Suicide and Life-Threatening BEHAVIOR
THE OFFICIAL JOURNAL OF THE AMERICAN ASSOCIATION OF SUICIDOLOGY

Suicide and Life-Threatening Behavior 50 (2) April 2020

DOI: 10.1111/sltb.12609

A Preliminary Precision Treatment Rule for Remission of Suicide Ideation

Ronald C. Kessler, PhD, Samantha A. Chalker, MA, Alex R. Luedtke, PhD, Ekaterina Sadikova, MPH and David A. Jobes, PhD

Objective: There is growing interest in the development of composite precision treatment rules (PTRs) to guide the selection of the treatments most likely to be helpful for individual patients. We present here the results of an effort to develop a preliminary PTR for Collaborative Assessment and Management of Suicidality (CAMS) relative to enhanced-care as usual based on secondary analysis of the Operation Worth Living (OWL) randomized controlled trial. The outcome of interest is eliminating suicide ideation (SI) within 3 months of initiating treatment.
Method: A state-of-the-art ensemble machine learning method was used to develop the PTR among the n = 148 U.S. Soldiers (predominately male and White, age range 18–48) OWL patients.
Results: We estimated that CAMS was the better treatment for 77.8% of patients and that treatment assignment according to the PTR would result in a 13.6% (95% CI: 0.9%–26.3%) increase in 3-month SI remission compared to random treatment assignment.
Conclusions: Although promising, results are limited by the small sample size, restrictive baseline assessment, and inability to evaluate effects on suicidal behaviors or disaggregate based on history of suicidal behaviors. Replication is needed in larger samples with comprehensive baseline assessments, longer-term follow-ups, and more extensive outcomes.

Approximately 47,000 Americans died by suicide in 2017 (Drapeau & McIntosh, 2018), 1.4 million made suicide attempts, and a striking 10.6 million had serious suicidal ideation (SI; SAMHSA, 2017). There are a number of proven suicide-focused interventions for reducing suicide attempts (SA; Brown et al., 2005; Bryan et al., 2017; DeCou, Comtois, & Landes, 2019; Gysin-Maillart, Schwab, Soravia, Megert, & Michel, 2016; Rudd et al., 2015; Stanley et al., 2018). However, these treatments do not show strong effects in reducing SI relative to controls. The Collaborative Assessment and Management of Suicidality (CAMS), in comparison, has reliably demonstrated replicated effectiveness in reducing SI relative to controls across a number of clinical trials (Comtois et al., 2011; Ellis et al., 2015, 2017; Jobes, Wong, Conrad, Drozd & Neal-Walden, 2005, 2017; Ryberg, Zahl, Diep, Landro, & Fosse, 2019). As argued by Jobes and Joiner

Ronald C. Kessler, and Ekaterina Sadikova, Department of Health Care Policy, Harvard Medical School, Boston, MA, USA; Samantha A. Chalker, and David A. Jobes, Department of Psychology, The Catholic University of America, Washington, DC, USA; Alex R. Luedtke, Department of Statistics, University of Washington, Seattle, WA, USA and Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.
Address correspondence to R. C. Kessler, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115; E-mail: kessler@hcp.med.harvard.edu

(2019), SI is an outcome of interest in its own right in that it is a marker of severe distress and is also a strong predictor of suicidal behavior even though we know that the predictors of SI are different from the predictors of the transition from SI to suicidal behavior (Klonsky, May, & Saffer, 2016).

Differences in the estimated effects of suicide-specific clinical interventions across RCT's may be due to different mechanisms and different underlying study samples. For example, Dialectical Behavior Therapy (see meta-analysis by Decou et al., 2019) and the two forms of suicide-focused cognitive behavioral therapy (Brown et al., 2005; Rudd et al., 2015) specifically emphasize the development of skills (DBT) and changing coping behaviors within the "suicide mode" model (CBT-based care) and often focus on patients who made prior suicide attempts (SAs). In contrast, hypothesized mechanisms for CAMS emphasize changing suicidal cognitions (e.g., Ellis et al., 2015) and the installation of hope (e.g., Comtoise et al., 2011). While there are mixed data about CAMS reducing SA and self-harm behaviors (Andreasson et al., 2016), many of the CAMS trials have studied relatively lower risk suicidal ideators (with more limited attempt histories) in comparison to the previously noted suicide-focused treatments that specifically recruit suicide attempters (e.g., Brown et al., 2005).

Differential treatment responses across the clinical trials cited above raise the very real possibility that different suicidal populations should be targeted with different treatments for optimal clinical outcomes and health care cost savings (Jobes, 1995; Jobes, Gregorian, & Colborn, 2018). For example, the optimal treatment to promote remission of SI among patients without a history of SA might be quite different from the optimal treatment to prevent recurrence of SA. These are the goals of precision medicine: to understand how the effects of treatment are modified by patient characteristics and to develop "precision treatment rules" (PTRs) based on this understanding to determine which of the treatments under consideration is likely to yield the best outcome for each patient or fine-grained patient subgroup (Kessler, Bossarte, Luedtke, Zaslavsky, & Zubizarreta, 2019). However, we lack guidance on such optimal treatment matching for suicide-focused interventions.

We present here the results of an attempt to develop a preliminary PTR for the effect of CAMS relative to enhanced-care as usual (E-CAU) in the elimination of SI within 3 months of initiating treatment in a secondary analysis of the "Operation Worth Living" (OWL) randomized controlled trial (Jobes et al., 2017). CAMS is a well-studied suicide-specific therapeutic framework that targets and treats patient-defined "suicidal-drivers" (Jobes, 2016; Jobes, Comtois, Brenner, Gutierrez, & O'Connor, 2016). In the OWL trial, 148 U.S. Army Soldiers with significant SI (>13 on the Beck Scale for Suicide Ideation) were randomized to either CAMS (n = 73) or enhanced-care as usual (E-CAU; n = 75). Significant aggregate treatment responses were seen in both arms of the RCT across a range of outcomes at four follow-up assessments (1, 3, 6, and 12 months). However, the only significant difference between CAMS and E-CAU was that SI was resolved more quickly in the CAMS arm than the E-CAU arm, as documented by a significantly lower rate of continued SI among CAMS than E-CAU patients at 3 months (Jobes et al., 2017). Although no significant difference in suicidal behaviors was found between patients in the CAMS and E-CAU arms, the sample was underpowered to detect such differences. And an even larger sample would be needed to evaluate the hypothesis that any effect of CAMS in reducing suicidal behavior is due to CAMS resolving SI more quickly than E-CAU.

A subsequent secondary analysis of the OWL data evaluated the effects of nine potential treatment moderators assessed at baseline on six suicide-related and mental health outcomes, including four sociodemographics (age, sex, race, and marital status), two aspects of military history (time in service and combat deployments), borderline personality disorder, history of one or more prior suicide attempts at the beginning of

treatment, and level of suicidality at the beginning of treatment (Huh et al., 2018). Eight significant moderator interactions were found, six of them favoring CAMS. However, no attempt was made to combine these interactions to develop a PTR. This is the goal of the current study. We also expand the potential modifiers to include a wider range of variables assessed in OWL intake interviews. The focus on 3-month SI remission is based on this being the only primary outcome that showed significant separation between CAMS and E-CAU within the aggregate OWL RCT analyses.

A number of prior studies have attempted to combine multiple interactions along the same basic lines as we do here to develop composite PTRs for treatment of major depression (reviewed by Cohen & DeRubeis, 2018). Similar efforts are also beginning to appear in studies of other mental disorders (e.g., Fernandes et al., 2017; van den Oord, Chan, & Aberg, 2018). However, up until now there has been no research on PTRs for suicide-focused interventions. This is unfortunate because it would be reasonable to expect more heterogeneity in the effects of interventions focused on suicidality than on specific mental disorders because suicidality occurs in conjunction with manifold mental disorders and interventions focused on suicidality target intermediate outcomes (e.g., increased perceptions of belongingness in caring text interventions) that might be issues for only a subset of patients, leading to heterogeneity in treatment effects (Kessler, 2019).

METHODS

Participants and Setting

The 148 OWL participants were Soldiers seeking treatment in the Division of Behavioral Health at a U.S. Army installation in the Southeastern U.S. who reported SI at intake, were judged by intake workers to be at elevated suicide risk, and subsequently had scores of at least >13 on the Scale for Suicide Ideation (SSI; Beck, Kovacs, & Weissman, 1979). Exclusion criteria included clinical indicators that psychosocial care was inappropriate (e.g., psychosis, cognitive impairment), involuntary treatment, separation or transfer expected in the next 12 months, and pregnancy (as required by the treatment site institutional review board). All study procedures were reviewed and approved by the Department of Defense's Institutional Review Board (IRB) and Human Research Protection Office (HRPO), as well as the University of Washington, The Catholic University of America, and the Denver Veterans Affairs IRBs. A data safety monitoring board (DSMB) oversaw the study.

Clinicians and Treatment

The study therapists were all psychiatric social workers except for one MA-level mental health counselor. In order to adjust for differences in intensity of treatment across arms, E-CAU required at least four sessions offered to patients and therapists were provided with clinical consultation meetings, but no training was offered on type of psychotherapy required. The content of CAMS intervention is described in detail elsewhere (Jobes, 2016). CAMS clinicians received a 1.5-day training from DAJ and participated in weekly group consultation calls. Review of video recordings of initial CAMS sessions with all patients and random checks with 10% of patients were used to confirm between-group fidelity, treatment adherence, and absence of drift (Corona, Gutierrez, Wagner, & Jobes, 2018).

Measures

The predictors considered in developing the PTR included diverse measures of patient sociodemographics, military history, suicidality, psychopathology, and other predictors obtained in baseline OWL intake interviews. Importantly, only predictors assessed at intake were included in the PTR model.

Sociodemographics. These included patient age, gender, race (categorized as

Black, other, and White), and dichotomous variables for being currently married, having at least one child, U.S. nativity, and having completed at least some college.

Military Status And Service History. These included information about time since entering service, number of deployments (0–3+), and a dummy variable indicating a history of at least one combat deployment to Afghanistan or Iraq. The Treatment History Interview—Military (THI-M; Linehan, 1993) was used to capture information about prior behavioral health care.

Suicidality. The primary outcome, 3-month SI, was assessed with the Scale for Suicide Ideation-Current (Beck, Brown, & Steer, 1997). Baseline assessments of current SI and lifetime worst SI were made with the same scale and used as predictors of the 3-month outcome. In addition, the Suicide Attempt Self-Injury Count (Linehan & Comtois, 1996; Linehan, Comtois, Brown, Heard, & Wagner, 2006) was administered at intake and used to create nested dichotomies for whether the patient had a lifetime history of SA and nonsuicidal self-injury (NSSI; separate yes–no dummy variables for one and more than one lifetime occurrence of each) and as well as reports about age at first, most recent, and most lethal SA and NSSI.

Psychopathology. Lifetime and current baseline prevalence of DSM-IV major depressive disorder, PTSD, other anxiety disorders, and substance use disorders were assessed with the Structured Clinical Interview (SCID) for DSM-IV (First, 1997a). Borderline personality disorder was assessed with the SCID for Axis II (First, 1997b). Severity of DSM-IV PTSD (range 17–85) was assessed with the PTSD Checklist—Military Version (PCL-M; Weathers, Huska, & Keane, 1991). The Outcome Questionnaire (OQ-45; Lambert et al., 1996) was used to calculate a total index of psychiatric distress (range 0–180) and the mental health symptom distress subscale (range 0–100). The Neurobehavioral Symptom Inventory (NSI; Kaplan, 2014; Meterko et al., 2012) was used to measure vestibular–sensory symptoms (range 0–44), cognitive symptoms (range 0–16), and mood symptoms (range 0–28). The SF-36 was used to assess eight subscales—general health, physical functioning, role-physical, bodily pain, vitality, social functioning, mental health, and role-emotional subscales, each with a 0–100 range (McHorney, Ware, Lu, & Sherbourne, 1994)—as well as mental component and physical component summary scores based on those subscales. The latter were standardized to a mean of 50 and standard deviation of 10 in a benchmark U.S. general population sample (Ware et al., 1995).

Other Predictors. The Connor–Davidson Resilience Scale (CD-RISC; Connor & Davidson, 2003) was used to assess patient-perceived resilience to stress (range 0–100). Patients and therapists were asked separately about their expectancies of treatment success prior to the first session (range 0–7).

Data Analyses

Balancing on Baseline Covariates. The baseline covariates used to develop PTRs typically differ somewhat across treatment arms of experiments, making it useful, although not necessary, to balance on these differences. We did this using the super learner (SL) ensemble machine learning (ML) method (van der Laan, Polley, & Hubbard, 2007) to create a propensity score and then used balancing weights to adjust the data for baseline imbalance (Li, Morgan, & Zaslavsky, 2017). The R package SuperLearner was used to implement the analysis (Polley, LeDell, Kennedy, Lendle, & van der Laan, 2018). SL uses cross-validation (CV) to select a weighted combination of predicted outcome scores across a collection (ensemble) of candidate ML algorithms that yields a nearly optimal weighted combination guaranteed to perform as well or nearly as well as the best component algorithm according to a prespecified criterion (van der Laan, Polley, & Hubbard, 2007).

Adjusting for Loss to Follow-up. Twenty-four of the 148 patients originally randomized were lost to follow-up by the 3-month postrandomization assessment.

Stepwise logistic regression was used to estimate the probability of remaining in the study through the 3-month assessment from baseline covariates. Inverse probability weighting (IPW) based on this model was used to adjust the remaining $n = 124$ cases for loss to follow-up (Mansournia & Altman, 2016).

Estimating the PTR. The approach we used to estimate the PTR is based on an extension of the same SL ML method that was used to balance baseline covariates. This approach begins by using SL to estimate the outcome separately in each of the two treatment arms and then generating a preliminary predicted outcome score for each patient under each of the two treatment conditions based on these results (Luedtke & van der Laan, 2017). A propensity-weighted difference between the two conditional predicted outcome scores is then used as the outcome in a second SL model, where interactions between treatment and baseline covariates are estimated directly (i.e., the outcome is an estimate of the within-patient difference in treatment response across treatment arms), avoiding the need to estimate main effects. The resulting individual-level predicted difference score is then used to estimate which treatment would be optimal for each patient. CAMS was estimated to be optimal if the predicted difference score was positive and E-CAU if the predicted difference score was negative. This approach has advantages over more conventional methods of estimating a PTR, most of which require correct specification of both the (possibly nonlinear) main effects and the (possibly complex nonlinear and higher-order) interactions to estimate the PTR, as only the interactions need to be specified correctly in the SL approach (Luedtke & van der Laan, 2017). This SL-based approach also improves on other approaches that estimate interactions directly by using ensemble ML rather than relying on any single algorithm to specify interactions correctly. The R package sg was used to implement the analysis (Luedtke, 2018).

Recent recommendations suggest using a wide range of candidate algorithms in the SL ensemble to optimize performance (LeDell, van der Laan, & Petersen, 2016). To this end, we used a library of 14 parametric and flexible candidate algorithms implemented in R in our ensemble (Table 1) both for baseline covariate balancing and to estimate the PTR: a generalized linear model, five penalized logistic regressions that differed in values of the mixing parameter, two splines (adaptive splines and adaptive polynomial splines), two decision trees (random forest and Bayesian additive regression trees), three support vector machines (with linear, polynomial, and radial kernels), and neural networks. A brief overview of these algorithms is presented in Table 1. For each algorithm, we used two different methods to screen predictors for entry into SL, a univariate p-value screen with $p = .05$ and an elastic net screen with the mixing parameter set to 0.5, for a total of 28 different combinations between algorithms and screeners. The estimated aggregate effect of optimal treatment allocation was quantified by using a cross-validated targeted minimum loss-based estimator (CV-TMLE) of the attained improvement in the outcome based on optimized treatment compared to randomized care (van der Laan & Luedtke, 2015). The use of CV-TMLE results in a nested cross-validation scheme, wherein the outer cross-validation was used to validate the performance of the treatment rule fitted via the inner cross-validation strategy used by the SL ensemble algorithm. See van der Laan and Luedtke (2015) for further details on this approach.

RESULTS

Baseline Distribution of Sociodemographic and Clinical Characteristics

Operation Worth Living participants were predominately male (80.4%), White (53.2%), married (50.7%), and attended some college (52.7%; Table 2). In terms of military descriptors, over two thirds (69.6%) were of junior enlisted rank (E1–E4) and 41.5% never had a combat deployment. Half (50.0%) had never made a suicide attempt in their lifetime.

TABLE 1
Summary of Machine Learning Methods

Algorithm	R package	Description
I. Super learner	*SuperLearner* (van der Laan, Polley, & Hubbard, 2007)	Super learner is an ensemble machine learning approach that uses cross-validation (CV) to select a weighted combination of predicted outcome scores across a collection of candidate algorithms (learners) to yield an optimal combination according to a prespecified criterion that performs as least as well as the best component algorithm
II. PTR estimation and evaluation	*sg* (Luedtke & van der Laan, 2017)	This doubly robust framework uses SuperLearner to estimate the conditional average treatment effect to build a PTR and uses a cross-validated targeted maximum likelihood estimator to evaluate the marginal improvement resulting from implementing the PTR in a population
III. Learners in the super learner library		
a. Generalized linear models	*stats* (Nelder & Wedderburn, 1972)	Maximum likelihood estimation with flexible link function
b. Elastic net MPP = 0.1 MPP = 0.3 MPP = 0.5 MPP = 0.7 MPP = 0.9	*glmnet* (Friedman, Hastie, & Tibshirani, 2010)	Elastic net is a regularization method that minimizes the problem of overlap among predictors by explicitly penalizing over-fitting with a composite penalty $\lambda\{\text{MPP} \times \text{Plasso} + (1 - \text{MPP}) \times \text{Pridge}\}$, where MPP is a mixing parameter penalty with values between 0 and 1 that controls relative weighting between the lasso penalty (Plasso) and the ridge penalty (Pridge). The parameter λ controls the total amount of penalization. The ridge penalty handles multicollinearity by shrinking all coefficients smoothly toward 0 but retains all variables in the model. The lasso penalty allows simultaneous coefficient shrinkage and variable selection, tending to select at most one predictor in each strongly correlated set, but at the expense of giving unstable estimates in the presence of high multicollinearity. The elastic net approach of combining the ridge and lasso penalties has the advantage of yielding more stable and accurate estimates than either ridge or lasso alone while maintaining model parsimony
c. Spline		
Adaptive splines	*earth* (Milborrow, Hastie, Tibshirani, Miller, & Lumley, 2016)	Adaptive spline regression flexibly captures both linear and piecewise nonlinear associations as well as interactions among these associations by connecting linear segments (splines) of varying slopes and smoothed to create piecewise curves (basis functions). Final fit is built using a stepwise procedure that selects the optimal combination of basis functions

(continued)

TABLE 1
(continued)

Algorithm	R package	Description
Adaptive polynomial splines	*polspline* (Kooperberg, 2015)	Adaptive polynomial splines are similar to adaptive splines, but differ in the order which basis functions (e.g., linear versus nonlinear) are added to build the final model
d. Decision tree Random forest	*randomForest* (Liaw & Wiener, 2002)	Independent variables are partitioned (based on values) and stacked to build decision trees and ensemble an aggregate "forest." Random forests builds numerous trees in bootstrapped samples and generates an aggregate tree by averaging across trees, thereby reducing over-fitting
e. Neural network	*nnet* (Ripley & Venables, 2016)	Feed-forward neural network with a single hidden layer comprising of 2 nodes, used for multinomial log-linear models
f. Bayesian additive regression trees	*bartMachine* (Chipman, George & McCulloch, 2010)	"Sum-of-trees" model where the comprising trees are constrained by a regularization prior. Fitting and inference by Bayesian backfitting MCMC
g. Support vector machine	*svm* (Schölkopf, Platt, Shawe-Taylor, Smola & Williamson, 2001)	Classification algorithm that aims to create maximum distance between classes of outcome separated by an optimally selected hyperplane (linear, polynomial, or radial kernel)

Two thirds (62.6%) met criteria for a major depressive disorder at baseline, 50.7% for baseline post-traumatic stress disorder (PTSD), and 48.9% for some other baseline anxiety disorder. Smaller proportions met criteria for borderline personality disorder (27.5%), alcohol abuse or dependence (15.8%), drug abuse or dependence (4.3%), or bipolar disorder (3.6%). More details on sample composition are reported elsewhere (Jobes et al., 2017; Table S1).

Balancing on Baseline Covariates

As noted above, baseline variables typically differ somewhat across randomized treatment arms. OWL was no exception in this regard. Although most of these differences were small, a few were relatively substantial. Most notably, prevalence of current substance use disorder (alcohol or drug abuse or dependence) was significantly higher in the CAMS arm than the E-CAU arm (23.3% vs. 10.7%, $\chi^2_1 = 4.2$, $p = .041$), the mean SF-36 role-physical score was significantly higher in the CAMS arm than the E-CAU arm (42.5 vs. 37.7, $t = 7.4$, $p = .007$), and the mean NSI mood subscale score was significantly higher in the E-CAU arm than the CAMS arm (18.8 vs. 16.9, $t = 4.0$, $p = .047$). These differences were no longer significant after a balancing weight was used to adjust for differences in baseline covariates ($p = 30–.66$). Similar improvements were seen in balance of baseline covariates with less extreme initial imbalance (Table S1).

Adjusting for Loss to Follow-up

Only one baseline covariate, the presence of lifetime substance disorder (OR = 3.4, 95% CI = 1.4–8.3), was significantly associated with remaining in the study through the 3-month assessment after adjusting for treatment assignment. There was no significant interaction between treatment assignment and substance disorder in predicting loss to follow-up. The IPW generated

TABLE 2
Baseline Sociodemographic and Clinical Characteristics: Overall and by Treatment Condition[a]

	Overall		CAMS		E-CAU	
	N	%	*N*	%	*N*	%
I. Sociodemographics						
Gender						
Male	119	80.4	56	76.7	63	84.0
Female	29	19.6	17	23.3	12	16.0
Ethnicity						
White/Caucasian	75	53.2	37	51.4	38	55.1
Black/African American	34	24.1	17	23.6	17	24.6
Latino/a	15	3.6	12	16.7	3	4.3
Asian or Pacific Islander	5	10.6	2	2.8	3	4.3
Other	12	8.5	4	5.6	8	11.6
Marital status						
Single, never married	38	26.0	20	28.2	18	24.0
Married	74	50.7	35	49.3	39	52.0
Separated or divorced	33	22.6	16	22.5	17	22.7
Widowed	1	0.7	0	0.0	1	1.3
Education						
Some high school	1	0.7	1	1.4	0	0.0
High school graduate or GED	57	39.0	30	42.3	27	36.0
Some college, AA, or technical training	77	52.7	33	22.5	44	58.7
Bachelor's or graduate degree	11	7.5	7	0.0	4	5.3
II. Military status and service history						
Rank						
Junior enlisted (E1–E4)	103	69.6	51	69.9	52	69.3
Noncommissioned officer (E5–E9)	41	27.7	21	28.8	20	26.7
Officer (W2–O3)	4	2.7	1	1.4	3	4.0
Number of combat deployments						
0	61	41.5	31	42.5	30	40.5
1	38	25.9	18	24.7	20	27.0
2	28	19.0	17	23.3	11	14.9
3+	20	13.6	7	9.6	13	17.6
III. Suicidality						
Lifetime suicide attempts						
0	74	50.0	37	50.7	37	49.3
1	34	23.0	16	21.9	18	24.0
2+	40	27.0	20	27.4	20	26.7
Lifetime self-inflicted injuries						
0	57	38.8	27	37.0	30	40.5
1	29	19.7	11	15.1	18	24.3
2+	61	41.5	35	47.9	26	35.2
IV. Current psychopathology						
Major depressive disorder	87	62.6	40	56.3	47	69.1
PTSD	69	50.7	31	45.6	38	55.9
Anxiety disorder (excluding PTSD)	68	48.9	34	47.9	34	50.0
Borderline personality disorder	38	27.5	21	30.0	17	25.0
Alcohol abuse or dependence	22	15.8	15	21.1	7	10.3
Drug abuse or dependence	6	4.3	4	5.6	2	2.9
Bipolar disorder	5	3.6	3	4.2	2	2.9

[a]There were no statistically significant differences between treatment conditions with respect to the variables presented in Table 1.

from a simple logistic model containing dummy variables for treatment assignment and lifetime substance disorder as the only predictors was used to adjust for the difference in the distribution of this variable between the $n = 124$ cases who completed the 3-month assessment and the $n = 24$ cases who were lost to follow-up.

Estimating the PTR

Algorithm Weights to Predict Optimal Treatment. Two of the 28 combinations between algorithms and screeners dominated in the final SL solution to predict optimal treatment: random forest (weight = 0.89 based on the p-value screener) and adaptive splines (weight = 0.11 based on the elastic net screener).

The Distribution of Predicted Differences in Probability of Remission by Treatment. Inspection of the distribution of differences in patient-level predicted probabilities of remission between the two treatments showed that E-CAU was predicted to be optimal for 22.2% of patients and CAMS for the remaining 77.8% (Figure 1). However, the two probabilities were similar (i.e., less than 10% difference in probability of remission) for about 20% of patients. The median estimated benefit of E-CAU among patients for whom E-CAU was predicted to be better was about 15%, whereas the median estimated benefit of CAMS among patients for whom CAMS was predicted to be preferred was about 35%.

The Predicted Impact of Using the PTR on the Overall Remission Rate. We estimated that optimal treatment assignment would result in a 13.6% (95% CI: 0.9%–26.3%) increase in the proportion of patients with 3-month SI remission compared to random treatment assignment. This effect was the difference between the 64.1% (95% CI: 50.0%–78.2%) expected remission rate if all patients were treated optimally (i.e., 77.8% were treated with CAMS) and the expected remission rate under randomization. However, based on the low marginal increase in predicted probability of remission seen in Figure 1 for about 10% of patients for whom CAMS was optimal, the predicted remission rate would be very close to optimal (62.4%; 95% CI: 47.5%–77.3%) if CAMS was provided only to the 65% of patients estimated to profit most from this treatment.

Significant Predictors of Optimal Treatment. Although it is difficult to evaluate the importance of individual predictors in ensemble ML models, a rough idea of variable importance can be obtained by examining the associations of baseline covariates with the patient-level predicted difference scores that define the PTR. We did this using penalized regression, where the mixing parameter was set to 0.9 to produce a sparse model. All variables, both predictors and the outcome, were standardized to a mean of 0 and variance of 1.0 to facilitate comparison across predictors and evaluate strength of associations. An ordinary least squares (OLS) regression model was estimated in parallel with the penalized regression model to generate standard errors, which are not used in penalized regression.

Three variables stood out as being especially strong predictors of having a higher probability of remission with CAMS than E-CAU: White or Black race (compared to others; $\beta = 0.36$–0.39 in elastic net and OLS, respectively), having good physical functioning at the beginning of treatment ($\beta = 0.33$–0.39), and not having a lifetime history of any anxiety disorder (including PTSD) at the beginning of treatment ($\beta = 0.18$–0.20; Table 3). All three of these predictors were statistically significant in the OLS model. The fourth strongest predictor in the elastic net model, having fewer than three prior lifetime deployments, was also significant in the OLS model ($\beta = 0.07$–0.14). In addition, four other predictors that were not among the stronger predictors in the elastic net model were significant in the OLS model: young age at first self-inflicted injury ($\beta = 0.11$); not having a lifetime history of major depressive disorder at the beginning of treatment ($\beta = 0.10$); having borderline personality disorder ($\beta = 0.10$); and the patient having low expectancy for treatment prior to initiating treatment ($\beta = 0.10$).

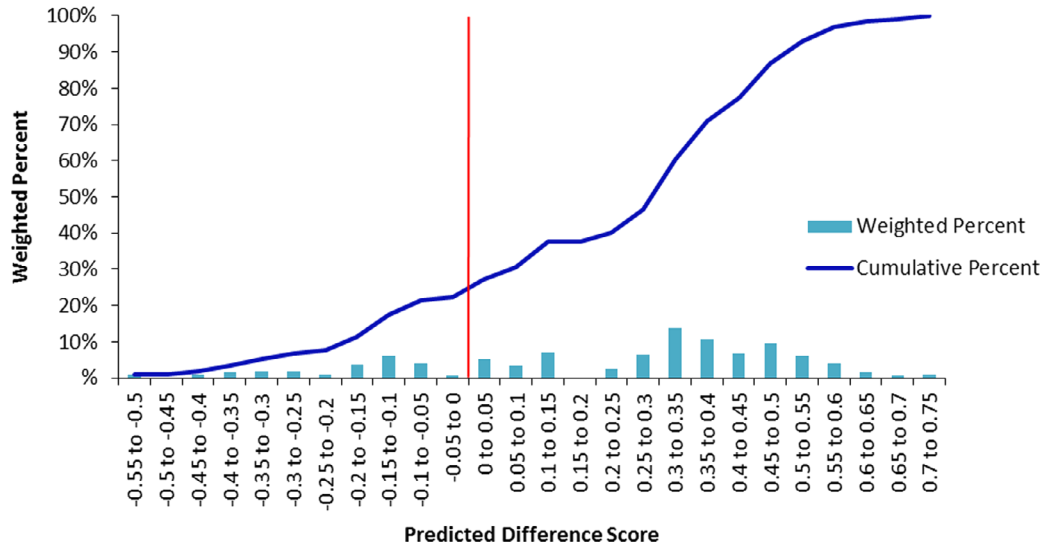

Figure 1. Weighted Distribution of Predicted Difference Score. [Color figure can be viewed at wileyonlinelibrary.com]

DISCUSSION

Our analysis was limited by the fact that OWL was not designed to develop a PTR. This means that the investigators made no attempt to include a comprehensive set of hypothesized prescriptive predictors of remission in their baseline assessment and that the trial was powered to detect a main effect, not the effect of a PTR. We have shown elsewhere that adequate statistical power to develop a PTR for interventions with relatively complex multivariate effects of prescriptive predictors on common mental disorders requires a sample of 300–500 patients per treatment arm (Luedtke, Sadikova, & Kessler, 2019). The fact that the OWL sample was much smaller than this explains why the confidence interval around our estimate of the incremental benefit of our PTR is wide. The point estimate of the incremental benefit might also have been biased downward by the fact that we lacked the richer set of baseline predictors we would have included in a trial designed specifically to develop a PTR.

Based on the above limitations, the PTR developed here needs to be considered no more than preliminary. It is beyond the scope of the current report to discuss whether CAMS would be considered a valuable intervention to eliminate SI given that a PTR exists that could pinpoint a subset of patients for whom CAMS would increase probability of SI remission by 35% beyond E-CAU. It is likely, though, that such an evaluation would require a deeper understanding of the extent to which CAMS is also associated with change in conditional risk of suicidal behavior separately in the presence and absence of SI remission as well as the extent to which CAMS yields outcomes different from those of other suicide-focused interventions separately among patients with and without a history of prior SA. In the ideal case, larger trials with more diverse samples would be carried out that developed PTRs for CAMS in addition to a range of other suicide-focused interventions as well as usual care to determine which of these interventions were optimal for which patients in leading to SI remission and conditional reductions in suicide-related behaviors. It might well be that CAMS is optimal for the largest proportion of patients with a history of SI but not SA, whereas other SRB-focused interventions are optimal among patients with a history of prior SAs. Given the focus of both CAMS and recently approved intranasal

TABLE 3
Baseline Predictors of the Individual-level Differences in Probability of Remission that Define the PTR Selected by an Elastic Net Penalized Regression Model with the Mixing Parameter Set to 0.9 Along with Ordinary Least Squares Regression Coefficients Using the Same Predictors

	Elastic net	Ordinary least squares		
	β	β	*SE*	*p*
I. Sociodemographics				
White race (compared to black)	0.03	0.03	0.06	.64
Other race (compared to black)	−0.36	−0.39	0.06	<.0001
Married (compared to either never or previously married)	−0.02	−0.01	0.06	.82
Born in the U.S.	0.00	0.06	0.05	.21
II. Military status and service history				
3+ Lifetime deployments (compared to 0–2)	−0.07	−0.14	0.05	.008
III. Suicidality				
Age at first self-inflicted injury[a]	−0.02	−0.11	0.05	.05
History of multiple suicide attempts (compared to 0–1)[b]	0.00	0.04	0.06	.45
IV. Psychopathology				
Lifetime major depressive disorder	−0.03	−0.10	0.05	.07
Lifetime anxiety or PTSD	−0.18	−0.20	0.06	.0005
Borderline personality disorder	0.00	−0.11	0.05	.043
Square root of baseline low mood (NSI)	−0.07	−0.06	0.06	.34
V. Other predictors				
Square root of SF-36 physical functioning score	0.33	0.39	0.06	<.0001
Square root of SF-36 social functioning score	0.03	0.07	0.06	.25
Client expectancy of treatment prior to 1st session	0.00	−0.10	0.05	.07

[a]Self-inflicted injuries include either nonsuicidal self-injury or suicide attempt.
[b]Although, as shown in Table 2, history of suicide attempts was initially trichotomized (i.e., none, one, multiple), the significant predictor of individual-level differences in probability of remission was history of multiple suicide attempts compared to 0-1.

esketamine on rapid reduction in SI (Canuso et al., 2018), it would be of special interest to compare the separate and joint effects of these two treatments.

When interpreting the significant predictors of our preliminary CAMS PTR, it is important to remember that these predictors were selected with penalized regression. This means that we need to think of the predictors as standing in for the other predictors with which they were exogenously correlated. The association we found between 0 and 2 deployments and CAMS being preferable to E-CAU, for example, might reflect broader associations with related covariates such as short time in service, young age, and *low* rank. The associations of good physical health and absence of lifetime anxiety and depression with optimality of CAMS, in comparison, might reflect broader associations with low overall disorder persistence, severity, and comorbidity, possibly linked to the fact that CAMS is especially helpful relative to E-CAU among patients with early-onset self-injury and borderline personality disorder. Given the size of the OWL sample, we made no attempt to search for such clusters or interactions in the multivariate data, but analyses of this sort should be central to any future efforts to expand on this work. Such replications would be needed if a more stable PTR is to be developed to provide clinical decision support to treatment providers in deciding when CAMS might be more useful than usual care for particular patients.

Even though the PTR developed here is only provisional, the results are useful in suggesting that substantial variation exists in

the relative effectiveness of CAMS and E-CAU. As discussed in more detail elsewhere (Kessler, 2019; Kessler, Bernecker et al., 2019; Kessler, Bossarte et al., 2019), such heterogeneity of treatment effects (HTE) is likely to exist in many interventions focused on prevention of suicide-related ideation and behaviors (SRIBs). This is true because SRIBs occur in conjunction with manifold mental disorders and because SRIB-focused interventions target intermediate outcomes that might be issues for only a subset of patients, leading inevitably to weak aggregate effects even if large effects exist among the subset of patients who are helped by the intervention. Because of this likely presence of HTE, it is important not to reject SRIB-focused interventions prematurely based on their generally weak aggregate effects. For example, if HTE existed in the "caring text" RCT recently reported by Comtois et al. (2019), then the relatively modest aggregate effects dismissed in an editorial of that study as *predominantly null* (Hoge, 2019) might have masked a (perhaps substantially) stronger effect in the (possibly small) subset of patients helped by the intervention, which might make that intervention a valuable component in a multimodal SRIB prevention strategy. However, it is unknown whether HTE existed because the trial was not designed to study this possibility (either in terms of targeted sampling or comprehensive assessment of hypothesized baseline modifiers).

The same argument could be made for a number of other SRIB-focused trials. For example, CAMS is only one of several outpatient therapies that have been shown in replicated RCTs to reduce SRIBs (Jobes, Au, & Siegelman, 2015), although with concerns about publication bias for some of them (Tarrier, Taylor, & Gooding, 2008). But these trials typically had modest aggregate effect sizes that might reflect dampening due to HTE. Yet none of these trials was designed to investigate which patients benefit from these therapies or which specific SRIB-focused therapies might be optimal for which patients. Clinicians consequently must rely on trial-and-error in deciding whether to use these treatments, and suicidal patients often must suffer through multiple courses of ineffective treatment before they find one that works. A substantial proportion of suicidal patients give up before an effective treatment is found, sometimes with tragic results. The development of comprehensive PTRs using the precision medicine analysis methods described in the current report could be of enormous value in addressing this pressing problem and potentially help bring significant clinical relief to millions of suicidal patients.

REFERENCES

ANDREASSON, K., KROUGH, J., WENNEBERG, C., JESSEN, H. K., KRAKAUER, K., GLUUD, C., ET AL. (2016). Effectiveness of dialectical behavior therapy versus collaborative assessment and management of suicidality treatment for reduction of self-harm in adults with borderline personality traits and disordered—A randomized observer-blind clinical trial. *Depression and Anxiety*, *33*, 520–530.

BECK, A., BROWN, G., & STEER, R. (1997). Psychometric characteristics of the scale for suicide ideation with psychiatric outpatients. *Behaviour Research and Therapy*, *35*, 1039–1046.

BECK, A. T., KOVACS, M., & WEISSMAN, A. (1979). Assessment of suicidal intention: The scale for suicide ideation. *Journal of Consulting and Clinical Psychology*, *47*, 343–352.

BROWN, G. K., TEN HAVE, T., HENRIQUES, G. R., XIE, S. X., HOLLANDER, J. E., & BECK, A. T. (2005). Cognitive therapy for the prevention of suicide attempts: A randomized controlled trial. *Journal of the American Medical Association*, *294*, 563–570. https://10.1001/jama.294.5.563.

BRYAN, C. J., MINTZ, J., CLEMANS, T. A., LEESON, B., BURCH, T. S., WILLIAMS, S. R., ET AL. (2017). Effect of crisis response planning vs. contracts for safety on suicide risk in U.S. Army Soldiers: A randomized clinical trial. *Journal of Affective Disorders*, *212*, 64–72. https://10.1016/j.jad.2017.01.028.

CANUSO, C. M., SINGH, J. B., FEDGCHIN, M., ALPHS, L., LANE, R., LIM, P., ET AL. (2018). Efficacy and safety of intranasal esketamine for the rapid

reduction of symptoms of depression and suicidality in patients at imminent risk for suicide: results of a double-blind, randomized, placebo-controlled study. *American Journal of Psychiatry*, *175*(7), 620–630. https://doi.org/10.1176/appi.ajp.2018.17060720.

CHIPMAN, H. A., GEORGE, E. I., & MCCULLOCH, R. E. (2010). BART: Bayesian additive regression trees. *The Annals of Applied Statistics*, *4*, 266–298.

COHEN, Z. D., & DERUBEIS, R. J. (2018). Treatment selection in depression. *Annual Review of Clinical Psychology*, *14*, 209–236.

COMTOIS, K. A., JOBES, D. A., O'CONNOR, S., ATKINS, D. C., JANIS, K., CHESSEN, C., ET AL. (2011). Collaborative Assessment and Management of Suicidality (CAMS): Feasibility trial for next-day appointment services. *Depression and Anxiety*, *28*, 963–972.

COMTOIS, K. A., KERBRAT, A. H., DECOU, C. R., ATKINS, D. C., MAJERES, J. J., BAKER, J. C., ET AL. (2019). Effect of augmenting standard care for military personnel with brief caring text messages for suicide prevention: A randomized clinical trial [published online]. *JAMA Psychiatry*, *76*(5), 474. https://doi.org/10.1001/jamapsychiatry.2018.4530.

CONNOR, K. M., & DAVIDSON, J. R. (2003). Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety*, *18*, 76–82. https://doi.org/10.1002/da.10113.

CORONA, C. D., GUTIERREZ, P. M., WAGNER, B. M., & JOBES, D. A. (2018). The psychometric properties of the Collaborative Assessment and Management of Suicidality rating scale. *Journal of Clinical Psychology*, *75*(1), 190–201. https://doi.org/10.1002/jclp.22699.

DECOU, C. R., COMTOIS, K. A., & LANDES, S. J. (2019). Dialectical behavior therapy is effective for the treatment of suicidal behavior: A meta-analysis. *Behavior Therapy*, *50*, 60–72.

DRAPEAU, C. W., & MCINTOSH, J. L. (2018). *U.S.A. Suicide 2017: Official final data*. Washington, DC: American Association of Suicidology. Retrieved December 10, 2018, from https://www.suicidology.org.

ELLIS, T. E., RUFINO, K. A., & ALLEN, J. G. (2017). A controlled comparison trial of the Collaborative Assessment and Management of Suicidality (CAMS) in an inpatient setting: Outcomes at discharge and six months follow up. *Psychiatry Research*, *249*, 252–260.

ELLIS, T. E., RUFINO, K. A., ALLEN, J. G., FOWLER, J. C., & JOBES, D. A. (2015). Impact of a suicide specific intervention within inpatient psychiatric care: The Collaborative Assessment and Management of Suicidality. *Suicide and Life-Threatening Behavior*, *45*, 556–566.

FERNANDES, B. S., WILLIAMS, L. M., STEINER, J., LEBOYER, M., CARVALHO, A. F., & BERK, M. (2017). The new field of 'precision psychiatry'. *BMC Medicine*, *15*(1), 80.

FIRST, M. B. (1997a). *Structured clinical interview for DSM-IV axis I disorders SCID-I*. Washington, DC: American Psychiatric Press.

First, M. B. (Ed.) (1997b). *User's guide for the structured clinical interview for DSM-IV axis II personality disorders SCID-II*. Washington, DC: American Psychiatric Press.

FRIEDMAN, J., HASTIE, T., & TIBSHIRANI, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*, 1–22.

GYSIN-MAILLART, A., SCHWAB, S., SORAVIA, L., MEGERT, M., & MICHEL, K. (2016). A novel brief therapy for patients who attempt suicide: A 24-months follow-up randomized controlled study of the Attempted Suicide Short Intervention Program (ASSIP). *PLoS Med*, *13*, e1001968. https://doi.org/10.1371/journal.pmed.1001968.

HOGE, C. W. (2019). Suicide reduction and research efforts in service members and veterans – sobering realities [published online]. *JAMA Psychiatry*, *76*(5), 464. https://doi.org/10.1001/jamapsychiatry.2018.4564.

HUH, D., JOBES, D. A., COMTOIS, K. A., KERBRAT, A. H., CHALKER, S. A., GUTIERREZ, P. M., ET AL. (2018). The Collaborative Assessment and Management of Suicidality (CAMS) versus enhanced care as usual (E-CAU) with suicidal soldiers: Moderator analyses from a randomized controlled trial. *Military Psychology*, *30*, 495–506. https://doi.org/10.1080/08995605.2018.1503001.

JOBES, D. A. (1995). The challenge and the promise of clinical suicidology. *Suicide and Life-Threatening Behavior*, *25*, 437–449.

JOBES, D. A. (2016). *Managing suicidal risk: A collaborative approach* (2nd ed.). New York, NY: Guilford Press.

JOBES, D. A., AU, J., & SIEGELMAN, A. (2015). Psychological approaches to suicide treatment and prevention. *Current Treatment Options in Psychiatry*, *2*, 363–370. https://doi.org/10.1007/s40501-015-0064-3.

JOBES, D. A., COMTOIS, K. A., BRENNER, L. A., GUTIERREZ, P. M., & O'CONNOR, S. S. (2016). Lessons learned from clinical trials of the Collaborative Assessment and Management of Suicidality (CAMS). In R. C. O'Connor, & J. Pirkis (Eds.), *International handbook of suicide prevention* (2nd ed., pp. 431–439). Chichester, UK: Wiley-Blackwell.

JOBES, D. A., COMTOIS, K. A., GUTIERREZ, P. M., BRENNER, L. A., HUH, D., CHALKER, S. A., ET AL. (2017). A randomized controlled trial of the Collaborative Assessment and Management of Suicidality versus enhanced care as usual with suicidal

soldiers. *Psychiatry*, *80*, 339–356. https://doi.org/10.1080/00332747.2017.1354607.

JOBES, D. A., GREGORIAN, M. J., & COLBORN, V. A. (2018). A stepped care approach to clinical suicide prevention. *Psychological Services*, *15*, 243–250.

JOBES, D. A., & JOINER, T. E. (2019). Reflections on suicidal ideation. *Crisis*, *40*, 227–230. https://doi.org/10.1027/0227-5910/a000615.

JOBES, D. A., WONG, S. A., CONRAD, A. K., DROZD, J. F., & NEAL-WALDEN, T. (2005). The Collaborative Assessment and Management of Suicidality versus treatment as usual: A retrospective study with suicidal outpatients. *Suicide and Life-Threatening Behavior*, *35*, 483–497. https://doi.org/10.1521/suli.2005.35.5.483.

KAPLAN, D. (2014). *Neurobehavioral Symptom Inventeory (NSI): Recommendations for Scoring and Serial Administration for Concussion Health Care Outcomes Standardization (MCMR-DCV)*. Silver Spring, MD. Retrieved February 15, 2019, from https://dvbic.dcoe.mil/sites/default/files/DVBIC_-_NSI_Information_Paper_Final.pdf

KESSLER, R. C. (2019). Clinical epidemiological research on suicide-related behaviors: Where we are and where we need to go. *JAMA Psychiatry*, *76*(8), 777.

KESSLER, R. C., BERNECKER, S. L., BOSSARTE, R. M., LUEDTKE, A. R., MCCARTHY, J. F., NOCK, M. K., ET AL. (2019). The role of big data analytics in predicting suicide. In I. C. Passos, B. Mwangi, & F. Kapczinski (Eds.), *Personalized psychiatry: Big data analytics in mental health* (pp. 77–98). Cham, Switzerland: Springer Nature Switzerland AG.

KESSLER, R. C., BOSSARTE, R. M., LUEDTKE, A., ZASLAVSKY, A. M., & ZUBIZARRETA, J. R. (2019). Machine learning methods for developing precision treatment rules with observational data. *Behaviour Research and Therapy*, *120*, 103412.

KLONSKY, D. E., MAY, A. M., & SAFFER, B. Y. (2016). Suicide, suicide attempts, and suicidal ideation. *Annual Review of Clinical Psychology*, *12*, 307–330.

KOOPERBERG, C. (2015). *polspline: Polynomial Spline Routines*. R package version 1.1.12. Retrieved March 2, 2019, from http://CRAN.R-project.org/package=polspline.

LAMBERT, M. J., BURLINGAME, G. M., UMPHRESS, V., HANSEN, N. B., VERMEERSCH, D. A., CLOUSE, G. C., ET AL. (1996). The reliability and validity of the outcome questionnaire. *Clinical Psychology and Psychotherapy*, *3*, 249–258.

LEDELL, E., VAN DER LAAN, M. J., & PETERSEN, M. (2016). AUC-maximizing ensembles through metalearning. *The International Journal of Biostatistics*, *12*, 203–218. https://doi.org/10.1515/ijb-2015-0035.

LI, F., MORGAN, K., & ZASLAVSKY, A. (2017). Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, *113*(521), 390–400. https://doi.org/10.1080/01621459.2016.1260466.

LIAW, A., & WIENER, M. (2002). Classification and regression by randomForest. *R News*, *2*, 18–22.

LINEHAN, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York, NY: Guilford Press Inc.

LINEHAN, M. M., & COMTOIS, K. A. (1996). *Lifetime parasuicide history*. Seattle, WA: University of Washington. Retrieved January 12, 2019, from http://depts.washington.edu/uwbrtc/wp-content/uploads/L-SASI-Count.pdf.

LINEHAN, M. M., COMTOIS, K. A., BROWN, M. Z., HEARD, H. L., & WAGNER, A. (2006). Suicide Attempt Self-Injury Interview (SASII): Development, reliability, and validity of a scale to assess suicide attempts and intentional self-injury. *Psychological Assessment*, *18*, 303–312. https://doi.org/10.1037/1040-3590.18.3.303.

LUEDTKE, A. (2018). *sg*. Retrieved February 18, 2019, from https://github.com/alexluedtke12/sg.

LUEDTKE, A., SADIKOVA, E., & KESSLER, R. C. (2019). Sample size requirements for multivariate models to predict between-patient differences in best treatments of major depressive disorder. *Clinical Psychological Science*, 7, 445–461. https://doi.org/10.1177/2167702618815466.

LUEDTKE, A. R., & VAN DER LAAN, M. J. (2017). Evaluating the impact of treating the optimal subgroup. *Statistical Methods in Medical Research*, *26*, 1630–1640. https://doi.org/10.1177/0962280217708664.

MANSOURNIA, M. A., & ALTMAN, D. G. (2016). Inverse probability weighting. *BMJ*, *352*, i189. https://doi.org/10.1136/bmj.i189.

MCHORNEY, C. A., WARE, J. E., LU, J. F., & SHERBOURNE, C. D. (1994). The MOS 36-Item Short-Form Health Survey (SF-36®): III. Tests of data quality, scaling assumptions and reliability across diverse patient groups. *Medical Care*, *32*, 40–66.

METERKO, L. M., BAKER, M. E., STOLZMANN, D. K., HENDRICKS, L. A., CICERONE, L. K., & LEW, L. H. (2012). Psychometric assessment of the Neurobehavioral Symptom Inventory-22: The structure of persistent postconcussive symptoms following deployment-related mild traumatic brain injury among veterans. *Journal of Head Trauma Rehabilitation*, *27*, 55–62. https://doi.org/10.1097/HTR.0b013e318230fb17.

MILBORROW, S., HASTIE, T., TIBSHIRANI, R., MILLER, A., & LUMLEY, T. (2016). *earth: Multivariate adaptive regression splines*. R package version

4.4.5. Retrieved February 23, 2019, from http://CRAN.R-project.org/package=earth.

NELDER, J., & WEDDERBURN, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)*, *135*(3), 370–384. https://10.2307/2344614.

POLLEY, E., LEDELL, E., KENNEDY, C., LENDLE, S., & VAN DER LAAN, M. (2018). *SuperLearner: Super Learner Prediction version 2.0-24*. Retrieved January 29, 2019, from https://cran.r-project.org/web/packages/SuperLearner/index.html.

RIPLEY, B., & VENABLES, W. (2016). *nnet: Feed-forward neural networks and multinominal log-linear models*. R package version 7.3-12. Retrieved March 3, 2019, from https://cran.r-project.org/web/packages/nnet/index.html.

RUDD, M. D., BRYAN, C. J., WERTENBERGER, E. G., PETERSON, A. L., YOUNG-MCCAUGHAN, S., MINTZ, J., ET AL. (2015). Brief cognitive-behavioral therapy effects on post-treatment suicide attempts in a military sample: Results of a randomized clinical trial with 2-year follow-up. *American Journal of Psychiatry*, *172*, 441–449. https://10.1176/appi.ajp.2014.14070843.

RYBERG, W., ZAHL, P.-H., DIEP, D. M., LANDRO, N. I., & FOSSE, R. (2019). Managing suicidality within specialized care: A randomized controlled trial. *Journal of Affective Disorders*, *249*, 112–120.

SAMHSA (2017). *Suicidal thoughts and behaviors among adults*. NSDUH Annual National Report. Retrieved February 5, 2019, from https://www.samhsa.gov/data/report/2017-nsduh-annual-national-report.

SCHÖLKOPF, B., PLATT, J., SHAWE-TAYLOR, J., SMOLA, A. J., & WILLIAMSON, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, *13*, 1443–1471. https://doi.org/10.1162/089976601750264965.

STANLEY, B., BROWN, G. K., BRENNER, L. A.,GALFALVY, H. C., CURRIER, G. W., KNOX, K. L., ET AL. (2018). Comparison of safety planning intervention with follow-up vs. usual care of suicidal patients treated in the emergency department. *JAMA Psychiatry*, *75*, 894–900. https://doi.org/10.1001/jamapsychiatry.2018.1776.

TARRIER, N., TAYLOR, K., & GOODING, P. (2008). Cognitive-behavioral interventions to reduce suicide behavior: a systematic review and meta-analysis. *Behavior Modification*, *32*, 77–108. https://doi.org/10.1177/0145445507304728.

VAN DEN OORD, E. J. C. G., CHAN, R. F., & ABERVAN DER LAANG, K. A. (2018). Successes and challenges in precision medicine in psychiatry. *JAMA Psychiatry*, *75*(12), 1269–1270.

VAN DER LAAN, M. J., & LUEDTKE, A. R. (2015). Targeted learning of the mean outcome under an optimal dynamic treatment rule. *Journal of Causal Inference*, *3*, 61–95. https://doi.org/10.1515/jci-2013-0022.

VAN DER LAAN, M. J., POLLEY, E. C., & HUBBARD, A. E. (2007). Super learner. *Statistical Applications in Genetics and Molecular Biology*, *6*, Article25. 10.2202/1544-6115.1309.

WARE, J., KOSINSKI, M., BAYLISS, M. S., MCHORNEY, C. A., ROGERS, W. H., & RACZEK, A. (1995). Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: Summary of results from the Medical Outcomes Study. *Medical Care*, *33*, AS264-AS279.

WEATHERS, F. W., HUSKA, J. A., & KEANE, T. M. (1991). *PCL-M for DSM-IV*. Boston, MA: National Center for PTSD – Behavioral Science Division.

Manuscript Received: May 31, 2019
Revision Accepted: October 10, 2019

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. The Complete Set of Baseline Variables (Excluding Square Root Transformations of Continuous Variables) by Randomly Assigned Treatment (n = 124, Weighted Up to N = 148 to Account for Dropout by 3 Months)