



Published in final edited form as:

Depress Anxiety. 2018 November ; 35(11): 1073–1080. doi:10.1002/da.22807.

Pre-deployment predictors of psychiatric disorder-symptoms and interpersonal violence during combat deployment

Anthony J. Rosellini, PhD¹, Murray B. Stein, MD, MPH, FRCPC^{2,3}, David M. Benedek, MD⁴, Paul D. Bliese, PhD⁵, Wai Tat Chiu, AM⁶, Irving Hwang, MA⁶, John Monahan, PhD⁷, Matthew K. Nock, PHD⁸, Nancy A. Sampson, BA⁶, Amy E. Street, PhD^{9,10}, Alan M. Zaslavsky, PhD⁶, Robert J. Ursano, MD⁴, and Ronald C. Kessler, PhD^{6,*}

¹Center for Anxiety and Related Disorders, Boston University, Boston, MA, USA

²Departments of Psychiatry and Family Medicine & Public Health, University of California San Diego, La Jolla, CA, USA

³VA San Diego Healthcare System, San Diego, CA, USA

⁴Center for the Study of Traumatic Stress, Department of Psychiatry, Uniformed Services University School of Medicine, Bethesda, MD, USA

⁵Darla Moore School of Business, University of South Carolina, Columbia, SC, USA

⁶Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

⁷School of Law, University of Virginia, Charlottesville, VA, USA

⁸Department of Psychology, Harvard University, Cambridge, MA, USA

⁹National Center for PTSD, VA Boston Healthcare System, Boston, MA, USA

¹⁰Department of Psychiatry, Boston University School of Medicine, Boston, MA, USA

Abstract

Background—Preventing suicides, mental disorders, and non-combat related interpersonal violence during deployment are priorities of the U.S. Army. We used pre-deployment survey and administrative data to develop actuarial models to identify soldiers at high risk of these outcomes during combat deployment.

Methods—The models were developed in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS) Pre-Post Deployment Study, a panel study of soldiers deployed to Afghanistan in 2012–2013. Soldiers completed self-administered questionnaires before deployment and one (T1), three (T2), and nine months (T3) after deployment, and consented to

*Corresponding author: Ronald C. Kessler, PhD, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA USA 02115; 617-432-3587 (voice); 617-432-3588 (fax); Kessler@hcp.med.harvard.edu.

Conflict of interest disclosures: Dr. Stein has been a consultant for Care Management Technologies, received payment for his editorial work from UpToDate and *Depression and Anxiety*, and had research support for pharmacological imaging studies from Janssen. Dr. Monahan is a co-owner of the Classification of Violence Risk (COVR), Inc. In the past 3 years, Dr. Kessler received support for his epidemiological studies from Sanofi Aventis; was a consultant for Johnson & Johnson Wellness and Prevention, Shire, Takeda; and served on an advisory board for the Johnson & Johnson Services Inc. Lake Nona Life Project. Kessler is a co-owner of DataStat, Inc., a market research firm that carries out healthcare research. The remaining authors report nothing to disclose.

administrative data linkage. Seven during-deployment outcomes were operationalized using the post-deployment surveys. Two overlapping samples were used because some outcomes were assessed at T1 (n=7,048) and others at T2–T3 (n=7,081). Ensemble machine learning was used to develop a model for each outcome from 273 pre-deployment predictors, which were compared to simple logistic regression models.

Results—The relative improvement in area under the receiver operating characteristic curve (AUC) obtained by machine learning compared to the logistic models ranged from 1.11 (major depression) to 1.83 (suicidality). The best-performing machine learning models were for major depression (AUC=0.88), suicidality (0.86), and generalized anxiety disorder (0.85). Roughly 40% of these outcomes occurred among the 5% of soldiers with highest predicted risk.

Conclusions—Actuarial models could be used to identify high risk soldiers either for exclusion from deployment or preventive interventions. However, the ultimate value of our approach depends on the associated costs, competing risks (e.g., stigma), and the effectiveness to-be-determined interventions.

Keywords

Army; military; deployment; predictive modeling; risk assessment; mental disorder; violence

Introduction

Preventing mental disorders, suicide, and non-combat related interpersonal violence (e.g., fights among soldiers) during deployment are priorities of the U.S. Army (Department of the Army, 2015, 2016). The Army uses a two-phase screening approach during the pre-deployment period to detect some risk factors of these outcomes, where soldiers who screen positive on a brief questionnaire (phase one) are further assessed by mental health providers (phase two) to determine if the soldier should be excluded from deployment. An extension of this approach would be to use administrative records and comprehensive risk factor surveys prior to deployment to develop actuarial models to predict risk of mental disorders, suicidality, and interpersonal violence during deployment. Models of this sort have been developed successfully to define small groups of female soldiers at high risk of sexual assault victimization (Street et al., 2016), male soldiers at high risk of physical violence perpetration (Rosellini et al., 2017), and soldiers in treatment who are at high risk of suicide (Kessler et al., 2017; Kessler et al., 2015). If similar models based on pre-deployment data could be developed to predict negative outcomes during deployment, results could be used to target soldiers judged to be high risk for various outcomes for diverse preventive interventions either prior to deployment (e.g., a multi-session cognitive-behavioral program for depression/anxiety, Buntrock et al., 2016; Topper, Emmelkamp, Watkins, & Ehring, 2017; anger management, Shea, Lambert, & Reddy, 2013) or during deployment (e.g., assigning a battle buddy). We report here results of an attempt to build models predicting adverse outcomes during deployment using information obtained from administrative records and a risk factor survey prior to deployment.

Materials and Methods

Sample

Data are from the Pre-Post Deployment Study (PPDS) of the *Army Study to Assess Risk and Resilience in Servicemembers* (Army STARRS) (Ursano et al., 2014). The PPDS was a four-wave panel survey of three Army Brigade Combat Teams (BCTs) assessed 1–2 months before deployment to Afghanistan (average deployment length of 10 months) in 2012 (baseline [T0]) and three times after returning from deployment (within one month after [T1], two months after T1 [T2], and six months after T2 [T3]). Retrospective questions about mental disorders and interpersonal violence experienced during deployment were asked in T1–T3. Participants gave written informed consent to have their de-identified administrative records linked to their de-identified survey responses for purposes of analysis. Recruitment, consent, and data protection procedures were approved by the Institution Review Boards at all collaborating organizations.

9,488 of the 9,949 soldiers in the BCTs (95.3%) consented to participate in the T0 survey, with 8,558 (86.0%) completing that survey and consenting to link their survey responses to their administrative records. 7,742 (90.5%) of the latter respondents subsequently deployed to Afghanistan. Mental disorders during deployment (defined below) were assessed at T2 and T3, whereas interpersonal violence was assessed at T1. Analyses were consequently carried out in two overlapping samples of soldiers who completed (i) T0 and either T2 and/or T3 (n=7,081) and (ii) T0 and T1 (n=7,048). All analyses were weighted to adjust for baseline differences between soldiers who completed versus non-completers at T0, agreed versus did not agree to administrative record linkage, and completed versus non-completers of post-deployment surveys. Additional details on PPDS design, sampling, and weighting are reported elsewhere (Kessler, Colpe, et al., 2013; Kessler, Heeringa, et al., 2013; Rosellini et al., 2015).

Outcomes

The T1 survey was a short paper and pencil self-administered questionnaire (SAQ) administered in a group-administration format in the three BCTs. However, as many respondents were no longer in their original units as of T2, and in some cases were separated from Army service, a two-phase mixed-mode design was used for T2–T3. In Phase 1, emails and text messages were sent to T0 respondents (whether or not they completed the T1 survey) to ask them to complete web-based SAQs. In Phase 2, attempts were made to contact SAQ non-respondents for interviewer-administered telephone surveys.

Mental disorders—The T2–T3 surveys assessed major depressive episode (MDE) and generalized anxiety disorder (GAD) with the Composite International Diagnostic Interview screening scales (CIDI-SC) (Kessler, Calabrese, et al., 2013). An Army STARRS clinical reappraisal study (Kessler, Santiago, et al., 2013) found good concordance of CIDI-SC diagnoses with independent clinical diagnoses based on blinded Structured Clinical Interviews for *DSM-IV* (SCID) (First, Spitzer, Gibbon, & Williams, 2002). Questions were worded to assess symptoms of MDE and GAD during the month of deployment when soldiers “*had the largest number of problems*” with the symptoms. Individuals were also

asked “*how many months*” during the deployment they had these symptoms. MDE and GAD were coded yes if the soldier met criteria and reported having the disorder for at least three months during deployment.

Anger attacks—Anger attacks were assessed using two questions from the CIDI-SC asking soldiers to report the number of attacks: (i) “*when all of a sudden you lost control and either hit, kicked, or tried to hurt someone,*” and (ii) “*when all of a sudden you lost control and either broke or smashed something worth more than a few dollars or threatened someone.*” The outcome for anger attacks was coded yes if the soldier reported having six or more attacks during deployment. This number was set higher than the *DSM-IV* requirement of “several” behavioral anger attacks (i.e., involving physical violence or destruction of property) to define a group of soldiers with clear evidence of persistent-serious anger control problems.

Suicidality—Suicidality was assessed at T2–T3 using a modified version of the Columbia-Suicide Severity Rating Scale (Posner et al., 2011). The outcome for suicidality was coded yes if the soldier reported (i) “*thoughts of killing yourself,*” or “*wishing you were dead or would go to sleep and never wake up*” at least *a little of the time* during the deployment, and (ii) at least *a little* difficulty “*controlling those thoughts or pushing them out of your mind when you wanted to,*” and (iii) if they “*had an intention to act*” on the thoughts about suicide, or if they at least “*sometimes*” “*did dangerous things related to wishing you were dead*” during the deployment.

Head injury—T2–T3 asked about head injuries during deployment. Head injury was coded yes if the soldier reported having “*a head, neck, or blast injury*” that “*knocked you out,*” or “*caused you to be dazed or see stars,*” or “*perforated your eardrum,*” or “*caused a lapse in memory.*”

Violence—Two violence outcomes were operationalized using questions in the T1 SAQ that asked how many times during deployment the soldier experienced: (i) being bullied or hazed (“*you were bullied or hazed by one or more members of your unit*”), and (ii) a fight (“*you got into a fight either with someone in the military or with a civilian*”). These outcomes were coded yes if the soldier reported the event occurring one or more times.

Predictors

Given the focus on targeting preventive interventions prior to deployment, we focused on developing models using only T0 predictors. We considered all potential predictors for which there was evidence in the literature for significant associations with the outcomes. As very few prior studies examined risk factors for mental disorders or interpersonal violence *during* deployment (Bliese, Thomas, McGurk, McBride, & Castro, 2011; Booth-Kewley, Highfill-McRoy, Larson, Garland, & Gaskin, 2012; MHAT-9, 2013; Mulligan et al., 2010), we operationalized a broad range of potential predictors found in both military and civilian literatures on risk-protective factors for mental disorders (Ramchand, Rudavsky, Grant, Tanielian, & Jaycox, 2015; Seal et al., 2009), suicidality (Afifi et al., 2016; Nock et al., 2013), traumatic brain injury (Cassidy et al., 2004; Elmasry, Boivin, Feng, Packnett, &

Cowan, 2017), and violence perpetration-victimization (Dahlberg, 1998; Elbogen et al., 2010). For a complete description of all 273 predictors, see Supplemental Table 1.

Survey variables—The T0 survey was a computerized SAQ administered in group-administration settings in the three BCTs. 158 survey variables or scales were included in the analysis to operationalize potential predictors in six risk/protective factor domains: socio-demographics (e.g., sex, marital status), self-reported mental disorders (e.g., lifetime and 30-day GAD, MDE, PTSD, substance abuse-dependence; standardized scales reflecting severity of symptoms), exposure to stressors (e.g., childhood adversities, lifetime traumatic stressors, past-year stressful life events), personality (e.g., neuroticism, impulsivity, secure attachment), social networks (e.g., number of friends), and suicidality/non-suicidal self-injury (e.g., lifetime and 30-day suicide ideation, plans, attempts, and non-suicidal self-injury).

Administrative variables—115 administrative variables assessed prior to deployment operationalized potential predictors in three risk factor domains: Army career factors (e.g., age of enlistment, rank), prior crime perpetration-victimization (e.g., documented perpetration or victimization of physical or sexual violence), and treatment of mental and physical disorders (outpatient or inpatient treatment for specific mental disorders or broad classes of physical disorders based on ICD-9-CM codes).

Analytic plan

Super learning—Analysis was carried out remotely by Harvard Medical School analysts on the secure University of Michigan Army STARRS Data Coordination Center server. As the goal of the analysis was to maximize prediction accuracy, we considered all 273 potential predictors for each outcome. Super learning (van der Laan, Polley, & Hubbard, 2007) was used to develop an actuarial model for each outcome. Super learning is an ensemble machine learning approach that tries out a range of different classifiers selected by the researcher (e.g., conventional regression analysis, decision trees, support vector machines) to predict an outcome of interest and combines the individual-level predictions based on each algorithm into an optimally weighted average.

We used a super learning library of 14 classifiers each implemented using three different sets of predictors. Based on recent recommendations (LeDell, van der Laan, & Peterson, 2016), and because it was unknown which classifier would result in optimized prediction, we selected a range of classifiers: logistic regression (Hosmer, Lemeshow, & Sturdivant, 2013), six penalized regression classifiers (i.e., elastic net with varying mixing parameter penalties, Friedman, Hastie, & Tibshirani, 2010), two spline regressions (adaptive splines, Friedman, 1991; adaptive polynomial splines, Stone, Hansen, Kooperberg, & Truong, 1997); two decision tree methods (random forests, Breiman, 2001; Bayesian additive regression trees, Chipman George, & McCulloch, 2010), support vector machines (linear kernel, Steinwart & Christmann, 2008), generalized boosting (Freund & Schapire, 1999), and neural nets (Venables & Ripley, 2002). Detailed descriptions of the classifiers can be found in Supplemental Table 2.

The 14 classifiers were implemented using the full set of independent variables as well as two smaller variables sets based on the fact that the use of fewer predictors can improve the performance of some algorithms (e.g., if there is collinearity in the full set of predictors). The two smaller sets of predictors were based on two screening methods built into the R package that implements super learning. The first (screen.corP) identified the subsets of predictors with marginally significant ($p < 0.10$) univariate associations with each outcome. The second (screen.glmnet) identified the non-redundant predictors that would not have their coefficients shrunk to zero in a lasso regression for each outcome.

Model evaluation—Once the final super learner model for each outcome was estimated, externally cross-validated individual-level predicted probabilities were used to calculate area under the receiver operating characteristic curve (AUC). The AUC of the super learner model for the outcome was then compared to the AUC of simple logistic regression models that included a limited number of predictors consisting of the presence or severity of similar disorders or problems prior to deployment. For example, we specified simple models to predict MDE during deployment from T0 30-day MDE and a standardized T0 composite scale of nine 30-day MDE and GAD symptoms assessed for all soldiers. See Table 2 footnote for further descriptions of the simple model predictors. These simple models were implemented via 10-fold cross validation to provide a conservative comparison with super learning, which also used 10-fold cross validation.

We then prioritized for further consideration the subset of outcomes for which super learning outperformed the simple model and the observed proportion of cases was at least 15% among soldiers in the top-risk ventile (i.e., three times the expected value). For outcomes meeting these criteria, the sample was divided into 20 groups of equal size (ventiles) for each outcome and rank-ordered based on predicted probabilities of the outcome. The observed proportion of total cases of each outcome, and prevalence of each outcome, was then calculated in each ventile of predicted risk.

Results

Outcome prevalence and correlations

The most common outcome during deployment was head injury, reported by roughly one-fifth of respondents (21.1%). (Table 1) The least common outcome was suicidality (1.5%). MDE, GAD, and anger attacks were similar to each other in prevalence (5.5–5.7%) and intermediate in prevalence between head injury and suicidality. Several outcomes were moderately-to-strongly inter-correlated (tetrachoric correlations), particularly MDE with GAD (.81), and suicidality with MDE (.58) (see Supplemental Table 3).

Overall accuracy

The relative contributions of each classifier (using each predictor set) to each super learner model is presented in Supplemental Table 4. All super learner models had an AUC above .70, with MDE (0.88), suicidality (0.86), and GAD (0.85) achieving the highest AUCs (Table 2). None of the 13 simple logistic regression models achieved an AUC as high as super learning (simple model AUCs=0.47–0.79). The model using a composite scale of

MDE/GAD symptoms to predict MDE during deployment achieved the highest AUC (0.79) of all the simple models. The relative improvement in AUC obtained via super learning over the simple models (i.e., super learner AUC divided by simple model AUC) ranged from 1.11 (MDE) to 1.83 (suicidality).

Accuracy in identifying true cases

Super learning models for 6 of the 7 outcomes (all except head injury) had proportions of observed cases among soldiers in the top risk ventile above the minimum pre-specified level of 15%. Roughly two-fifths (37.5–44.8%) of all soldiers with MDE, GAD, and suicidality during deployment, one-third of all soldiers with anger attacks during deployment (30.7%), one-fourth of all soldiers who were bullied/hazed during deployment (27.1%), and one-fifth of all soldiers who got into a fight during deployment (19.6%), were among the 5% in the highest predicted risk ventile for those outcomes. Across these six models, this proportion generally remained above the expected value (5%) in the top six predicted risk ventiles. (Figure 1) Prevalence of the outcomes among soldiers in the highest-risk ventile (i.e., positive predictive value) was above 30.0% for all outcomes except suicidality (13.0%).

Discussion

This is the first study to our knowledge to develop actuarial models predicting adverse mental disorder and violence outcomes occurring during combat deployment. Using conventional interpretive guidelines (Hosmer, Lemeshow, & Sturdivant, 2013), super learner AUCs all were either *excellent* (> 0.80) or *acceptable* (> 0.70). In comparison, *poor* AUCs (< 0.70) were achieved by all of the simple logistic models using dichotomous T0 predictors and all but three of the logistic models that used continuous T0 predictors. *Acceptable* AUCs were achieved for MDE, GAD, and suicidality predicted by a composite scale of 30-day pre-deployment MDE/GAD symptoms. For MDE, GAD, and anger attacks, the relative improvement of using all available predictors and ensemble machine learning was nontrivial but modest (e.g., relative improvement < 1.15 ; AUC > 0.10). This suggests that numerous pre-deployment factors contribute to risk for these outcomes during deployment, but that a large proportion of risk may be accounted for by linear associations between the outcome and severity of similar problems prior to deployment.

For all outcomes except head injury, 15.0% of observed cases of the outcomes occurred among the 5% of soldiers with highest predicted super learner risk. For MDE, GAD, and suicidality, roughly two-thirds of all observed cases occurred among soldiers in the top three ventiles of predicted risk. Roughly half of all observed cases of anger attacks and bullying/hazing were in the top three ventiles of predicted risk. In other words, sensitivity would be ~66% for MDE, GAD, and suicidality and ~50% for anger attacks and being bullied/hazed, if the third ventile (top 15%) of predicted super learner risk was used as the cut-point for determining “high risk.”

The two-phase pre-deployment screening approach currently used in the U.S. Army to assess soldiers at risk has been found to be effective in reducing rates of occupationally impairing mental disorder symptoms during deployment (Warner, Appenzeller, Parker, Warner, & Hoge, 2011). The accuracy achieved by the current models argues for their

potential value in improving on this current two-phase screening approach, albeit with the realization that a range of logistical and ethical considerations would have to be addressed if models like those developed here were used as first-stage screens to target soon-to-deploy soldiers for more in-depth clinical assessments. For example, would clinicians be made aware of each soldier's risks, some of which would be based on information obtained from data systems (e.g., criminal justice) not otherwise available to mental health professionals? If so, what would the privacy implications be of doing so? If such information was not made available to clinicians, how should the clinician approach the task of carrying out in-depth risk assessments in the absence of knowing the basis for the soldier being targeted as high-risk? Complex issues of this sort would need to be considered before models of the sort developed here could be used to guide clinical assessments. However, other types of interventions that would not require in-depth clinical assessment (e.g., selectively assigning soldiers to battle buddies based on actuarially-defined risks) could be implemented in ways that circumvented such logistical and ethical considerations.

We are unaware of intensive targeted prevention programs that have been developed specifically for deploying military populations. However, the current models could be used by military researchers aiming to develop and evaluate new preventive interventions for soldiers in combat units. Given the rarity of the outcomes and (likely) small effect of prevention on low-risk soldiers, targeted prevention programs could be developed in conjunction with prediction tools that identify soldiers most likely to need or benefit from prevention. It will not be clear if the performance of the current models is strong enough to justify targeted preventive intervention until these programs have been developed and evaluated in combat units. It would be necessary to weigh the financial costs of a program with the number needed to treat (to prevent one soldier from experiencing the outcome; NNT) prior to using actuarial models for targeted prevention.

Other costs and competing risks also would need to be carefully weighed prior to using actuarial model to identify high-risk soldiers for targeted preventive intervention. The Army currently requires periodic health surveys (Military Health System, 2018), which could be expanded (for deployable units) to include measures used here. However, it could be expensive and logistically challenging to consolidate survey data with records from a number of different administrative data systems. It would also be necessary to consider stigma and privacy concerns associated with being labeled as high-risk or being recommended to participate in targeted preventive intervention. Several strategies exist to reduce stigma associated with mental disorders (Greene-Shortridge, Brit, & Castro, 2007) and could be adapted for targeted prevention (e.g., education about empirically-supported approaches to prevention). Nevertheless, high-risk status would need to be treated as protected health information.

Three study limitations must be acknowledged. First, the "got into a fight" outcome was broad and could have included verbal or physical violence perpetration *or* victimization. Model performance predicting fights might have been better had this outcome been more specific given that risk factors vary depending on the severity and circumstances of violence (Elbogen et al., 2014; Gallaway, Fink, Millikan, & Bell, 2012; Sullivan & Elbogen, 2014). Second, all outcomes were assessed using retrospective reports obtained after returning from

deployment. Responses may have been different had the survey been administered during deployment, although it is noteworthy that the rates of MDE, GAD, and head injury were generally similar to those obtained in anonymous surveys of combat soldiers during deployment (Mental Health Advisory Team 9, 2013). The current models must also be distinguished from others developed to predict *post-deployment* outcomes (Karstoft, Statnikov, Andersen, Madsen, & Galatzer-Levy, 2015), which will be an avenue of future Army STARRS research. Third, soldiers completed the surveys knowing their responses would be confidential. If models like those we developed here were to be used to target preventive interventions in the future, soldiers would have to be made aware of this prior to completing a survey, which might reduce model performance.

Conclusions

Efforts to prevent negative emotional and behavioral outcomes during combat deployment have been limited by a lack of actuarial models to identify high risk soldiers. Our models could be used to assist in making determinations about who should be excluded from deployment. In addition, actuarial models could be used to identify soldiers for who preventive intervention is needed and should be included in research developing and evaluating targeted prevention programs for deploying combat units (i.e., distinct programs for mental disorders/suicide and anger/violence). However, the ultimate value of these models will depend on their costs, effectiveness of the preventive interventions used with high risk soldiers, and competing risks, all of which have to be weighed by Army leaders as they decide whether or not to develop such models for practical use.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The data analyzed in this report were collected as part of the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). Army STARRS was sponsored by the Department of the Army and funded under cooperative agreement number U01MH087981 with the U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Mental Health (NIH/NIMH). This research was conducted by Harvard Medical School and is funded by the Department of Defense, Office of the Assistant Secretary for Defense for Health Affairs, Defense Health Program (OASD/HA), awarded and administered by the U.S. Army Medical Research & Materiel Command (USAMRMC), at Fort Detrick, MD, under Contract Number: (Award # W81XWH-12-2-0113 and W81XWH-13-1-0037). The funders/sponsors were not involved in any aspect of the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, or approval of the manuscript. Although a draft of this manuscript was submitted to the Department of Defense, the Department of the Army, and NIMH for review and comment prior to submission, this was with the understanding that comments would be no more than advisory.

References

- Afifi TO, Taillieu T, Zamorski MA, Turner S, Cheung K, Sareen J. Association of child abuse exposure with suicidal ideation, suicide plans, and suicide attempts in military personnel and the general population in Canada. *JAMA Psychiatry*. 2016; 73(3):229–238. [PubMed: 26817953]
- Bliese PD, Thomas JL, McGurk D, McBride S, Castro CA. Mental health advisory teams: a proactive examination of mental health during combat deployments. *International Review of Psychiatry*. 2011; 23(2):127–134. [PubMed: 21521081]

- Booth-Kewley S, Highfill-McRoy RM, Larson GE, Garland CF, Gaskin TA. Anxiety and depression in Marines sent to war in Iraq and Afghanistan. *Journal of Nervous and Mental Disease*. 2012; 200(9): 749–757. [PubMed: 22922233]
- Breiman L. Random forests. *Machine Learning*. 2001; 45(1):5–32.
- Buntrock C, Ebert DD, Lehr D, Smit F, Riper H, Berking M, Cuijpers P. Effect of a web-based guided self-help intervention for prevention of major depression in adults with subthreshold depression: a randomized clinical trial. *JAMA*. 2016; 315(17):1854–1863. [PubMed: 27139058]
- Cassidy JD, Carroll LJ, Peloso PM, Borg J, von Holst H, Holm L, ... Coronado VG. Incidence, risk factors and prevention of mild traumatic brain injury: results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *Journal of Rehabilitation Medicine*. 2004; 43(Suppl):28–60.
- Chipman HA, George EI, McCulloch RE. BART: Bayesian additive regression trees. *The Annals of Applied Statistics*. 2010; 4(1):266–298.
- Colpe LJ, Naifeh JA, Aliaga PA, Sampson NA, Heeringa SG, Stein MB, ... Kessler RC. Mental health treatment among soldiers with current mental disorders in the Army Study to Assess Risk and Resilience in Service Members (Army STARRS). *Military Medicine*. 2015; 180(10):1041–1051. [PubMed: 26444466]
- Dahlberg LL. Youth violence in the United States. Major trends, risk factors, and prevention approaches. *American Journal of Preventive Medicine*. 1998; 14(4):259–272. [PubMed: 9635070]
- Department of the Army. AR-600-63 Army Health Promotion. 2015. from https://www.army.mil/e2/downloads/rv7/r2/policydocs/r600_63.pdf
- Department of the Army. ATP 6-22.5 A leader's guide to Soldier health and fitness. 2016. from http://www.apd.army.mil/epubs/DR_pubs/DR_a/pdf/web/atp6_22x5.pdf
- Drasgow F, Stark S, Chernyshenko OS, Nye CD, Hulin CL, White LA. Technical Report 1311: development of the Tailored Adaptive Personality Assessment System (TAPAS) to support Army selection and classification decisions. Fort Belvoir, Virginia: United States Army Research Institute for the Behavioral and Social Sciences; 2012.
- Elbogen EB, Cueva M, Wagner HR, Sreenivasan S, Brancu M, Beckham JC, Van Male L. Screening for violence risk in military veterans: predictive validity of a brief clinical tool. *American Journal of Psychiatry*. 2014; 171(7):749–757. [PubMed: 24832765]
- Elbogen EB, Fuller S, Johnson SC, Brooks S, Kinneer P, Calhoun PS, Beckham JC. Improving risk assessment of violence among military veterans: an evidence-based approach for clinical decision-making. *Clinical Psychology Review*. 2010; 30(6):595–607. [PubMed: 20627387]
- Elmasry H, Boivin MR, Feng X, Packnett ER, Cowan DN. Preenlistment and early service risk factors for traumatic brain injury in the Army and Marine Corps: FY 2002–2010. *Journal of Head Trauma Rehabilitation*. 2017; 32(1):1–7. [PubMed: 26709579]
- First M, Spitzer R, Gibbon M, Williams J. Structured Clinical Interview for DSM-IV Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP). New York, NY: Biometrics Research, New York State Psychiatric Institute; 2002.
- Freund Y, Schapire RE. A short introduction to boosting. Paper presented at the Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence; San Francisco, CA. 1999.
- Friedman JH. Multivariate adaptive regression splines. *The Annals of Statistics*. 1991; 19(1):1–141.
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*. 2010; 33(1):1–22. [PubMed: 20808728]
- Galloway MS, Fink DS, Millikan AM, Bell MR. Factors associated with physical aggression among US Army soldiers. *Aggressive Behavior*. 2012; 38(5):357–367. [PubMed: 22898873]
- Greene-Shortridge TM, Brit TW, Castro CA. The stigma of mental health problems in the military. *Military Medicine*. 2007; 172(2):157–161. [PubMed: 17357770]
- Hosmer D, Lemeshow S, Sturdivant R. *Applied Logistic Regression*. 3. Hoboken, NJ: John Wiley & Sons; 2013.
- Karstoft KI, Statnikov A, Andersen SB, Madsen T, Galatzer-Levy IR. Early identification of posttraumatic stress following military deployment: Application of machine learning methods to a prospective study of Danish soldiers. *Journal of Affective Disorders*. 2015; 184:170–175. [PubMed: 26093830]

- Kessler RC, Calabrese JR, Farley PA, Gruber MJ, Jewell MA, Katon W, ... Wittchen HU. Composite International Diagnostic Interview screening scales for DSM-IV anxiety and mood disorders. *Psychological Medicine*. 2013; 43(8):1625–1637. [PubMed: 23075829]
- Kessler RC, Colpe LJ, Fullerton CS, Gebler N, Naifeh JA, Nock MK, ... Heeringa SG. Design of the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *International Journal of Methods in Psychiatric Research*. 2013; 22(4):267–275. [PubMed: 24318217]
- Kessler RC, Heeringa SG, Colpe LJ, Fullerton CS, Gebler N, Hwang I, ... Ursano RJ. Response bias, weighting adjustments, and design effects in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *International Journal of Methods in Psychiatric Research*. 2013; 22(4):288–302. [PubMed: 24318218]
- Kessler RC, Santiago PN, Colpe LJ, Dempsey CL, First MB, Heeringa SG, ... Ursano RJ. Clinical reappraisal of the Composite International Diagnostic Interview Screening Scales (CIDI-SC) in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *International Journal of Methods in Psychiatric Research*. 2013; 22(4):303–321. [PubMed: 24318219]
- Kessler RC, Stein MB, Petukhova MV, Bliese P, Bossarte RM, Bromet EJ, ... Ursano RJ. Predicting suicides after outpatient mental health visits in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *Molecular Psychiatry*. 2017; 22(4):544–551. [PubMed: 27431294]
- Kessler RC, Warner CH, Ivany C, Petukhova MV, Rose S, Bromet EJ, ... Ursano RJ. Predicting suicides after psychiatric hospitalization in US Army soldiers: the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *JAMA Psychiatry*. 2015; 72(1):49–57. [PubMed: 25390793]
- LeDell E, van der Laan MJ, Peterson M. AUC-maximizing ensembles through metalearning. *The International Journal of Biostatistics*. 2016; 12(1):203–218. [PubMed: 27227721]
- Mental Health Advisory Team 9 (MHAT 9). Mental Health Advisory Team 9 (MHAT 9) Operation Enduring Freedom (OEF) 2013: Afghanistan. Office of the Surgeon General U.S. Army Medical Command, Office of the Command Surgeon Headquarters U.S. Army Central Command and Afghanistan; Washington, DC: 2013. from, http://armymedicine.mil/Documents/MHAT_9_OEF_Report.pdf
- Military Health System. Periodic Health Assessment. 2018. Retrieved from <https://health.mil/Military-Health-Topics/Health-Readiness/Reserve-Health-Readiness-Program/Our-Services/PHA>
- Mulligan K, Jones N, Woodhead C, Davies M, Wessely S, Greenberg N. Mental health of UK military personnel while on deployment in Iraq. *British Journal of Psychiatry*. 2010; 197(5):405–410. [PubMed: 21037218]
- Nock MK, Deming CA, Fullerton CS, Gilman SE, Goldenberg M, Kessler RC, ... Ursano RJ. Suicide among soldiers: a review of psychosocial risk and protective factors. *Psychiatry*. 2013; 76(2):97–125. [PubMed: 23631542]
- Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, ... Mann JJ. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *American Journal of Psychiatry*. 2011; 168(12):1266–1277. [PubMed: 22193671]
- Ramchand R, Rudavsky R, Grant S, Tanielian T, Jaycox L. Prevalence of, risk factors for, and consequences of posttraumatic stress disorder and other mental health problems in military populations deployed to Iraq and Afghanistan. *Current Psychiatry Reports*. 2015; 17(5):37. [PubMed: 25876141]
- Rosellini AJ, Stein M, Benedek D, Bliese P, Chiu W, Hwang I, ... Sampson N. Using self-report surveys at the beginning of service to develop multi-outcome risk models for new soldiers in the US Army. *Psychological Medicine*. 2017; 47(13):2275–2287. [PubMed: 28374665]
- Rosellini AJ, Stein MB, Colpe LJ, Heeringa SG, Petukhova MV, Sampson NA, ... Kessler RC. Approximating a DSM-5 diagnosis of PTSD using DSM-IV criteria. *Depression and Anxiety*. 2015; 32(7):493–501. [PubMed: 25845710]
- Seal KH, Metzler TJ, Gima KS, Bertenthal D, Maguen S, Marmar CR. Trends and risk factors for mental health diagnoses among Iraq and Afghanistan veterans using Department of Veterans Affairs health care, 2002–2008. *American Journal of Public Health*. 2009; 99(9):1651–1658. [PubMed: 19608954]

- Shea MT, Lambert J, Reddy MK. A randomized pilot study of anger treatment for Iraq and Afghanistan veterans. *Behaviour Research and Therapy*. 2013; 51(10):607–613. [PubMed: 23916629]
- Steinwart I, Christmann A. *Support Vector Machines*. New York, NY: Springer; 2008.
- Street AE, Rosellini AJ, Ursano RJ, Heeringa SG, Hill ED, Monahan J, ... Kessler RC. Developing a risk model to target high-risk preventive interventions for sexual assault victimization among female U.S. Army Soldiers. *Clinical Psychological Science*. 2016; 4(6):939–956. [PubMed: 28154788]
- Stone CJ, Hansen M, Kooperberg C, Truong YK. Polynomial splines and their tensor products in extended linear modeling (with discussion). *Annals of Statistics*. 1997; 25:1371–1470.
- Sullivan CP, Elbogen EB. PTSD symptoms and family versus stranger violence in Iraq and Afghanistan veterans. *Law and Human Behavior*. 2014; 38(1):1–9. [PubMed: 23646917]
- Topper M, Emmelkamp PM, Watkins E, Ehling T. Prevention of anxiety disorders and depression by targeting excessive worry and rumination in adolescents and young adults: a randomized controlled trial. *Behaviour Research and Therapy*. 2017; 90:123–136. [PubMed: 28049069]
- Ursano RJ, Colpe LJ, Heeringa SG, Kessler RC, Schoenbaum M, Stein MB. The Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *Psychiatry*. 2014; 77(2):107–119. [PubMed: 24865195]
- van der Laan MJ, Polley EC, Hubbard AE. Super learner. *Statistical Applications in Genetics and Molecular Biology*. 2007; 6 Article25.
- Venables W, Ripley Brian D. *Modern applied statistics with S (4th ed., Statistics and computing)*. New York: Springer; 2002.
- Warner CH, Appenzeller GN, Parker JR, Warner CM, Hoge CW. Effectiveness of mental health screening and coordination of in-theater care prior to deployment to Iraq: a cohort study. *American Journal of Psychiatry*. 2011; 168(4):378–385. [PubMed: 21245086]

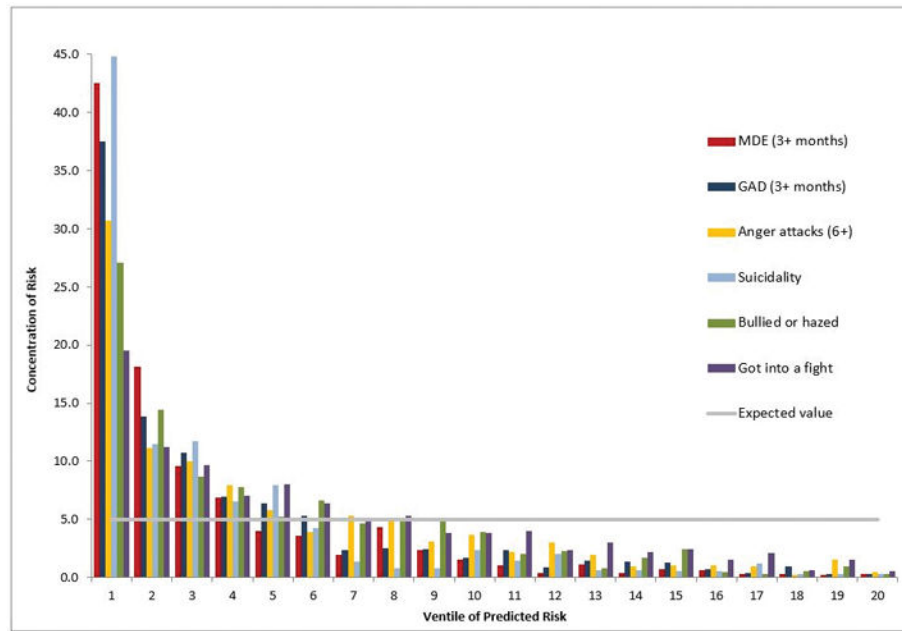


Figure 1. Proportion of outcomes (observed) within each ventile of predicted risk derived from the final super learner models^a

^aVentiles are 20 groups created by dividing the sample into 20 equally sized groups defined by rank order of predicted risk from the final penalized models.

Table 1

Prevalence of adverse during-deployment outcomes in the Pre-Post Deployment Study

	%	SE	(n)
Mental disorders^a			
MDE (3+ months)	5.5	0.3	395
GAD (3+ months)	5.6	0.3	400
Other symptoms^a			
Anger attacks (6+)	5.7	0.3	409
Suicidality	1.5	0.2	107
Head injury	21.1	0.9	1,537
(n) ^a	(7,081)		
Violence^b			
Bullied or hazed by unit	6.4	0.3	494
Got into a fight	11.2	0.5	829
(n) ^b	(7,048)		

Abbreviations: SE, standard error; MDE, major depressive episode; GAD, generalized anxiety disorder;

^aMental disorders and other symptoms occurring during deployment were assessed at T2 and T3. A total of 7,081 soldiers completed T0, T2 and/or T3.

^bViolence occurring during deployment was assessed at T1. A total of 7,048 soldiers completed T0 and T1.

Table 2

Comparison of cross-validated area under the receiver operating characteristic curves, super learner versus simple logistic regressions

	Logistic regression AUC		Super learner AUC
	T0 disorder/problem ^a	T0 symptom severity ^b	
Mental disorders^c			
MDE (3+ months)	0.67	0.79	0.88
GAD (3+ months)	0.64	0.75	0.85
Other symptoms^c			
Anger attacks (6+)	0.59	0.69	0.79
Suicidality	0.47	0.72	0.86
Head injury	--	0.58	0.74
Violence during deployment^d			
Bullied or hazed	0.61	0.56	0.80
Got into a fight	0.54	0.63	0.75

Abbreviations: AUC, area under the receiver operating characteristic curve; MDE, major depressive episode; GAD, generalized anxiety disorder.

^aOutcome was regressed onto a counterpart T0 dichotomous predictor: During deployment MDE was predicted by T0 30-day MDE; GAD was predicted by T0 30-day GAD; Anger attacks was predicted by T0 30-day anger attacks; Suicidality was predicted by T0 30-day suicidality; Bullied/hazed was predicted by T0 report of any (mild, moderate, severe, or very severe) bullying by unit members in the prior 12-months; and Got into a fight was predicted by T0 30-day anger attacks. We were not able to identify a face valid dichotomous T0 disorder/problem to predict head injury during deployment.

^bOutcome was regressed onto T0 symptom severity: Standardized scales were created based on factor analyses of T0 symptoms that were rated using Likert-type scales by all participants (i.e., no skip outs): a 9-item scale composed of five T0 30-day GAD symptoms and four T0 30-day MDE symptoms, a 19-item scale composed of various 30-day health symptoms (e.g., memory problems, sleep problems, pain, headaches), and a 9-item scale assessing the general frequency of anger, irritability, and arguments. During deployment MDE, GAD, and Suicidality were then predicted using the MDE/GAD scale; Anger attacks and Got into a fight were predicted using the irritability scale; Head injury was predicted using the 30-day health symptom scale; and Bullied/Hazed was predicted using all five severity categories (none, mild, moderate, severe, very severe) for bullying by unit members in the prior 12-months.

^cMental disorders and other symptoms occurring during deployment were assessed at T2 and T3. A total of 7,081 soldiers completed T0, T2 and/or T3.

^dViolence occurring during deployment was assessed at T1. A total of 7,048 soldiers completed T0 and T1.

Table 3

Performance of the final super learner for each outcome

	Concentration of risk in ventiles predicted to have highest risk ^a			Outcome rate in ventiles predicted to have highest risk ^b		
	Top 5%	Top 10%	Top 15%	Top 5%	Top 10%	Top 15%
Mental disorders^c						
MDE (3+ months)	42.5	60.6	70.2	46.7	33.2	25.7
GAD (3+ months)	37.5	51.4	62.2	41.6	28.5	23.0
Other symptoms^c						
Anger attacks (6+)	30.7	41.8	51.8	34.9	23.8	19.7
Suicidality	44.8	56.3	68.0	13.0	8.2	6.6
Head injury	13.2	24.0	34.3	55.5	50.7	48.3
Violence during deployment^d						
Bullied or hazed	27.1	41.5	50.2	34.7	26.6	21.5
Got into a fight	19.6	30.8	40.4	44.0	34.6	30.3

Abbreviations: AUC, area under the receiver operating characteristic curve; MDE, major depressive episode; GAD, generalized anxiety disorder.

^aConcentration of risk refers to the percent of all observed occurrences of the outcome accounted for by soldiers in a subset of the predicted risk distribution (i.e., sensitivity among soldiers in the top 5%, 10%, and 15% of predicted risk).

^bPrevalence refers to the rate of the outcome among soldiers in a subset of the predicted risk distribution (i.e., positive predictive value among soldiers in the top 5%, 10%, and 15% of predicted risk).

^cMental disorders and other symptoms occurring during deployment were assessed at T2 and T3. A total of 7,081 soldiers completed T0, T2 and/or T3.

^dViolence occurring during deployment was assessed at T1. A total of 7,048 soldiers completed T0 and T1.