicide (AUC = 0.546, 95%CI 0.36–0.73) (Randall et al., 2019). Moreover, clinician assessment was not significantly better at assessing the risk of suicide in comparison to the C-CASA classification system. Likewise, the Convergent Functional Information for Suicidality tool had the best diagnostic accuracy (AUC = 0.81, 95%CI 0.76–0.87) in comparison to clinician prediction of risk, which had modest diagnostic accuracy (Randall et al., 2019).

Only one study conducted analyses by level of clinician training (Wang et al., 2016). It was found that clinicians' ability to predict future suicidal attempts with greater accuracy as compared to traditional risk assessment instruments was linked to their level of seniority, with senior psychiatric residents and staff psychiatrists demonstrating greater accuracy than junior psychiatric residents (AUC = 0.78 vs 0.76 respectively, $p < 0.001$).

Healthcare Utilization Outcomes

Chang and Tan (2015) investigated the ability of C-SSRS, Beck SSI, SPS, and the Patient Health Questionnaire 9 (PHQ-9) to predict adverse events in the emergency department, following a presentation for suicidal ideation. They found that SPS was significantly better at predicting hospital admission ($p = 0.009$) and stay ($p = 0.006$) but not near-term adverse events in the emergency department. These included the "need for unscheduled psychiatric or sedating medications, physical restraints, or security staff intervention" (Chang & Tan, 2015; p.1681). The remaining instruments demonstrated poor predictive value for adverse events in the emergency department and psychiatric admissions. Likewise, Saunders et al. (2014) found that SPS failed to identify most patients who presented to the emergency department following self-harm and went on to require psychiatric hospital admission or community psychiatric aftercare. Both studies concluded that currently available suicide risk assessment tools should not be routinely used in the emergency department to identify those at greatest risk (Chang & Tan, 2015; Saunders et al., 2014).

DISCUSSION

This systematic review examined the effect of suicide and self-harm risk assessment tools on predicting suicide and self-harm outcomes among adult healthcare service users. Overall, limited evidence was found to support the use of standalone risk assessment tools in healthcare settings. Of the 21 included studies, six evaluated SPS or modified SPS. All studies, except for one (Wu et al., 2014), advised against the use of SPS to screen patients presenting to hospitals with self-harm. Various other scales were evaluated including the Beck SSI, the Beck SIS, BDI, HCR-20, C-SSRS, C-CASA and CHRT scales, with promising, yet limited and weak evidence relating to their sensitivity and specificity. It was also found that combining two or more risk assessment tools was more effective than using a single tool (Glassmire et al., 2016; Reilly-Harrington et al., 2016; Stefansson et al., 2015), and that self-report measures can be potentially effective in predictive future suicide and self-harm (Glassmire et al., 2016; Reilly-Harrington et al., 2016; Villegas et al., 2018). Furthermore, studies measuring healthcare utilization

outcomes advised against using suicide risk assessment tools such as SPS, the Beck SSI, PHQ-9, and C-SSRS routinely in emergency departments (Chang & Tan, 2015; Saunders et al., 2014).

Findings from the seven reviews discussed in the introduction support findings from our current review (Chan et al., 2016; O'Shea & Dickens, 2014; Quinlivan et al., 2016; Runeson et al., 2017; Thom et al., 2020; Warden et al., 2014; Woodford et al., 2019). Overall, there was insufficient evidence to support the use of SPS and START in assessing or predicting suicidal behavior (O'Shea & Dickens, 2014; Warden et al., 2014). In fact, SPS and modified SPS repeatedly failed to identify patients requiring psychiatric admission or community psychiatric aftercare, predict repetition of self-harm, and accurately predict future suicide attempts (Bolton, Spiwak, & Sareen, 2012; Stefansson et al., 2015). Therefore, SPS was judged as not being of clinical value and should not be used alone to assess for self-harm risk in acute care. It was also found that unstructured clinician risk assessment was too inaccurate to be clinically useful, and that after-care should be allocated based on a need rather than risk assessment (Woodford et al., 2019).

Structured professional judgment is a widely accepted approach to clinical risk assessment and management (Fagan et al., 2009). It is considered as the third generation of risk assessment, combining unstructured clinical judgment (first generation) and actuarial assessment (second generation) (Higgins, Morrissey, Doyle, Bailey, & Gill, 2015). Structured clinical judgment frameworks can assist practitioners in moving beyond the use of intuition and risk assessment tools; however, such frameworks are not elaborated upon in detail to provide sound clinical guidance for practitioners (Higgins et al., 2015).

A recent review by Hanratty, Kilicaslan, Wilding, and Castle (2019) found limited evidence regarding the effectiveness of Collaborative Assessment and Management of Suicidality in reducing suicide risk and deliberate self-harm in adults. However, evidence from the present review was divided between studies favoring clinician assessment of risk (Quinlivan et al., 2017; Wang et al., 2016), and others where clinician assessment of risk did not significantly predict suicide attempts (Brucker et al., 2019; Harrison et al., 2018), or death from suicide (Randall et al., 2019).

Implications and Recommendations

Data to support the utilization of risk assessment tools and their impact on predicting suicide and self-harm are sparse, therefore the use of risk assessment tools in isolation as a predictor needs to be recognized. Indeed, no one scale was found to have sufficient evidence to support its use in clinical practice. It is argued that contemporary discourse in the patient safety literature on risk assessment tools needs to shift to reflect this lack of empirical evidence. The focus on risk assessment tools may be deterring the development of sound clinical judgment frameworks. Furthermore, risk assessment without the development and implementation of clinical judgment frameworks is an arbitrary practice and a shift in paradigm across all healthcare sectors is needed. Kapur and Goldney (2019) argue that clinicians need to urgently recognize the "fallacy" of risk assessment, recognizing that assessment tools are more likely to be serving the organization instead of the patient.

While not meeting the criteria for inclusion in this systematic review, a number of best practice and policy guidelines for the assessment of risk were sourced from the gray literature search. Overall, it was clear that for any recommendations relating to the assessment of suicide and self-harm risk to be implemented, a whole system, multi-agency, and collaborative approach is needed (Department of Health & Human Services, 2016; Health Service Executive, 2017; Queensland Mental Health Commission, 2015; Ridani et al., 2016; SANE Australia, 2014; Welsh Government, 2015). However, while these recommendations were made in the policy and guidance documents internationally, there was a clear lack of specificity as to how to implement the recommendations in practice. In addition, no single model of risk assessment was discussed in more than one document, which supports findings from our review.

It is recommended that research needs to move beyond trying to determine the efficacy of risk assessment tools as predictors of self-harm and suicide. As corroborated by this latest review, there is insufficient evidence to support the use of risk assessment tools as a standalone assessment method.

Strengths and Limitations

To the best of the authors' knowledge, this is the most up-to-date systematic review to evaluate and compare various suicide and self-harm risk assessment tools, inclusive of clinician assessment of risk. Rigor was sought in the conduct and reporting of this review and studies were sourced from various electronic databases and the gray literature. Moreover, record screening, data extraction, and quality appraisal were cross-checked by independent reviewers to ensure accuracy and minimize the risk of reporting bias.

Given this review is limited to prospective studies, some publications may have been missed. Most of the included studies were well-conducted cohort studies with a low risk of confounding or bias. However, quality appraisal of the included studies determined that, while all studies used valid exposure and reliable outcome measures, almost half of the studies inadequately identified or addressed potential confounders. The challenges of appropriate confounding control are particularly problematic in such studies, as exposure is established by a complex interaction between various patient, physician, and healthcare system factors and information (Brookhart, Stürmer, Glynn, Rassen, & Schneeweiss, 2010). While outcomes in some of the reviewed studies were described as self-harm, it was unclear how this term was operationalized and whether there were any distinctions made between suicidal and non-suicidal self-harm.

CONCLUSION

Findings from this systematic review indicate that there is insufficient evidence to support the use of any one clinical risk assessment tool, inclusive of clinician assessment of risk, for self-harm and suicidality in clinical settings. This review also found limited evidence pertaining to the effect of risk assessment on healthcare utilization due to self-harm or suicide. As such, it is timely that the discourse in relation to risk assessment moves toward a broader discussion on the safety of patients who have suicidal ideation

and those who attempt self-harm or suicide. Findings from this review underscore the need to develop and evaluate clinical judgment frameworks that are evidence-based, and responsive to individual patient needs.

ACKNOWLEDGEMENTS

The authors would like to thank members of the Connecting for Life project working group steering group in Ireland.

DISCLOSURE STATEMENT

There are no relevant financial or non-financial competing interests to report.

AUTHOR NOTES

Mohamad M. Saab, Margaret Murphy, and Elaine Meehan, Catherine McAuley School of Nursing and Midwifery, University College Cork, Cork, Ireland. Christina B. Dillon, Environmental Research Institute/School of Public Health, University College Cork, Cork, Ireland. Selena O'Connell, Josephine Hegarty, and Sinead Heffernan, Catherine McAuley School of Nursing and Midwifery, University College Cork, Cork, Ireland. Sonya Greaney, Southern Area Health Service Executive, Cork, Ireland. Caroline Kilty, John Goodwin, Irene Hartigan, and Maidy O'Brien, Catherine McAuley School of Nursing and Midwifery, University College Cork, Cork, Ireland. Derek Chambers and Una Twomey, Southern Area Health Service Executive, Cork, Ireland. Aine O'Donovan, Catherine McAuley School of Nursing and Midwifery, University College Cork, Cork, Ireland.

Correspondence concerning this article should be addressed to Dr. Aine O'Donovan, Catherine McAuley School of Nursing and Midwifery, University College Cork, Cork, Ireland. Email: aine.odonovan@ucc.ie

FUNDING

This work was supported by Ireland's Health Service Executive (HSE)—Mental Health Section.

ORCID

Mohamad M. Saab http://orcid.org/0000-0002-7277-6268
Aine O'Donovan http://orcid.org/0000-0001-6377-4140

DATA AVAILABILITY STATEMENT

The authors confirm that the data supporting the findings of this review are available within the article and its supplementary materials.

REFERENCES

American Foundation for Suicide Prevention. (2020, March 1). *Suicide statistics*. Retrieved from https://afsp.org/suicide-statistics/.

Bolton, J. M., Spiwak, R., & Sareen, J. (2012). Predicting suicide attempts with the SAD PERSONS scale: A longitudinal analysis. *The Journal of Clinical Psychiatry*, *73*(06), e735–e741. doi:10.4088/JCP.11m07362

Bossuyt, P., Davenport, C., Deeks, J., Hyde, C., Leeflang, M., & Scholten, R. (2008). Cochrane handbook for systematic reviews of diagnostic test accuracy. *The Cochrane Collaboration*. Retrieved from: https://methods.cochrane.org/sites/methods.cochrane.org.sdt/files/public/uploads/DTA%20Handbook%20Chapter%2011%20201312.pdf.

Boudreaux, E. D., Camargo, C. A., Arias, S. A., Sullivan, A. F., Allen, M. H., Goldstein, A. B., … Miller, I. W. (2016). Improving suicide risk screening and detection in the emergency department. *American Journal of Preventive Medicine*, *50*(4), 445–453. doi:10.1016/j.amepre.2015.09.029

Brookhart, M. A., Stürmer, T., Glynn, R. J., Rassen, J., & Schneeweiss, S. (2010). Confounding control in healthcare database research: Challenges and potential approaches. *Medical Care*, *48*(6), S114–S120. doi:10.1097/MLR.0b013e3181dbebe3

Brucker, K., Duggan, C., Niezer, J., Roseberry, K., Le-Niculescu, H., Niculescu, A. B., & Kline, J. A. (2019). Assessing risk of future suicidality in emergency department patients. *Academic Emergency Medicine*, *26*(4), 376–383. doi:10.1111/acem.13562

Campbell, L., & Beech, A. (2018). Do scores on the HCR-20 and FAM predict frequency of self-harm in females within a secure psychiatric hospital? *The Journal of Forensic Psychiatry & Psychology*, *29*(6), 914–933. doi:10.1080/14789949.2018.1477975

Centers for Disease Control and Prevention. (2010, November 16). *Defining suicide*. Injury Prevention & Control: Division of Violence Prevention. Retrieved from: https://vetoviolence.cdc.gov/defining-suicide.

Centre for Reviews and Dissemination. (2009). *CRD's guidance for undertaking reviews in healthcare*. York Publ. Services. Retrieved from: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf.

Chan, M. K. Y., Bhatti, H., Meader, N., Stockton, S., Evans, J., O'Connor, R. C., … Kendall, T. (2016). Predicting suicide following self-harm: Systematic review of risk factors and risk scales. *British Journal of Psychiatry*, *209*(4), 277–283. doi:10.1192/bjp.bp.115.170050

Chang, B. P., & Tan, T. M. (2015). Suicide screening tools and their association with near-term adverse events in the ED. *The American Journal of Emergency Medicine*, *33*(11), 1680–1683. doi:10.1016/j.ajem.2015.08.013

Cipriano, A., Cella, S., & Cotrufo, P. (2017). Nonsuicidal self-injury: A systematic review. *Frontiers in Psychology*, *8*, 1946. doi:10.3389/fpsyg.2017.01946

Cochrane Community. (2020). *Covidence*. Retrieved from: https://community.cochrane.org/help/tools-and-software/covidence.

de Beurs, D. P., de Groot, M. H., de Keijser, J., Mokkenstorm, J., van Duijn, E., de Winter, R. F., & Kerkhof, A. J. (2015). The effect of an e-learning supported Train-the-Trainer programme on implementation of suicide guidelines in mental health care. *Journal of Affective Disorders*, *175*, 446–453. doi:10.1016/j.jad.2015.01.046

Department of Health and Human Services. (2016). *Tasmanian suicide prevention strategy (2016-2020) working together to prevent suicide*. Tasmania, Australia. Retrieved from: https://www.dhhs.tas.gov.au/__data/assets/pdf_file/0014/214412/151152_DHHS_Suicide_Prevention_Strategy_Final_WCAG.pdf.

Dickens, G. L., & O'Shea, L. E. (2015). How short should short-term risk assessment be? Determining the optimum interval for START reassessment in a secure mental health service. *Journal of Psychiatric and Mental Health Nursing*, *22*(6), 397–406. doi:10.1111/jpm.12232

Fagan, J., Ijaz, A., Papaconstantinou, A., Lynch, A., O'Neill, H., & Kennedy, H. G. (2009). The suicide risk assessment and management manual (S-RAMM) validation study II. *Irish Journal of Psychological Medicine*, *26*(3), 107–113. doi:10.1017/S0790966700000380

Glassmire, D. M., Tarescavage, A. M., Burchett, D., Martinez, J., & Gomez, A. (2016). Clinical utility of the MMPI-2-RF SUI items and scale in a forensic inpatient setting: Association with interview self-report and future suicidal behaviors. *Psychological Assessment*, *28*(11), 1502. doi:10.1037/pas0000220

Godin, K., Stapleton, J., Kirkpatrick, S. I., Hanning, R. M., & Leatherdale, S. T. (2015). Applying systematic review search methods to the grey literature: A case study examining guidelines for school-based breakfast programs in Canada. *Systematic Reviews*, *4*(1), 138. doi:10.1186/s13643-015-0125-0

Green, K. L., Brown, G. K., Jager-Hyman, S., Cha, J., Steer, R. A., & Beck, A. T. (2015). The predictive validity of the beck depression inventory suicide item. *The Journal of Clinical Psychiatry*, *76*(12), 1683–1686. doi:10.4088/JCP.14m09391

Hanratty, D., Kilicaslan, J., Wilding, H., & Castle, D. (2019). A systematic review of efficacy of Collaborative Assessment and Management of Suicidality (CAMS) in managing suicide risk and deliberate self-harm in adult populations. *Australasian Psychiatry*, *27*(6), 559–564. doi:10.1177/1039856219848832

Harrison, D. P., Stritzke, W. G., Fay, N., & Hudaib, A. R. (2018). Suicide risk assessment: Trust an implicit probe or listen to the patient? *Psychological Assessment*, *30*(10), 1317–1329. doi:10.1037/pas0000577

Hawes, M., Yaseen, Z., Briggs, J., & Galynker, I. (2017). The Modular Assessment of Risk for Imminent Suicide (MARIS): A proof of concept for a multi-informant tool for evaluation of short-term suicide risk. *Comprehensive Psychiatry*, *72*, 88–96. doi:10.1016/j.comppsych.2016.10.002

Health Service Executive. (2017). *HSE Integrated Risk Management Policy—Incorporating an overview of the risk management process*. Dublin, Ireland. Retrieved from: https://www.hse.ie/eng/about/qavd/riskmanagement/risk-management-documentation/hse%20integrated%20risk%20management%20policy%202017.pdf.

Healthcare Improvement Scotland. (2019, November). *SIGN 50: A guideline developer's handbook*. Retrieved from: https://www.sign.ac.uk/assets/sign50_2019.pdf.

Hegarty, J., Flaherty, S. J., Saab, M. M., Goodwin, J., Walshe, N., Wills, T., … Naughton, C. (2020). An international perspective on definitions and terminology used to describe serious reportable patient safety incidents: A systematic review. *Journal of Patient Safety*. Advance online publication. doi:10.1097/PTS.0000000000000700

Higgins, A., Morrissey, J., Doyle, L., Bailey, J., & Gill, A. (2015). *Best practice principles for risk assessment and safety planning for nurses working in mental health services*. Retrieved from: https://www.lenus.ie/bitstream/handle/10147/559255/Best+Principles+Risk+Assessment.pdf?sequence=1.

Jeffs, L., Berta, W., Lingard, L., & Baker, G. R. (2012). Learning from near misses: From quick fixes to closing off the Swiss-cheese holes. *BMJ Quality & Safety*, *21*(4), 287–294. doi:10.1136/bmjqs-2011-000256

Jobes, D. A. (2012). The Collaborative Assessment and Management of Suicidality (CAMS): An evolving evidence-based clinical approach to suicidal risk. *Suicide and Life-Threatening Behavior*, *42*(6), 640–653. doi:10.1111/j.1943-278X.2012.00119.x

Jordan, J. T., & McNiel, D. E. (2018). Characteristics of a suicide attempt predict who makes another attempt after hospital discharge: A decision-tree investigation. *Psychiatry Research*, *268*, 317–322. doi:10.1016/j.psychres.2018.07.040

Kapur, N., & Goldney, R. D. (2019). *Suicide prevention*. Oxford: Oxford University Press.

Katz, C., Randall, J. R., Sareen, J., Chateau, D., Walld, R., Leslie, W. D., … Bolton, J. M. (2017). Predicting suicide with the SAD PERSONS scale. *Depression and Anxiety*, *34*(9), 809–816. doi:10.1002/da.22632

King, C. A., Horwitz, A., Czyz, E., & Lindsay, R. (2017). Suicide risk screening in healthcare settings: Identifying males and females at risk. *Journal of Clinical Psychology in Medical Settings*, *24*(1), 8–20. doi:10.1007/s10880-017-9486-y

Klonsky, E. D., May, A. M., & Saffer, B. Y. (2016). Suicide, suicide attempts, and suicidal ideation. *Annual Review of Clinical Psychology*, *12*(1), 307–330. doi:10.1146/annurev-clinpsy-021815-093204

Luoma, J. B., Martin, C. E., & Pearson, J. L. (2002). Contact with mental health and primary care providers before suicide: A review of the evidence. *American Journal of Psychiatry*, *159*(6), 909–916. doi:10.1176/appi.ajp.159.6.909

Madan, A., Frueh, B. C., Allen, J. G., Ellis, T. E., Rufino, K. A., Oldham, J. M., & Fowler, J. C. (2016). Psychometric reevaluation of the Columbia-Suicide Severity Rating Scale: Findings from a prospective, inpatient cohort of severely mentally ill adults. *The Journal of Clinical Psychiatry*, *77*(07), e867-73–e873. doi:10.4088/JCP.15m10069

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma, G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. doi:10.1371/journal.pmed.1000097

Moola, S., Munn, Z., Sears, K., Sfetcu, R., Currie, M., Lisy, K., … Mu, P. (2015). Conducting systematic reviews of association (etiology): The Joanna Briggs Institute's approach. *International Journal of Evidence-Based Healthcare*, *13*(3), 163–169. doi:10.1097/XEB.0000000000000064

O'Rourke, M. C., Jamil, R. T., & Siddiqui, W. (2021). *Suicide screening and prevention*. Treasure Island, FL: StatPearls Publishing. Retrieved from: https://www.ncbi.nlm.nih.gov/books/NBK531453/.

O'Shea, L. E., & Dickens, G. L. (2014). Short-Term Assessment of Risk and Treatability (START): Systematic review and meta-analysis. *Psychological Assessment*, *26*(3), 990–1002. doi:10.1037/a0036794

O'Shea, L. E., Picchioni, M. M., Mason, F. L., Sugarman, P. A., & Dickens, G. L. (2014). Predictive validity of the HCR-20 for inpatient self-harm. *Comprehensive Psychiatry*, *55*(8), 1937–1949. doi:10.1016/j.comppsych.2014.07.010

Oyesanya, M., Lopez-Morinigo, J., & Dutta, R. (2015). Systematic review of suicide in economic recession. *World Journal of Psychiatry*, *5*(2), 243. doi:10.5498/wjp.v5.i2.243

Pritchard, C., & Hansen, L. (2015). Examining undetermined and accidental deaths as source of 'under-reported-suicide' by age and sex in twenty Western countries. *Community Mental Health Journal*, *51*(3), 365–376. doi:10.1007/s10597-014-9810-z

Queensland Mental Health Commission. (2015). *Queensland suicide prevention action plan 2015–2017*. Queensland Mental Health Commission. Retrieved from: https://www.qmhc.qld.gov.au/sites/default/files/wp-content/uploads/2015/09/Queensland-Suicide-Prevention-Action-Plan-2015-17_WEB.pdf.

Quinlivan, L., Cooper, J., Davies, L., Hawton, K., Gunnell, D., & Kapur, N. (2016). Which are the most useful scales for predicting repeat self-harm? A systematic review evaluating risk scales using measures of diagnostic accuracy. *BMJ Open*, *6*(2), e009297. doi:10.1136/bmjopen-2015-009297

Quinlivan, L., Cooper, J., Meehan, D., Longson, D., Potokar, J., Hulme, T., … Kapur, N. (2017). Predictive accuracy of risk scales following self-harm: Multicentre, prospective cohort study. *British Journal of Psychiatry*, *210*(6), 429–436. doi:10.1192/bjp.bp.116.189993

Randall, J. R., Sareen, J., Chateau, D., & Bolton, J. M. (2019). Predicting future suicide: Clinician opinion versus a standardized assessment tool. *Suicide and Life-Threatening Behavior*, *49*(4), 941–951. doi:10.1111/sltb.12481

Reilly-Harrington, N. A., Shelton, R. C., Kamali, M., Rabideau, D. J., Shesler, L. W., Trivedi, M. H., … Nierenberg, A. A. (2016). A tool to predict suicidal ideation and behavior in bipolar disorder: The Concise Health Risk Tracking Self-Report. *Journal of Affective Disorders*, *192*, 212–218. doi:10.1016/j.jad.2015.12.036

Ridani, R., Torok, M., Shand, F., Holland, C., Murray, S., Borrowdale, K., … Christensen, H. (2016). *An evidence-based systems approach to suicide prevention: Guidance on planning, commissioning, and monitoring*. Sydney, Australia: Black Dog Institute. Retrieved from: https://www.blackdoginstitute.org.au/wp-content/uploads/2020/04/an-evidence-based-systems-approach-to-suicide-prevention.pdf?sfvrsn=0.

Runeson, B., Odeberg, J., Pettersson, A., Edbom, T., Adamsson, I. J., & Waern, M. (2017). Instruments for the assessment of suicide risk: A systematic review evaluating the certainty of the evidence. *PLoS One*, *12*(7), e0180292. doi:10.1371/journal.pone.0180292

SANE Australia. (2014). *Suicide prevention and recovery guide: A resource for mental health professionals*. Australia. Retrieved from: https://www.sane.org/images/PDFs/2779_SANE_SPRG_2016_06.pdf.

Saunders, K., Brand, F., Lascelles, K., & Hawton, K. (2014). The sad truth about the SADPERSONS Scale: An evaluation of its clinical utility in self-harm patients. *Emergency Medicine Journal*, *31*(10), 796–798. doi:10.1136/emermed-2013-202781

Stefansson, J., Nordström, P., Runeson, B., Åsberg, M., & Jokinen, J. (2015). Combining the suicide intent scale and the Karolinska interpersonal violence scale in suicide risk assessments. *BMC Psychiatry*, *15*(1), 226. doi:10.1186/s12888-015-0607-6

Suicide Prevention Resource Center. (2014, September). *Suicide screening and assessment*. Retrieved from: http://www.sprc.org/sites/default/files/migrate/library/RS_suicide%20screening_91814%20final.pdf.

The Joanna Briggs Institute. (2017). *The Joanna Briggs Institute Critical Appraisal tools for use in JBI Systematic Reviews Checklist for Cohort Studies*. Retrieved from: https://joannabriggs.org/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Cohort_Studies2017_0.pdf.

Thom, R., Hogan, C., & Hazen, E. (2020). Suicide risk screening in the hospital setting: A review of brief validated tools. *Psychosomatics*, *61*(1), 1–7. doi:10.1016/j.psym.2019.08.009

United Nations Development Programme. (2019). *Human development report 2019*. Retrieved from: http://hdr.undp.org/sites/default/files/hdr2019.pdf.

US Preventive Services Task Force. (2014). *Suicide risk in adolescents, adults and older adults: Screening*. Retrieved from: https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/suicide-risk-in-adolescents-adults-and-older-adults-screening.

Vasiliadis, H. M., Ngamini-Ngui, A., & Lesage, A. (2015). Factors associated with suicide in the month following contact with different types of health services in Quebec. *Psychiatric Services*, *66*(2), 121–126. doi:10.1176/appi.ps.201400133

Villegas, A. C., DuBois, C. M., Celano, C. M., Beale, E. E., Mastromauro, C. A., Stewart, J. G., … Hoeppner, B. B. (2018). A longitudinal investigation of the Concise Health Risk Tracking Self-Report (CHRT-SR) in suicidal patients during and after hospitalization. *Psychiatry Research*, *262*, 558–565. doi:10.1016/j.psychres.2017.09.044

Wang, Y., Bhaskaran, J., Sareen, J., Bolton, S. L., Chateau, D., & Bolton, J. M. (2016). Clinician prediction of future suicide attempts: A longitudinal study. *The Canadian Journal of Psychiatry*, *61*(7), 428–432. doi:10.1177/0706743716645287

Warden, S., Spiwak, R., Sareen, J., & Bolton, J. M. (2014). The SAD PERSONS scale for suicide risk assessment: A systematic review. *Archives of Suicide Research*, *18*(4), 313–326. doi:10.1080/13811118.2013.824829

Welsh Government. (2015). *Talk to me 2: Suicide and self harm prevention strategy and action plan for wales 2015–2020*. Cardiff, Wales. Retrieved from: https://gov.wales/sites/default/files/publications/2019-06/talk-to-me-2-suicide-and-self-harm-prevention-action-plan-for-wales-2015-2020.pdf.

Woodford, R., Spittal, M. J., Milner, A., McGill, K., Kapur, N., Pirkis, J., … Carter, G. (2019). Accuracy of clinician predictions of future self-harm: A systematic review and meta-analysis of predictive studies. *Suicide and Life-Threatening Behavior*, *49*(1), 23–40. doi:10.1111/sltb.12395

World Health Organization. (2019, September 27). *Suicide data. Mental health*. Retrieved from: https://www.who.int/mental_health/prevention/suicide/suicideprevent/en/.

Wu, C. Y., Huang, H. C., Wu, S. I., Sun, F. J., Huang, C. R., & Liu, S. I. (2014). Validation of the Chinese SAD PERSONS Scale to predict repeated self-harm in emergency attendees in Taiwan. *BMC Psychiatry*, *14*(1), 44. doi:10.1186/1471-244X-14-44