HHS Public Access

Author manuscript

J Clin Psychol. Author manuscript; available in PMC 2021 December 01.

Published in final edited form as:

J Clin Psychol. 2020 December ; 76(12): 2264–2282. doi:10.1002/jclp.22994.

Routinized categorization of suicide risk into actionable strata: Establishing the validity of an existing suicide risk assessment framework in an outpatient sample

Austin J. Gallyer[1], **Carol Chu**[1,2], **Kelly M. Klein**[1,3], **Jazmine Quintana**[1], **Corinne Carlton**[1], **Sean P. Dougherty**[1], **Thomas E. Joiner**[1]

[1]Department of Psychology, Florida State University, Tallahassee, Florida, USA

[2]Department of Psychology, Harvard University, Cambridge, Massachusetts, USA

[3]VA Boston Healthcare System, Brockton Division, Brockton, Massachusetts, USA

Abstract

Objective—The Suicide Risk Assessment and Management Decision Tree (DT) is a clinician-administered assessment that leads to risk categorizations that correspond with actionable strata. This study investigated the construct validity and test–retest reliability of the DT risk categories across two time points.

Method—Outpatients (N = 731) completed a battery of self-report measures. Spearman's correlations were used to examine the relationships between DT suicide risk level and suicidal symptoms, theory-based risk factors, psychiatric correlates, and DT suicide risk level at Timepoint 2. Correlations were analyzed for significant differences to examine the divergent validity of the DT.

Results—Results, overall, were in line with hypotheses, with the exception of depression and thwarted belongingness.

Conclusions—Findings provide evidence for the reliability, convergent validity, and discriminant validity of the DT. This clinician-administered suicide risk assessment may be useful for standardization of the assessment and management of suicide risk in outpatient clinical settings.

Keywords

outpatient; risk; suicidal ideation; suicide attempt; suicide risk assessment; validity

Correspondence Austin J. Gallyer, Department of Psychology, Florida State University, 1107 West Call St., Tallahassee, FL 32306, USA, Gallyer@psy.fsu.edu.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

1 | INTRODUCTION

Suicide is a leading cause of death, with over 47,000 individuals dying by suicide in the United States each year (Centers for Disease Control, 2017). Predicting who will attempt suicide has been a concern of scientists across mental health fields. Indeed, determining who is at risk for suicide is a critical suicide prevention strategy (U.S. Department of Health and Human Services, 2012) because the assessment and management of suicide risk is an avenue for connecting at-risk individuals with evidence-based mental healthcare services (Hom, Stanley, & Joiner, 2015). A crucial component of suicide risk assessment is risk stratification (e.g., low vs. high, long-term vs. imminent; Boudreaux & Horowitz, 2014). Such categorizations not only reflect the nature of suicide risk (Rufino et al., 2018; Witte, Holm-Denoma, Zuromski, Gauthier, & Ruscio, 2017) but can also be used to determine actionable steps to mitigate risk.

A common approach to the assessment of suicide risk is the use of self-report scales, such as the Beck Scale for Suicide Ideation (BSS; Beck, Kovacs, & Weissman, 1979). Though there is utility in screening for suicide risk using self-report scales (Batterham et al., 2015), there are notable limitations associated with self-report measures, such as respondents' idiosyncratic interpretations of items, as well as clinicians' limited abilities to both clarify respondents' inconsistencies and appraise the validity of criterion endorsement (Barker, Pistrang, & Elliott, 2005). Further, research has suggested that a clinician-administered suicide risk assessment interview may be a more accurate reflection of standardized nomenclature than a self-report item of suicidality (Hom, Joiner, & Bernert, 2016). Differences in nomenclature (e.g., what constitutes a "suicide attempt" vs. "nonsuicidal self-injury;" Crosby, Ortega, & Melanson, 2011) are important to clarify, given recent work showing differences in risk factor severity among those with aborted, interrupted, or actual suicide attempts (Rogers, Hom, Dougherty, Gallyer, & Joiner, 2018). Thus, a suicide risk assessment conducted by a trained clinician (cf. McNiel et al., 2008) who follows a standardized risk assessment protocol may yield more accurate findings than suicide risk assessments based on self-report scales alone.[1]

Several clinician-administered suicide risk assessment frameworks exist—such as the University of Washington Risk Assessment Protocol (UWRAP; Linehan, Comtois, & Ward-Ciesielski, 2012), the Chronological Assessment of Suicidal Events (CASE; Shea, 2002), and the Columbia Suicide Severity Rating Scale (C-SSRS; Posner et al., 2011). However, these existing suicide risk assessment frameworks have important clinical limitations. Specifically, while the UWRAP has valuable components (e.g., provision of emergency numbers, assessment of suicide intent), it is a broad assessment of mood and general distress, thus requiring more time to administer (Linehan et al., 2012). Moreover, the UWRAP protocol lacks clear, titrated actions on the basis of suicide risk, making administration difficult for individuals without sufficient clinical experience with patients experiencing suicidal symptoms. Other assessments, including the C-SSRS, lack clear guidelines for managing suicidal symptoms based on the determined level of suicide risk.

[1]We acknowledge that, in many settings, there may not be adequate infrastructure in place to conduct thorough clinician-administered suicide risk assessments (Horowitz, Ballard, & Pao, 2009); we nonetheless emphasize the importance of prioritizing patient safety.

The Suicide Risk Assessment and Management Decision Tree (hereafter referred to as the Decision Tree; Chu et al., 2015; Joiner, Walker, Rudd, & Jobes, 1999) addresses these limitations. The Decision Tree (DT) was developed to be a comprehensive, brief, and efficient assessment of suicide risk with clear, risk-dependent instructions for managing patients experiencing suicidal symptoms.

The DT is a semistructured clinical interview during which chronic (e.g., suicide attempt history) risk factors, acute (e.g., agitation) risk factors, and access to means (e.g., firearms), are considered, and risk is fluidly categorized into *low*, *moderate*, *severe*, or *extreme* risk levels, based on provided guidelines.[2] Moreover, there is decisional latitude when categorizing an individual at the bridge of two groups (e.g., *low-to-moderate* risk), and risk categorization can shift as an individual's clinical presentation changes. Unique to the DT is that its risk categories correspond to standardized, actionable suicide risk mitigation strategies. These strategies include encouraging social support for *low* (and higher) risk, safety planning and phone check-ins for *moderate* (and higher) risk, and consideration of involuntary hospitalization for *severe* or *extreme* risk. For a full discussion of the risk categories and corresponding clinical actions, see Chu et al. (2015).

Though the DT assesses risk factors that have been empirically supported, there is no research, to our knowledge, that has formally tested the validity and reliability of the DT. One study, conducted by Van Orden, Witte, Gordon, Bender, and Joiner (2008), reported that DT suicide risk categories were strongly correlated with Beck Scale for Suicidal Ideation scores ($r = .64$, $p < .001$; Beck & Steer, 1991). Thus, though there is some evidence for the construct validity of the DT, a comprehensive examination of the construct validity of the DT is needed. To assess inter-rater reliability, Van Orden and colleagues also used independent raters blind to the original DT suicide risk designation rate suicide risk level based on the information recorded from the DT. Using this approach, they found evidence of interrater reliability for the DT ($\kappa = .71$, $p < .001$). However, Van Orden et al.'s (2008) study examined a relatively small sample of outpatients ($n = 153$) and did not fully evaluate the construct validity, nor did they examine the test–retest reliability of the DT. Therefore, the present study sought to examine the construct validity and test–retest reliability of the DT in a large sample of individuals seeking outpatient treatment. To characterize the convergent and divergent validity of the DT, we evaluated the relationship between the DT and the following self-reported variables: (1) suicidal symptoms; (2) theory-based correlates; and (3) psychiatric correlates.

Given that the DT designation of suicide risk is largely determined by severity of suicidal ideation and behavior (e.g., plans, preparations, attempt history), we hypothesized that DT suicide risk category would have a moderate-to-strong (i.e., $\rho \geq 0.30$) positive association with self-report measures of suicidal ideation, attempt history, and attempt recency. Regarding theory-based correlates, we examined constructs common to the ideation-to-

[2] *Low risk* does not equate to zero risk. In describing an individual who denies any past or current suicide ideation, desire, intent, plans, preparations, and attempts, as well as nonsuicidal self-injury, we emphasize the importance of maintaining the *low*-risk designation. Indeed, the modal individual receiving a suicide risk assessment is likely to be receiving psychiatric or medical care in a health setting; in light of these circumstances, the individual may be likely to have some degree of risk. We prefer the term *low* risk, but we also see comparable utility in the terms *above-zero* or *near-zero* risk for some individuals.

action theories of suicide (Klonsky, Saffer, & Bryan, 2018). One of these models, called the interpersonal theory of suicide (IPTS; Joiner, 2005; Van Orden et al., 2010), posits that suicidal desire results when an individual feels an intractable sense of *perceived burdensomeness* and *thwarted belongingness*. The theory further proposes that, for an individual with suicidal desire to engage in suicidal behaviors, the individual must also possess the *capability for suicide*, which is composed of increased pain tolerance, fearlessness about death, and practical knowledge about the means of death. In line with the theory's predictions, a recent meta-analysis of the IPTS found that perceived burdensomeness was strongly associated with suicide risk ($r = .42$), thwarted belongingness was moderately associated with suicide risk ($r = .33$), and capability for suicide was not associated with suicide risk ($r = .09$; Chu et al., 2017). Regarding capability for suicide, it should be noted that it is not viewed as a risk factor for suicide in isolation. Rather, capability is viewed as being an important moderator between desire for suicide and suicide attempts, with those high in capability and high in desire engaging in suicidal behavior. Based on the predictions of the IPTS and the meta-analysis by Chu et al. (2017), we hypothesized that perceived burdensomeness would have a moderate to strong (i.e., $\rho \geq 0.30$) association with DT suicide risk. Moreover, we hypothesized that thwarted belongingness would have a moderate (i.e., $\rho = 0.30–0.40$) relationship with DT suicide risk, and that capability for suicide would have no association (i.e., $\rho = 0.00–0.10$) with DT suicide risk category.

We also examined the correlation between the DT and psychiatric constructs that are associated with suicide risk, such as insomnia, agitation, anxiety, depression, and borderline personality disorder symptoms (Bentley et al., 2016; Bernert, Kim, Iwata, & Perlis, 2015; Black, Blum, Pfohl, & Hale, 2004; Hall, Platt, & Hall, 1999; Rogers, Ringer, & Joiner, 2016). Given that these more distal risk factors may play a lesser role in suicide risk designations (Chu et al., 2015), we hypothesized that DT suicide risk category would have a small-to-moderate (i.e., $\rho = 0.10–0.30$) positive association with these variables. To provide evidence for discriminant validity, correlations between DT suicide risk and measures expected to be closely related to DT suicide risk should be significantly stronger than correlations between DT suicide risk and measures expected to be less closely related to DT suicide risk (Campbell & Fiske, 1959; Clark & Watson, 2019). Therefore, we hypothesized that the correlations between DT suicide risk and suicidal symptoms would be significantly greater than correlations between DT suicide risk and theoretical risk factors. We also hypothesized that the correlations between DT suicide risk and both thwarted belongingness and perceived burdensomeness would be significantly greater than the correlations between DT suicide risk and psychiatric correlates. We considered these two hypotheses at least partially supported if at least half of the correlation comparisons between groups were significant in the expected direction. Regarding capability for suicide, we hypothesized that the correlation between DT suicide risk and psychiatric correlates would be significantly greater than the correlation between DT suicide risk and capability for suicide.

Lastly, though suicidal symptoms have been shown to highly fluctuate over relatively short timeframes (i.e., hours or days; Kleiman et al., 2017), we hypothesized that the DT suicide risk category would demonstrate strong test–retest reliability across two timepoints (i.e., $\rho > 0.7$). Though suicidal symptoms may fluctuate, one of the most important risk factors

considered within the DT framework—suicide attempt history—is fairly stable, with the majority of individuals reporting only one suicide attempt (Liu, Zhang, & Sun, 2017). Moreover, recent evidence suggests that suicide risk is categorical, rather than dimensional (Rufino et al., 2018; Witte et al., 2017). Thus, while suicide symptoms fluctuate, suicide risk is expected to remain fairly stable.

Our analyses will inform whether the DT demonstrates acceptable construct validity. Though other forms of validity are also important (e.g., criterion validity), the present analyses reflect an important step toward examining whether the DT is effective in clinical practice, and measures what it purports to measure.

2 | METHOD

2.1 | Participants

The present sample is composed of outpatients (N = 731, 61.7% female) who presented to a training clinic affiliated with an American Psychological Association (APA)-accredited clinical science program within a large Southeastern university. Patients were not limited to college students; the modal patient presenting for services was not a student. The sample consists of adults (M = 27.39, SD = 10.76, range = 18–71) who completed the clinic application process between April 2005 and September 2015. The race/ethnicity of the sample is: 10.9% Hispanic, 73.3% White (non-Hispanic), 10.7% Black, 1.6% Asian/Pacific Islander, 0.4% Native-American/Alaska Native, and 0.8% did not report their race/ethnicity. Marital statuses include single/never married (78.5%), married (10.7%), separated (1.4%), widowed (0.7%), and divorced (8.8%). Regarding suicide attempt history, 76.5% of participants had no reported history of suicide attempts, 14.0% had one reported past attempt, and 8.6% had two or more reported past attempts; data on suicide attempt history were missing for 1.0% of participants.

All DT risk assessments and subsequent categorizations were performed by therapists working in the training clinic. Therapists were doctoral students in a clinical science program who were supervised by clinical psychologists and who had, at minimum, completed their first year of graduate training. Before working in the clinic, the therapists attended 2–3-h didactic seminars on the DT provided by the senior author (T. E. J.). These training sessions included review of the DT protocols and roleplaying; therapists were required to demonstrate proficiency in their administration of the DT assessment before working in the clinic. At this clinic, suicide risk assessments are routinely administered to all patients.

2.2 | Procedure

2.2.1 | Baseline (T1; screening)—Data for this study were collected as part of an ongoing endeavor to promote research on individuals receiving outpatient psychiatric services, and the University's Institutional Review Board approved all data collection procedures. As part of the clinic application process, patients seeking outpatient services completed a battery of self-report questionnaires that were administered online before a screening appointment. The average time between the completion of self-report

questionnaires and the screening appointment was 4.30 days (*SD* = 10.10 days, mode = 0 days). During the screening appointment, patients provided informed consent before entry into research at the clinic. All study patients participated in a 1- to 1.5-h screening appointment during which the screening therapist assessed suicide risk using the DT suicide risk assessment framework (Chu et al., 2015; Joiner et al., 1999). The screening therapists constituted a subset of all training therapists in this clinic; screening therapists were more advanced in their clinical training (i.e., second year of clinical training or above; third year in the clinical Ph.D. program or above).

2.2.2 | Follow-Up (T2; intake)—Following the screening appointment, the patient was assigned to a training therapist who managed the case under the supervision of a licensed clinical psychologist. This therapist conducted the intake session during which the therapist gathered more details regarding the patient's symptoms and assessed suicide risk using the DT. The average time between the screening appointment and the first intake session was 36.30 days (*SD* = 35.20 days, mode = 13 days). The wait time varied based on the patients' presenting problems and availability, the length of the waitlist, and the number of therapists requiring a new case.

2.3 | Measures

2.3.1 | Decision Tree Suicide Risk Levels (DT; Chu et al., 2015; Joiner et al., 1999)—The DT, a semi-structured clinical interview, was used to determine suicide risk level. Patients were categorized into one of seven risk categories: *low*, *low-to-moderate*, *moderate*, *moderate-to-severe*, *severe*, *severe-to-extreme*, and *extreme* by the assessor who conducted the interview at each timepoint. Patients' suicide risk levels were based chiefly on the severity of suicidal desire, ideation, intent, and resolved suicide plans and preparations. A patient's self-reported suicide attempt history and recency, as well as access to means, were all considered. Consistent with evidence suggesting that follow-up questions are needed to improve precision when assessing suicidal behavior history (Hom et al., 2016; Millner, Lee, & Nock, 2015), all therapists verified self-reported attempts with open-ended questions that aimed to assess the specific behavior, lethality, and intent. These main variables were examined in conjunction with the patient's overall clinical presentation and the presence of other significant risk factors (e.g., insomnia, agitation). Given constraints of the present study setting, at each timepoint, risk level was assessed by one therapist (i.e., T1 = Screening therapist; T2 = Treatment therapist). Thus, interrater reliability was not available for the current study. However, a previous study found support for the inter-rater reliability between therapists assessing suicide risk in this clinic ($\kappa = 0.71$, $p < .001$; Van Orden et al., 2008). Administration time of the DT ranges from 5 to 20 min, depending on risk level.

2.3.2 | BSS (Beck & Steer, 1991)—The BSS is a 21-item measure of suicidality, including current suicidal ideation, plans, preparations, and suicide attempt history. The first 19 items of this scale were summed and used as a measure of suicidal ideation, including thoughts about plans/preparations. Item 20, which was coded 0 = *no history of suicide attempts*, 1 = *history of one suicide attempt*, 2 = *history of more than one suicide attempt*, was used as a measure of suicide attempt history in the present study. In addition to these

measures, we created a summed measure of BSS items 1–7 and item 9 based on a previous study as a measure of suicidal desire (Dhingra, Klonsky, & Tapola, 2018). In an outpatient sample of individuals with mood disorders, the BSS was shown to have good reliability and construct validity (Beck, Epstein, Brown, & Steer, 1988; Beck, Steer, & Ranieri, 1988). In the present study, the internal consistency of items 1 through 19 was excellent ($\omega = 0.91$).[3]

2.3.3 | Suicide attempt timing—The timing of the most recent suicide attempt was assessed with a single self-report item developed by one of the authors: "If you have ever attempted suicide, when did your most recent suicide attempt occur?" This item was coded: 1 = *within the last month*, 2 = *more than 1 month ago but within the past year*, 3 = *More than 1 year ago but less than 5 years ago*, 4 = *more than 5 years ago*, and 5 = *I have never attempted suicide*.

2.3.4 | Depressive Symptom Index—Suicidality Subscale (DSI-SS; Joiner, Pfaff, & Acres, 2002)—The DSI-SS is a four-item subscale of the Hopelessness Depression Symptom Questionnaire (Metalsky & Joiner, 1997) that measures suicidal ideation and impulses over the past 2 weeks. Consistent with previous findings demonstrating good psychometric properties of the subscale (Joiner et al., 2002; Metalsky & Joiner, 1997), the DSI-SS exhibited excellent internal consistency in the present sample ($\omega = 0.92$).

2.3.5 | Interpersonal Needs Questionnaire (INQ-15; Van Orden, Cukrowicz, Witte, & Joiner, 2012)—The INQ-15 is a 15-item scale used to index perceived burdensomeness (INQ-PB) and thwarted belongingness (INQ-TB), which are two constructs derived from the IPTS. Items are rated on a 7-point Likert scale regarding how the individual has been feeling "recently," from 1 (*Not at all true for me*) to 7 (*Very true for me*). The perceived burdensomeness subscale consists of six items, and the thwarted belongingness subscale consists of nine items. In the current study, the reliabilities of the INQ-TB ($\omega = 0.97$) and the INQ-PB ($\omega = 0.97$) subscales were excellent. Data for the INQ were only available from 2012 to 2015.

2.3.6 | Acquired Capability for Suicide Scale—Fearlessness About Death (ACSS-FAD; Ribeiro et al., 2014)—The ACSS-FAD is a seven-item scale designed to measure fearlessness about death, a subconstruct of capability for suicide derived from the IPTS. Scores range from 0 to 28, with higher scores indicating greater fearlessness about death. In the current study, the ACSS-FAD demonstrated good reliability ($\omega = 0.87$). Data for the ACSS-FAD were only available from 2012 to 2015.

2.3.7 | Brief Agitation Measure (BAM; Ribeiro, Bender, Selby, Hames, & Joiner, 2011)—The BAM is a three-item measure of subjective agitation that has shown good reliability and validity in a clinical sample obtained from the same population as the

[3]We use ω (Zinbarg, Revelle, Yovel, & Li, 2005) as our measure of internal consistency throughout this report, rather than Cronbach's α. This choice is due to a large body of methodological evidence indicating that α relies on assumptions that are frequently violated and that result in inflated internal consistency measurements (see Dunn, Baguley, & Brunsden, 2014 for a review and discussion).

present study (Ribeiro et al., 2011). Reliability of this measure was good in this study (ω = 0.88). This measure was only available from 2012 to 2015.

2.3.8 | Insomnia Severity Index (ISI; Morin, 1993)—The ISI is a seven-item measure that assesses symptoms of insomnia and related distress (Bastien, Vallières, & Morin, 2001) and has demonstrated acceptable psychometric properties in both community and clinical samples (Bastien et al., 2001; Morin, Belleville, Bélanger, & Ivers, 2011). The internal consistency of the ISI was excellent in this study (ω = 0.92).

2.3.9 | Beck Anxiety Inventory (BAI; Beck et al., 1988)—The BAI is a 21-item questionnaire designed to measure anxiety symptoms, as independent from related symptoms of depression (Beck et al., 1988). The BAI has demonstrated good validity and reliability across numerous studies (Beck et al., 1988; Fydrich, Dowdall, & Chambless, 1992; Steer, Ranieri, Beck, & Clark, 1993) and excellent internal reliability in the present study (ω = 0.93).

2.3.10 | Beck Depression Inventory Second Edition (BDI-II; Beck, Steer, & Brown, 1996)—The BDI-II is a widely used 21-item measure of depressive symptoms. Many studies across diverse samples have provided support for the psychometric properties of the BDI-II (Dozois, Dobson, & Ahnberg, 1998; Osman, Kopper, Barrios, Gutierrez, & Bagge, 2004; Sprinkle et al., 2002; Steer, Ball, Ranieri, & Beck, 1997). In the current analyses, internal consistency of this measure was excellent (ω = 0.94)

2.3.11 | Personality Inventory for the DSM-5 (PID-5; Krueger, Derringer, Markon, Watson, & Skodol, 2012)—The PID-5 is a 220-item self-report measure that assesses 25 maladaptive traits across five broad domains on the basis of a dimensional DSM-5 model of personality disorders. In this study, the following individual PID-5 personality trait scale scores (from 0 to 3) were averaged to represent DSM-5 borderline personality disorder (BPD) symptoms: Emotional Lability, Anxiousness, Separation Insecurity, Hostility, Depressivity, Impulsivity, and Risk-taking. These PID-5 traits were selected based on Hopwood, Thomas, Markon, Wright and Krueger (2012) study, which provided preliminary support for PID-5 traits as a valid measure of personality pathology, including BPD. Previous studies indicated that reliability for all trait scores were adequate or better (i.e., $\alpha > .71$), and domain-level reliabilities were in the excellent range ($0.91 \leq \alpha \leq 0.96$; Krueger et al., 2012; Strickland, Drislane, Lucy, Krueger, & Patrick, 2013). In this study, internal consistency of the PID-5 BPD measure was excellent (ω = 0.91). Data for the PID-5 were available from 2013 to 2015.

2.4 | Data analysis

Manuscript preparation and analyses were conducted in R (Version 3.5.2; R Core Team, 2018) using the following R-packages: *tidyverse* (Version 1.2.1; Wickham, 2017), *papaja* (Version 0.1.0.9842; Aust & Barth, 2018), *haven* (Version 2.1.0; Wickham & Miller, 2019), *mice* (Version 3.3.0; van Buuren & Groothuis-Oudshoorn, 2011), *miceadds* (Version 3.0.16; Robitzsch, Grund, & Henke, 2018), *cocor* (Version 1.1.3; Diedenhofen & Musch, 2015),

scales (Version 1.0.0; Wickham, 2018), *DataExplorer* (Version 0.7.0; Cui, 2018), *lubridate* (Version 1.7.4; Grolemund & Wickham, 2011), and *psych* (Version 1.8.4; Revelle, 2018).

We first examined missing data. Most variables had below 4% missingness, including the BDI, BAI, BSS, suicide attempt history, DSI-SS, BAM, ACSS-FAD, and recency of suicide attempt. Variables with a greater percentage of missingness included: DT risk level at screening (9.71% missingness), INQ-TB and INQ-PB (17.65%), ISI total score (20.93%), DT risk level at intake (43.09%), and PID-5 BPD score (32.4%). Given that this is a considerable amount of missingness, we used multiple imputation for our main analyses. Specifically, we used the multivariate imputation by chained equations package (*mice*; van Buuren & Groothuis-Oudshoorn, 2011) to create 100 imputed data sets. To examine convergent and test–retest reliability, we then conducted Spearman correlations between all study variables in each of the data sets. We used Spearman's ρ correlation because it is appropriate for ordinal and continuous data, is more robust to nonnormal distributions, and has less variability than Pearson's r correlation (de Winter, Gosling, & Potter, 2016). Next, we pooled our Spearman's ρ results according to Rubin's rules (Rubin, 2004) using the *miceadds* package (Robitzsch et al., 2018). Research has indicated that, even at high rates of missingness (e.g., 90% missing), 100 imputations are able to provide unbiased parameter estimates and provide sufficient power to detect significant effects, comparable with full-information maximum likelihood (Graham, Olchowski, & Gilreath, 2007)[4]. In addition to examining test–retest reliability, we also examined changes in risk designation from T1 to T2. Specifically, we counted how many participants had the same risk level at T2 as T1, how many had increased risk at T2 compared with T1, and how many had decreased risk at T2 compared with T1.

To examine the divergent validity of the DT suicide risk assessment, we tested for significant differences between each of the correlations of DT risk level at screening (T1) with other study variables. For this procedure, we used a confidence interval approach described by Zou (2007) using the *cocor* package (Diedenhofen & Musch, 2015). In this approach, two correlations are considered to be significantly different in strength if the 95% confidence interval does not cross zero. We expected that DT suicide risk (T1) would be more strongly correlated with suicidal symptoms (i.e., BSS, DSI-SS, suicide attempt history, and suicide attempt recency), than with theoretical risk factors (i.e., INQ-PB, INQ-TB, and ACSS-FAD). Moreover, we expected that DT suicide risk (T1), would be more strongly correlated with INQ-PB and INQ-TB than with psychiatric correlates (i.e., BDI, BAI, BAM, ISI, and PID-5). Finally, we expected all variables to be more strongly related to DT suicide risk (T1) than ACSS-FAD.

3 | RESULTS

Spearman ρ correlations between all study variables and means and standard deviations are available in Table 1.

[4]We also ran our analyses using pairwise deletion; the interpretation of our results remained the same (see Table S1).

3.1 | Convergent validity and test–retest reliability

3.1.1 | Suicide symptoms—As hypothesized, we found that DT suicide risk was strongly correlated with measures of suicidal symptoms (i.e., BSS suicidal ideation, BSS suicidal desire, DSI-SS suicidal ideation, attempt history, and attempt recency), with $|\rho|$'s ranging from 0.43 to 0.63. Notably, based on how the recency variable was coded, DT suicide risk exhibited a strong negative relationship with suicide attempt recency. Thus, a more recent suicide attempt was associated with higher suicide risk.

3.1.2 | Theoretical risk factors—We found that DT suicide risk's correlation with theoretical risk factors ranged from nonsignificant to strong. Specifically, DT suicide risk was not related to fearlessness about death (ACSS-FAD; $\rho = 0.06$), weakly related with thwarted belongingness (INQ-TB; $\rho = 0.29$), and strongly related to INQ-PB perceived burdensomeness (INQ-PB; $\rho = 0.54$).

3.1.3 | Psychiatric correlates—In general support of our hypothesis, we found that DT suicide risk's correlations with psychiatric correlates were small (i.e., $\rho = 0.20-0.31$), with one exception: DT suicide risk was moderately correlated with depression as measured by the BDI ($\rho = 0.43$).

3.1.4 | Test–retest reliability—As expected, DT suicide risk demonstrated strong ($\rho = 0.76$) test–retest reliability from screening (T1) to intake (T2).

3.1.5 | Changes in suicide risk—From T1 to T2, 81.75% of participants stayed at the same risk level, 14.36% decreased in suicide risk, and 3.89% increased in suicide risk. Of those that were at least *moderate* risk at T1, 48.72% stayed at the same risk level, 44.87% decreased in suicide risk, and 6.41% increased in suicide risk.

3.2 | Discriminant validity

For all correlation pairwise comparisons, please see Table 2. In line with our hypothesis, we found that most suicide symptom variables were more strongly correlated with DT suicide risk (T1) than the theoretical variables were, with three exceptions: Suicide attempt history was not more strongly correlated with DT suicide risk (T1) than with INQ-PB (95% confidence interval [CI] = −0.07–0.06), BSS suicidal ideation was not more strongly related than the INQ-PB was (95% CI = −0.01 to 0.10), and BSS suicidal desire was not more strongly related than the INQ-PB (95% CI = −0.07–0.05). We also found that all suicide symptoms were more strongly correlated with DT suicide risk (T1) than all psychiatric correlates were. Regarding the comparison of correlations with the DT between psychiatric and theoretical correlates, we found that INQ-PB was more strongly correlated with DT suicide risk (T1) than all psychological correlates were. However, contrary to our hypothesis, INQ-TB was less correlated with DT suicide risk (T1) than the BDI was (95% CI = −0.20 to −0.08), and no significant differences were found between INQ-TB and the other psychiatric correlates. Finally, in line with our hypothesis, we found that the ACSS-FAD had a significantly smaller correlation with DT suicide risk (T1) than all of the psychiatric correlates did.

4 | DISCUSSION

Suicide risk assessment and management are critical suicide prevention strategies. However, few empirically informed suicide risk frameworks exist, and even fewer lead to risk categorizations that correspond with clinical action. For this reason, the DT (Chu et al., 2015; Joiner et al., 1999) was designed to routinize the assessment and management of suicide risk. In this study in a large clinical sample, we found support for the construct validity of the DT, as well as evidence that the DT is reliable across two assessments. The use of empirically supported risk assessment is vital in providing the highest quality of care for those at risk for suicidal behavior, and the convergence between clinician-administered and self-report measures is a foundational first step in establishing the validity of the DT.

Consistent with hypotheses, DT suicide risk was strongly correlated with measures of suicidal ideation, suicide attempt history, and timing of most recent suicide attempt. Specifically, patients categorized by the DT as having higher levels of suicide risk reported higher levels of suicidal ideation, more past suicide attempts, and a more recent suicide attempt. The strong correlation between DT suicide risk and self-reported suicidal ideation ($\rho = 0.68$) is similar to correlations observed in past research examining the relationship between interview-based risk assessment (i.e., C-SSRS) and self-report measures of suicidal ideation (Posner et al., 2011). Interestingly, we found that DT suicide risk was more strongly related to BSS suicidal ideation than to a measure of desire for suicide based on a subset of items from the BSS. As shown by previous work (Dhingra et al., 2018), the BSS is not a unidimensional measure, with some items assessing theoretically and empirically distinct constructs, including suicide plans. Because the DT was designed to be a measure of suicide risk rather than just desire for suicide, this further supports the construct validity of the DT. Encouragingly, we found that the DT was significantly more correlated with self-report measures of suicidal symptoms than with most theoretical risk factors, and all psychiatric correlates. Given that the DT was designed to be more strongly determined by current suicidal symptoms and history of suicidal behavior than by theoretical risk factors or psychiatric correlates, these findings lends support to the validity of the DT. Regarding the DT's association with theoretical risk factors, our results were also generally supportive.

Indeed, in line with hypotheses, the DT was strongly correlated with perceived burdensomeness and was not correlated with fearlessness about death ($\rho = 0.54$ and $\rho = 0.06$, respectively). However, the magnitude of the association between the DT and thwarted belongingness was slightly less than hypothesized ($\rho = 0.30–0.40$ vs. $\rho = .29$). Regardless of this latter correlation, most comparisons of correlations between DT suicide risk and theoretical risk factors, as well as between DT suicide risk and psychiatric correlates, were in the expected direction. Specifically, perceived burdensomeness was more strongly related to DT suicide risk than any psychiatric correlate was, and capability for suicide was less strongly related to DT suicide risk than all psychiatric correlates were. However, thwarted belongingness was less strongly related to DT suicide risk than depression was, and there were no significant differences in correlations with DT suicide risk between thwarted belongingness and any other psychiatric correlate.

Though the association between thwarted belongingness and DT suicide risk was slightly lower than expected, the magnitude of the correlation found in our study is supported by meta-analytic evidence. Specifically, a meta-analysis conducted by Chu et al. (2017) found that the 95% CI for the correlation between thwarted belongingness and suicide risk was 0.25–0.41. Thus, though we hypothesized that the relation between thwarted belongingness and DT suicide risk would fall between 0.30 and 0.40, the correlation observed in our study is well within the range of the best estimate of the true association between thwarted belongingness and suicide risk. On a similar note, though we hypothesized that the correlation between perceived burdensomeness and DT suicide risk would be strong in magnitude (i.e., $\rho > .30$), the association between perceived burdensomeness and suicide risk is slightly higher than the estimated upper bound observed by Chu et al. (i.e., $\rho = 0.54$ vs. $r = .49$, respectively; 2017). Though we cannot be certain about the cause of this finding, there are at least three potential explanations.

The first explanation is that, given that our sample consisted of outpatients at a psychology clinic, the individuals in our sample may have had higher suicide risk than the modal participant within the meta-analysis conducted by Chu et al. (2017). This possibility is evidenced by the high percentage of individuals with a previous suicide attempt in our sample (i.e., 22.79%). Evidence suggests that many risk factors demonstrate a stronger relationship with suicidal ideation at higher levels of ideation, with perceived burdensomeness exhibiting a particularly strong relationship with suicidal ideation at higher levels of suicidal ideation (Rogers & Joiner, 2017). Thus, if our sample had included more individuals with less severe suicidal symptoms, the relationship between perceived burdensomeness and DT suicide risk may have been even more comparable with the relationship observed in previous research (Chu et al., 2017). Another potential explanation is that our missing data technique biased the results. However, using pairwise comparisons yielded the same point-estimate correlation in our sample ($\rho = .54$). Importantly, research shows that case deletion usually leads to underestimation of correlation coefficients (Schafer & Graham, 2002).

Though we hypothesized that fearlessness about death would not have a relationship with suicide risk (i.e., $\rho = 0.00-0.10$), it could be argued that because our study used an outpatient sample that we should have observed a positive relationship between fearlessness about death and DT suicide risk. Specifically, the IPTS hypothesizes that in the context of high thwarted belongingness, perceived burdensomeness, and hopelessness about these two states, that fearlessness about death will be related to suicide risk/behavior (Joiner, 2005; Van Orden et al., 2010). In an outpatient sample like the one used in the present study, it is possible that thwarted belongingness, perceived burdensomeness, and hopelessness are high enough to show this relationship. Though we based our hypothesis largely on meta-analytic evidence (Chu et al., 2017), future studies may consider examining whether in very high-risk groups (e.g., after inpatient hospitalization) that the DT demonstrates a positive relationship with fearlessness about death.

Consistent with our hypotheses, we found that most psychiatric correlates demonstrated a small relationship with DT suicide risk, with the exception of depression, which exhibited a moderate relationship with DT suicide risk. The small relationships found between DT

suicide risk and psychiatric correlates are consistent with a meta-analysis that found that all psychiatric conditions demonstrated weak associations with suicide-related outcomes (Franklin et al., 2017). Taken together, the DT was associated with measures of suicidal symptoms, theoretical risk factors, and psychiatric correlates at magnitudes largely in line with our hypotheses. Therefore, the results of this study supports the construct validity of the DT. This is an important first step in the validation of a measure of suicide risk, as criterion validity cannot be assessed unless evidence supports that the instrument measures what it purports to measure.

Our results also revealed that the DT demonstrates strong test-retest reliability across two timepoints. Importantly, though the limits of the setting of the present study precluded examining the inter-rater reliability of the DT, the interrater reliability of the DT has previously been examined in a sample of over 150 outpatients (Van Orden et al., 2008). In that study, they used three independent raters blind to the original suicide risk designation to rate suicide risk level based on the information recorded by the first rater from the DT (e.g., suicide intent level). This interrater reliability evidence is particularly strong given that the second raters were not able to reinterview the patient, and had to make a suicide risk designation based on the information obtained from administering the DT. Thus, this study and our current results support the DT as a reliable measure of suicide risk.

Our test–retest results also show that, while suicidal ideation may fluctuate within short timeframes (Kleiman et al., 2017), suicide risk can remain relatively stable. Regarding this latter point, while DT suicide risk is highly informed by current suicidal ideation (Chu et al., 2015; Joiner et al., 1999), suicidal ideation is not the only factor taken into consideration when designating suicide risk. For example, presence and severity of suicide plans and preparations are also important indicators of risk within the DT framework (Chu et al., 2015; Joiner et al., 1999). Indeed, research has implicated the presence of a suicide plan as an important predictor for transitioning from suicidal ideation to suicide attempt (Nock et al., 2018). Moreover, evidence suggests that suicide risk is a categorical construct (Rufino et al., 2018; Witte et al., 2017). Therefore, while suicidal ideation may fluctuate, suicide risk is comprised of multiple constructs and, thus, may remain relatively stable across time. This is useful within clinical practice, as large changes in suicidal ideation—while concerning and warrants management according to best practices—does not necessarily indicate a severe change in suicide risk. Given this, the DT may help provide a clearer picture of suicide risk than self-report measures of suicidal symptoms alone.

Regarding changes in suicide risk across timepoints, over 80% of our sample remained at the same risk level from screening to the first intake appointment. Of those whose suicide risk did change, most individuals' suicide risk decreased over time. Given that in many settings suicide risk is assessed once with no follow-up, our results indicate that most will remain at similar risk from screening to intake. Even among those whose risk does change, our results suggest that, at least in our outpatient clinic, risk is more likely to decrease. Though this is the case, we want to emphasize that some individuals' risk did increase, with about 1% going from *low* risk to *moderate-to-severe* risk from screening to intake. Thus, while our results suggest relative suicide risk stability over a month span, it is important that suicide risk continue to be monitored. Moreover, the patterns seen in this sample may not generalize

to other samples and settings, such as inpatient populations. Therefore, to balance the critical need to monitor suicide risk over time and clinician and patient burden, we suggest that clinicians conduct the full DT at the first contact with a patient. At follow-ups, clinicians can then assess for changes to frequency and duration of death and suicidal ideation, instances of nonsuicidal self-injury, current intent and desire, any changes to plans or preparations, and any changes to access to means. Using this follow-up approach can take as little as a couple of minutes, and leads to corresponding clinical actions (e.g., means safety) based on the patient's suicide risk.

4.1 | Strengths, limitations, and future work

The current results are supportive of the construct validity and reliability of the DT framework. Our study was strengthened by using a relatively severe outpatient clinical sample. Consequently, patients with a wide variety of diagnoses and suicidal symptoms were included. Moreover, our sample mostly consisted of patients from the surrounding community, rather than college students. These strengths increase the likelihood of the generalizability of our results. Nevertheless, our study was limited by factors that suggest directions for future research. First, the present study did not examine the predictive validity of the DT. Therefore, research investigating the DT's ability to predict suicidal thoughts and behaviors would further clarify the clinical utility of the assessment.[5] Another limitation of the current study is the use of only self-report measures as evidence for the construct validity of the DT. Furthermore, examining how the DT's predictive validity compares to other interview-based assessments of suicide risk, such as the C-SSRS (Posner et al., 2011), would further elucidate the utility of the DT framework. Another area for future work pertains to the DT's corresponding clinical actions, which are determined by suicide risk designation. By making suicide risk designation correspond with evidence-based clinical actions, such as safety planning (Stanley & Brown, 2012), the DT seeks to improve patient care. Thus, the field would benefit from research examining whether the DT's approach to clinical actions is more effective at mitigating suicide risk than that of other, nonstandardized approaches. Last, though the DT assesses a wide variety of risk factors for designating suicide risk, future work may consider adding other relevant risk factors to the DT. For example, psychological pain is a construct that has shown strong relationships with suicidal ideation and attempts (Lambert et al., 2020; Montemarano, Troister, Lambert, & Holden, 2018). Moreover, psychological pain is featured in some classical and contemporary theories of suicide (Klonsky & May, 2015; Shneidman, 1993). Therefore, future work may consider adding psychological pain to the DT and examining whether this increases the DT's performance.

5 | CONCLUSIONS

We have examined the construct validity and test–retest reliability of a suicide risk assessment framework first developed by Joiner et al. (1999) and updated by Chu et al.

[5]Recent evidence suggests that many risk factors and tools are relatively inaccurate at making absolute determinations about whether an individual will or will not attempt suicide (Franklin et al., 2017; Gutierrez et al., n.d.). In this regard, other approaches, such as models developed using machine-learning algorithms, may be more effective (Ribeiro, Huang, Fox, Walsh, & Linthicum, 2019; Walsh, Ribeiro, & Franklin, 2017, 2018). However, there have been problems raised about this approach as well (e.g., Belsher et al., 2019).

(2015). Our results support both the construct validity and the test–retest reliability of the DT suicide risk assessment framework. While further work will help inform the scope of the DT's utility, the DT can be considered for use in clinical practice to routinize the assessment and management of suicide risk among patients with a variety of diagnoses and severity of suicidal symptoms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This study was supported by a grant by the National Institute of Mental Health T32MH093311. This study was also supported in part by the Military Suicide Research Consortium (MSRC), an effort supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award No. W81XWH-16-2-0004. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the MSRC or the Department of Defense.

Funding information

National Institute of Mental Health, Grant/Award Number: T32MH093311; U.S. Department of Defense, Grant/Award Number: W81XWH-16-2-0004

REFERENCES

Aust F, & Barth M (2018). papaja: Create APA manuscripts with R Markdown. Retrieved from https://github.com/crsh/papaja

Barker C, Pistrang N, & Elliott R (2005). Self-report methods. In Research methods in clinical psychology: An introduction for students and practitioners (2nd ed.). Chichester: John Wiley & Sons, Ltd 10.1002/0470013435.ch6

Bastien CH, Vallières A, & Morin CM (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Medicine, 2, 297–307. [PubMed: 11438246]

Batterham P, Ftanou M, Pirkis J, Brewer J, Mackinnon A, Beautrais A, … Christensen H (2015). A systematic review and evaluation of measures for suicidal ideation and behaviors in population-based research. Psychological Assessment, 27(2), 501–512. 10.1037/pas0000053 [PubMed: 25496086]

Beck AT, Epstein N, Brown G, & Steer RA (1988). An inventory for measuring clinical anxiety: Psychometric properties. Journal of Consulting and Clinical Psychology, 56(6), 893–897. [PubMed: 3204199]

Beck AT, Kovacs M, & Weissman A (1979). Assessment of suicidal ideation: The scale for suicide ideation. Journal of Consulting and Clinical Psychology, 47(2), 343–352. [PubMed: 469082]

Beck AT, & Steer RA (1991). Manual for the Beck Scale for Suicide Ideation. San Antonio, TX: Psychological Corporation.

Beck AT, Steer RA, & Brown GK (1996). Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation.

Beck AT, Steer RA, & Ranieri WF (1988). Scale for suicide ideation: Psychometric properties of a self-report version. Journal of Clinical Psychology, 44(4), 499–505. [PubMed: 3170753]

Belsher BE, Smolenski DJ, Pruitt LD, Bush NE, Beech EH, Workman DE, … Skopp NA (2019). Prediction models for suicide attempts and deaths: A systematic review and simulation. JAMA Psychiatry, 76(6), 642–651. 10.1001/jamapsychiatry.2019.0174 [PubMed: 30865249]

Bentley KH, Franklin JC, Ribeiro JD, Kleiman EM, Fox KR, & Nock MK (2016). Anxiety and its disorders as risk factors for suicidal thoughts and behaviors: A meta-analytic review. Clinical Psychology Review, 43, 30–46. 10.1016/j.cpr.2015.11.008 [PubMed: 26688478]

Black DW, Blum N, Pfohl B, & Hale N (2004). Suicidal behavior in borderline personality disorder: Prevalence, risk factors. Journal of Personality Disorders, 18(3), 226–239. 10.1521/pedi.18.3.226.35445 [PubMed: 15237043]

Bernert RA, Kim JS, Iwata NG, & Perlis ML (2015). Sleep disturbances as an evidence-based suicide risk factor. Current Psychiatry Reports, 17(75), 554 10.1007/s11920-015-0554-4 [PubMed: 25698339]

Boudreaux ED, & Horowitz LM (2014). Suicide risk screening and assessment. American Journal of Preventive Medicine, 47(3), S163–S169. 10.1016/j.amepre.2014.06.005 [PubMed: 25145734]

Campbell DT, & Fiske DW (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56(2), 81–105. [PubMed: 13634291]

Centers for Disease Control. (2017). WISQARS: Web-Based Injury Statistics Query and Reporting System. Government Database. Retrieved from www.cdc.gov/injury/wisqars

Chu C, Buchman-Schmitt JM, Stanley IH, Hom MA, Tucker RP, Hagan CR, … Joiner TE (2017). The interpersonal theory of suicide: A systematic review and meta-analysis of a decade of cross-national research. Psychological Bulletin, 143(12), 1313–1345. 10.1037/bul0000123 [PubMed: 29072480]

Chu C, Klein KM, Buchman-Schmitt JM, Hom MA, Hagan CR, & Joiner TE (2015). Routinized assessment of suicide risk in clinical practice: An empirically informed update. Journal of Clinical Psychology, 71(12), 1186–1200. 10.1002/jclp.22210 [PubMed: 26287362]

Clark LA, & Watson D (2019). Constructing validity: New developments in creating objective measuring instruments. Psychological Assessment, 31, 1412–1427. 10.1037/pas0000626 [PubMed: 30896212]

Crosby AE, Ortega L, & Melanson C (2011). Self-directed violence surveillance: Uniform definitions and recommended data elements, version 1.0. Atlanta: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control.

Cui B (2018). DataExplorer: Data explorer. Retrieved from https://CRAN.R-project.org/package=DataExplorer

de Winter JCF, Gosling SD, & Potter J (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. Psychological Methods, 21(3), 273–290. 10.1037/met0000079 [PubMed: 27213982]

Dhingra K, Klonsky ED, & Tapola V (2018). An empirical test of the three-step theory of suicide in U.K. university students. Suicide & Life-Threatening Behavior, 49, 478–487. 10.1111/sltb.12437 [PubMed: 29604120]

Diedenhofen B, & Musch J (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. PLOS One, 10(4), e0121945 10.1371/journal.pone.0121945 [PubMed: 25835001]

Dozois DJA, Dobson KS, & Ahnberg JL (1998). A psychometric evaluation of the Beck Depression Inventory-II. Psychological Assessment, 10(2), 83–89.

Dunn TJ, Baguley T, & Brunsden V (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. British Journal of Psychology, 105(3), 399–412. 10.1111/bjop.12046 [PubMed: 24844115]

Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, … Nock MK (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. Psychological Bulletin, 143(2), 187–232. 10.1037/bul0000084 [PubMed: 27841450]

Fydrich T, Dowdall D, & Chambless DL (1992). Reliability and validity of the beck anxiety inventory. Journal of Anxiety Disorders, 6(1), 55–61. 10.1016/0887-6185(92)90026-4

Graham JW, Olchowski AE, & Gilreath TD (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prevention Science, 8(3), 206–213. 10.1007/s11121-007-0070-9 [PubMed: 17549635]

Grolemund G, & Wickham H (2011). Dates and times made easy with lubridate. Journal of Statistical Software, 40(3), 1–25. Retrieved from http://www.jstatsoft.org/v40/i03/

Hall RCW, Platt DE, & Hall RCW (1999). Suicide risk assessment: A review of risk factors for suicide in 100 patients who made severe suicide attempts: Evaluation of suicide risk in a time of managed care. Psychosomatics, 40, 18–27. 10.1016/S0033-3182(99)71267-3 [PubMed: 9989117]

Hom MA, Joiner TE, & Bernert RA (2016). Limitations of a single-item assessment of suicide attempt history: Implications for standardized suicide risk assessment. Psychological Assessment, 28(8), 1026–1030. 10.1037/pas0000241 [PubMed: 26502202]

Hom MA, Stanley IH, & Joiner TE (2015). Evaluating factors and interventions that influence help-seeking and mental health service utilization among suicidal individuals: A review of the literature. Clinical Psychology Review, 40, 28–39. 10.1016/j.cpr.2015.05.006 [PubMed: 26048165]

Hopwood CJ, Thomas KM, Markon KE, Wright AGC, & Krueger RF (2012). DSM-5 personality traits and DSM-IV personality disorders. Journal of Abnormal Psychology, 121(2), 424–432. 10.1037/a0026656 [PubMed: 22250660]

Horowitz LM, Ballard ED, & Pao M (2009). Suicide screening in schools, primary care and emergency departments. Current Opinion in Pediatrics, 21(5), 620–627. 10.1097/MOP.0b013e3283307a89 [PubMed: 19617829]

Joiner TE (2005). Why people die by suicide. Cambridge: Harvard University Press.

Joiner TE, Pfaff JJ, & Acres JG (2002). A brief screening tool for suicidal symptoms in adolescents and young adults in general health settings: Reliability and validity data from the Australian National General Practice Youth Suicide Prevention Project. Behaviour Research and Therapy, 40, 471–481. [PubMed: 12008659]

Joiner TE, Walker RL, Rudd MD, & Jobes DA (1999). Scientizing and routinizing the assessment of suicidality in outpatient practice. Professional Psychology: Research and Practice, 30(5), 447–453.

Kleiman EM, Turner BJ, Fedor S, Beale EE, Huffman JC, & Nock MK (2017). Examination of real-time fluctuations in suicidal ideation and its risk factors: Results from two ecological momentary assessment studies. Journal of Abnormal Psychology, 126(6), 726–738. 10.1037/abn0000273 [PubMed: 28481571]

Klonsky ED, & May AM (2015). The three-step theory (3ST): A new theory of suicide in the "ideation-to-action" framework. International Journal of Cognitive Therapy, 8(2), 114–129. 10.1521/ijct.2015.8.2.114

Klonsky ED, Saffer BY, & Bryan CJ (2018). Ideation-to-action theories of suicide: A conceptual and empirical update. Current Opinion in Psychology, 22, 38–43. 10.1016/j.copsyc.2017.07.020 [PubMed: 30122276]

Krueger RF, Derringer J, Markon KE, Watson D, & Skodol AE (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. Psychological Medicine, 42(9), 1879–1890. 10.1017/S0033291711002674 [PubMed: 22153017]

Lambert Ce, Troister T, Ramadan Z, Montemarano V, Fekken GC, & Holden RR (2020). Psychache predicts suicide attempter status change in students starting university. Suicide and Life-Threatening Behavior, 10.1111/sltb.12624

Linehan MM, Comtois KA, & Ward-Ciesielski EF (2012). Assessing and managing risk with suicidal individuals. Cognitive and Behavioral Practice, 19(2), 218–232. 10.1016/j.cbpra.2010.11.008

Liu Y, Zhang J, & Sun L (2017). Who are likely to attempt suicide again? A comparative study between the first and multiple timers. Comprehensive Psychiatry, 78, 54–60. 10.1016/j.comppsych.2017.07.007 [PubMed: 28803042]

McNiel DE, Fordwood SR, Weaver CM, Chamberlain JR, Hall SE, & Binder RL (2008). Effects of training on suicide risk assessment. Psychiatric Services, 59(12), 1462–1465. 10.1176/appi.ps.59.12.1462 [PubMed: 19033175]

Metalsky GI, & Joiner TE (1997). The Hopelessness Depression Symptom Questionnaire. Cognitive Therapy and Research, 21, 359–384. 10.1023/A:1021882717784

Millner AJ, Lee MD, & Nock MK (2015). Single-item measurement of suicidal behaviors: Validity and consequences of misclassification. PLOS One, 10(10):e0141606. [PubMed: 26496707]

Montemarano V, Troister T, Lambert CE, & Holden RR (2018). A four-year longi tudinal study examining psychache and suicide ideation in elevated-risk undergraduates: A test of Shneidman's model of suicidal behavior. Journal of Clinical Psychology, 1–13. 10.1002/jclp.22639

Morin CM (1993). Insomnia: Psychological assessment and management. New York, NY: Guilford Press.

Morin CM, Belleville G, Bélanger L, & Ivers H (2011). The Insomnia Severity Index: Psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep, 34, 601–608. [PubMed: 21532953]

Nock MK, Millner AJ, Joiner TE, Gutierrez PM, Han G, Hwang I, … Smoller JW (2018). Risk factors for the transition from suicide ideation to suicide attempt: Results from the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). Journal of Abnormal Psychology, 127(2), 139–149. 10.1037/abn0000317 [PubMed: 29528668]

Osman A, Kopper BA, Barrios F, Gutierrez PM, & Bagge CL (2004). Reliability and validity of the Beck Depression Inventory-II with adolescent psychiatric inpatients. Psychological Assessment, 16(2), 120–132. 10.1037/1040-3590.16.2.120 [PubMed: 15222808]

Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, … Mann JJ (2011). The Columbia Suicide Severity Rating Scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. The American Journal of Psychiatry, 168(12), 1266–1277. 10.1176/appi.ajp.2011.10111704 [PubMed: 22193671]

R Core Team. (2018). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing Retrieved from https://www.R-project.org/

Revelle W (2018). Psych: Procedures for psychological, psychometric, and personality research. Evanston, IL: Northwestern University Retrieved from https://CRAN.R-project.org/package=psych

Ribeiro JD, Bender TW, Selby EA, Hames JL, & Joiner TE (2011). Development and validation of a brief self-report measure of agitation: The brief agitation measure. Journal of Personality Assessment, 93(6), 597–604. 10.1080/00223891.2011.608758 [PubMed: 21999383]

Ribeiro JD, Huang X, Fox KR, Walsh CG, & Linthicum KP (2019). Predicting imminent suicidal thoughts and nonfatal attempts: The role of complexity. Clinical Psychological Science, 5, 371 10.1177/2167702619838464

Ribeiro JD, Witte TK, Van Orden KA, Selby EA, Gordon KH, Bender TW, & Joiner TE (2014). Fearlessness about death: The psychometric properties and construct validity of the revision to the Acquired Capability for Suicide Scale. Psychological Assessment, 26(1), 115–126. 10.1037/a0034858 [PubMed: 24274043]

Robitzsch A, Grund S, & Henke T (2018). Miceadds: Some additional multiple imputation functions, especially for mice. Retrieved from https://CRAN.R-project.org/package=miceadds

Rogers ML, Hom MA, Dougherty SP, Gallyer AJ, & Joiner TE (2018). Comparing suicide risk factors among individuals with a history of aborted, interrupted, and actual suicide attempts. Archives of Suicide Research.

Rogers ML, & Joiner TE (2017). Severity of suicidal ideation matters: Reexamining correlates of suicidal ideation using quantile regression. Journal of Clinical Psychology, 74(3), 442–452. 10.1002/jclp.22499 [PubMed: 28672075]

Rogers ML, Ringer FB, & Joiner TE (2016). A meta-analytic review of the association between agitation and suicide attempts. Clinical Psychology Review, 48, 1–6. 10.1016/j.cpr.2016.06.002 [PubMed: 27348187]

Rubin DB (2004). Multiple imputation for nonresponse in surveys (Vol. 81). John Wiley & Sons.

Rufino KA, Marcus DK, Ellis TE, Boccaccini MT, Rufino C, Marcus KA, … Boccaccini TE (2018). Further evidence that suicide risk is categorical: A taxometric analysis of data from an inpatient sample. Psychological Assessment, 30(11), 1541–1547. 10.1037/pas0000613 [PubMed: 30070556]

Schafer JL, & Graham JW (2002). Missing data: Our view of the state of the art. Psychological Methods, 7(2), 147–177. 10.1037//1082-989X.7.2.147 [PubMed: 12090408]

Shea SC (2002). The practical art of suicide assessment: A guide for mental health professionals and substance abuse counselors. Hoboken, NJ: John Wiley & Sons, Ltd.

Shneidman ES (1993). Suicide as psychache. The Journal of Nervous and Mental Disease, 181(3), 145–147. 10.1097/00005053-199303000-00001 [PubMed: 8445372]

Sprinkle SD, Lurie D, Insko SL, Atkinson G, Jones GL, Logan AR, & Bissada NN (2002). Criterion validity, severity cut scores, and test-retest reliability of the Beck Depression Inventory-II in a

university counseling center sample. Journal of Counseling Psychology, 49(3), 381–385. 10.1037//0022-0167.49.3.381

Stanley B, & Brown GK (2012). Safety planning intervention: A brief intervention to mitigate suicide risk. Cognitive and Behavioral Practice, 19, 256–264. 10.1016/j.cbpra.2011.01.001

Steer RA, Ball R, Ranieri WF, & Beck AT (1997). Further evidence for the construct validity of the Beck Depression Inventory-II with psychiatric outpatients. Psychological Reports, 80, 443–446. [PubMed: 9129364]

Steer RA, Ranieri WF, Beck AT, & Clark DA (1993). Further evidence for the validity of the Beck anxiety inventory with psychiatric outpatients. Journal of Anxiety Disorders, 7(3), 195–205. 10.1016/0887-6185(93)90002-3

Strickland CM, Drislane LE, Lucy M, Krueger RF, & Patrick CJ (2013). Characterizing psychopathy using DSM-5 personality traits. Assessment, 20, 327–338. 10.1177/1073191113486691 [PubMed: 23620353]

U.S. Department of Health and Human Services. (2012). Office of the Surgeon General and National Action Alliance for Suicide Prevention. Washington, DC: National Strategy for Suicide Prevention: Goals and Objectives for Action, 2012.

van Buuren S, & Groothuis-Oudshoorn K (2011). mice: Multivariate imputation by chained equations in r. Journal of Statistical Software, 45(3), 1–67. Retrieved from https://www.jstatsoft.org/v45/i03/

Van Orden KA, Cukrowicz KC, Witte TK, & Joiner TE (2012). Thwarted belongingness and perceived burdensomeness: Construct validity and psychometric properties of interpersonal needs questionnaire. Psychological Assessment, 24(1), 197–215. [PubMed: 21928908]

Van Orden KA, Witte TK, Cukrowicz KC, Braithwaite SR, Selby EA, & Joiner TE (2010). The interpersonal theory of suicide. Psychological Review, 117(2), 575–600. 10.1037/a0018697 [PubMed: 20438238]

Van Orden KA, Witte TK, Gordon KH, Bender TW, & Joiner TE (2008). Suicidal desire and the capability for suicide: Tests of the interpersonal-psychological theory of suicidal behavior among adults. Journal of Consulting and Clinical Psychology, 76(1), 72–83. 10.1037/0022-006X.76.1.72 [PubMed: 18229985]

Walsh CG, Ribeiro JD, & Franklin JC (2017). Predicting risk of suicide attempts over time through machine learning. Clinical Psychological Science, 5(3), 457–469.

Walsh CG, Ribeiro JD, & Franklin JC (2018). Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning. Journal of Child Psychology and Psychiatry, 59(12), 1261–1270. 10.1111/jcpp.12916 [PubMed: 29709069]

Wickham H (2017). Tidyverse: Easily install and load the “tidyverse”. Retrieved from https://CRAN.R-project.org/package=tidyverse

Wickham H (2018). Scales: Scale functions for visualization. Retrieved from https://CRAN.R-project.org/package=scales

Wickham H, & Miller E (2019). Haven: Import and export“spss”, “stata” and “sas” files. Retrieved from https://CRAN.R-project.org/package=haven

Witte TK, Holm-Denoma JM, Zuromski KL, Gauthier JM, & Ruscio J (2017). Individuals at high risk for suicide are categorically distinct from those at low risk. Psychological Assessment, 29(4), 382–393. [PubMed: 27280745]

Zinbarg RE, Revelle W, Yovel I, & Li W (2005). Cronbach’s, Revelle’s, and Mcdonald’s H: Their relations with each other and two alternative conceptualizations of reliability. Psychometrika, 70(1), 123–133. 10.1007/s11336-003-0974-7

Zou GY (2007). Toward using confidence intervals to compare correlations. Psychological Methods, 12(4), 399–413. 10.1037/1082-989X.12.4.399 [PubMed: 18179351]

TABLE 1

Correlations between study variables and descriptive statistics (N = 731)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. DT suicide risk (T1)															
2. DT suicide risk (T2)	0.76														
3. SA history	0.53	0.55													
4. SA recency	−0.43	−0.45	−0.78												
5. BSS	0.58	0.51	0.31	−0.28											
6. BSS (desire)	0.53	0.51	0.30	−0.27	0.78										
7. DSI-SS	0.63	0.64	0.36	−0.30	0.63	0.64									
8. INQ-PB	0.54	0.59	0.36	−0.26	0.52	0.55	0.54								
9. INQ-TB	0.29	0.32	0.27	−0.24	0.34	0.35	0.37	0.56							
10. ACSS-FAD	0.06	−0.03	0.04	−0.07	0.05	0.02	0.00	−0.08	−0.06						
11. BDI	0.43	0.46	0.37	−0.33	0.46	0.56	0.50	0.64	0.60	−0.14					
12. BAI	0.29	0.31	0.25	−0.19	0.24	0.35	0.31	0.38	0.36	−0.25	0.65				
13. BAM	0.31	0.32	0.27	−0.22	0.31	0.37	0.43	0.44	0.49	−0.19	0.67	0.60			
14. ISI	0.25	0.27	0.21	−0.15	0.17	0.23	0.25	0.22	0.22	−0.10	0.49	0.46	0.39		
15. PID-5 BPD	0.20	0.16	0.14	−0.07	0.20	0.25	0.32	0.33	0.34	−0.07	0.35	0.29	0.33	0.17	
Mean	1.29	1.24	0.31	3.62	11.66	5.29	1.25	13.73	32.31	14.8	21.07	15.34	5.82	11.03	1.32
SD	1.06	0.95	0.62	0.85	6.55	3.28	2.05	8.95	12.80	6.51	12.64	11.82	4.04	6.58	0.50

Note: Shaded cells indicate significant correlations at $p < .05$. Correlations are pooled Spearman's correlations on 100 multiple imputation data sets. Also included are means and standard deviations of variables before multiple imputation.

Abbreviations: ACSS-FAD, Acquired Capability for Suicide Scale–Fearlessness About Death; BAI, Beck Anxiety Inventory; BAM, Brief Agitation Measure; BDI, Beck Depression Inventory; BSS, Beck Scale for Suicide Ideation; DSI-SS, Depressive Symptom Index–Suicidality Subscale; DT, Decision Tree; INQ-PB, Interpersonal Needs Questionnaire–Perceived Burdensomeness; INQ-TB, Interpersonal Needs Questionnaire–Thwarted Belongingness; ISI, Insomnia Severity Index; PID-5 BPD, Personality Inventory for the DSM-5 Borderline Personality Disorder Scale; SA, suicide attempt; T1, at screening; T2, at intake.

TABLE 2

Differences in correlations with DT suicide risk (T1) between study variables

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. DT suicide risk (T2)													
2. SA history	0.18, 0.28												
3. SA recency	1.12, 1.27	0.86, 1.07											
4. BSS	0.13, 0.23	−0.11, 0.02	−1.09, −0.93										
5. BSS (desire)	0.18, 0.29	−0.06, 0.07	−1.05, −0.88	0.01, 0.09									
6. DSI-SS	0.1, 0.18	−0.16, −0.03	−1.14, −0.98	−0.09, 0	−0.15, −0.05								
7. INQ-PB	0.18, 0.27	−0.07, 0.06	−1.05, −0.89	−0.01, 0.1	−0.07, 0.05	0.04, 0.14							
8. INQ-TB	0.41, 0.54	0.17, 0.32	−0.82, −0.62	0.22, 0.37	0.17, 0.31	0.27, 0.41	0.19, 0.31						
9. ACSS-FAD	0.62, 0.78	0.38, 0.56	−0.59, −0.4	0.44, 0.61	0.38, 0.56	0.48, 0.65	0.39, 0.57	0.13, 0.33					
10. BDI	0.28, 0.4	0.04, 0.18	−0.95, −0.76	0.09, 0.22	0.05, 0.16	0.14, 0.26	0.06, 0.17	−0.2, −0.08	−0.47, −0.26				
11. BAI	0.4, 0.54	0.16, 0.32	−0.82, −0.63	0.21, 0.37	0.16, 0.31	0.26, 0.41	0.18, 0.32	−0.08, 0.07	−0.34, −0.12	0.08, 0.19			
12. BAM	0.39, 0.52	0.15, 0.3	−0.83, −0.65	0.2, 0.35	0.15, 0.29	0.25, 0.39	0.16, 0.3	−0.09, 0.05	−0.35, −0.14	0.06, 0.17	−0.08, 0.05		
13. ISI	0.45, 0.59	0.21, 0.37	−0.77, −0.58	0.26, 0.42	0.21, 0.36	0.31, 0.46	0.21, 0.37	−0.04, 0.13	−0.29, −0.08	0.11, 0.25	−0.03, 0.12	−0.01, 0.14	
14. PID-5 BPD	0.49, 0.63	0.25, 0.41	−0.73, −0.54	0.3, 0.46	0.25, 0.4	0.35, 0.5	0.26, 0.41	0, 0.16	−0.25, −0.04	0.14, 0.3	0.01, 0.17	0.02, 0.18	−0.05, 0.13

Note: Shaded cells indicate significant difference between correlations with DT Suicide Risk (T1). Cells contain 95% confidence interval of the difference between correlations with DT suicide risk (T1).

Abbreviations: ACSS-FAD, Acquired Capability for Suicide Scale–Fearlessness About Death; BAI, Beck Anxiety Inventory; BAM, Brief Agitation Measure; BDI, Beck Depression Inventory; BSS, Beck Scale for Suicide Ideation; DSI-SS, Depressive Symptom Index–Suicidality Subscale; DT, Decision Tree; INQ-PB, Interpersonal Needs Questionnaire–Perceived Burdensomeness; INQ-TB, Interpersonal Needs Questionnaire–Thwarted Belongingness; ISI, Insomnia Severity Index; PID-5 BPD, Personality Inventory for the DSM-5 Borderline Personality Disorder Scale; SA, suicide attempt; T1, at screening; T2, at intake.