**SUPPLEMENTARY MATERIAL**

***A practical risk calculator for suicidal behavior among transitioning U.S. Army soldiers: Results from the Study to Assess Risk and Resilience in Servicemembers-Longitudinal Study (STARRS-LS)***

**Methodology Appendix**

**Sample**

Data come from the Army Study to Assess Risk and Resilience in Servicemembers-Longitudinal Study (STARRS-LS) surveys. Army STARRS was an epidemiological-neurobiological initiative designed to study risk and protective factors for Army suicidal behaviors (Ursano et al., 2014). The Longitudinal Study (LS) surveys were two panel surveys administered to a probability sample of soldiers and veterans who participated while in active service in one of three different baseline self-report surveys carried out as part of Army STARRS. The analysis focuses only on the subset of such respondents who provided informed consent to link their survey responses to their Army administrative records. As detailed elsewhere (Kessler, Colpe, et al., 2013),calibrationwas used to adjust the respondent sample for discrepancies with nonrespondents across a wide range of administrative variables available for all soldiers.

The baseline Army STARRS surveys, which have been described in detail elsewhere (Heeringa et al., 2013; Kessler, Colpe, et al., 2013; Kessler, Heeringa, et al., 2013; Ursano et al., 2014), were implemented as part of three component studies: the New Soldier Study (NSS), a 2011-2012 cross-sectional survey of n=50,765 soldiers, including members of the Army Reserve and Army National Guard, administered when new soldiers reported for Basic Combat Training; the All Army Study (AAS), a 2011-2013 cross-sectional survey of n=39,666 active soldiers serving throughout the world, including in combat deployments in Afghanistan; and the Pre-Post Deployment Study (PPDS), a 2012-2014 four-wave panel survey among n=9,415 soldiers in 3 Brigade Combat Teams deployed to Afghanistan, with a baseline survey administered before deployment (~2-3 weeks) and subsequent surveys administered between 1 and 9 months after returning from deployment.

 LS recruitment was based on a complex multi-stage nonproportional sampling scheme. The initial LS sampling frame divided baseline Army STARRS survey participants into three strata for differential sampling to maximize statistical power for fixed data collection costs that would not allow follow-up of all baseline respondents. Stratum 1 (n=22,176) included those who either met lifetime criteria for any of the mental disorders assessed in their Army STARRS survey, reported a lifetime history of suicidality in that survey, or were ever in the Special Forces. Stratum 2 (n= 26,833) included those not in Stratum 1 who were either female or in the Army Reserve or Army National Guard. Stratum 3 (n=23,378) included all remaining baseline survey respondents. The first STARRS-LS survey (LS1; September 2016-April 2018) attempted to administer a self-report follow-up survey to 100% of Stratum 1 and probability samples of 67% of Stratum 2 and 50% of Stratum 3. Efforts were made to contact and survey both soldiers still in service and those separated from service with an online self-administered questionnaire or, if preferred by the respondent, a telephone interview.

LS1 recruitment occurred in five phases. In Phase 1, participants were mailed a letter inviting them to participate in the interview. A $50 incentive was offered for participation. The letter included a link to the web interview. Participants were sent a series of emails and text invitations and reminders in Phase 2. Phone calls began in Phase 3 and continued to Phase 4. Participants were again given the option of completing the interview on the web or over the phone. Emails were also sent periodically throughout Phases 3 and 4. In Phase 4, the token of appreciation was increased from $50 to $100 for completing the interview (Supplementary Figure S1). LS1 data were weighted to include the nonresponse adjustment and post-stratification weights developed for the baseline Army STARRS surveys (Nock et al., 2013) along with an additional weight to adjust for the under-representation of difficult-to-recruit participants in LS1. A total of n=14,508 respondents completed LS1, with a weighted (for the under-sampling of difficult-to-recruit participants) response rate of 35.6%. This weighted sample was then post-stratified to adjust for differential response related to survey variables available for all baseline Army STARRS survey respondents and all administrative data available for these baseline respondents as of December 31, 2016. All LS1 respondents were eligible to complete the second STARRS-LS survey (LS2; April 2018 through July 2019). The same field procedures were used in LS2 as in LS1 (Supplementary Figure S2). The LS2 response rate was 83.7% (n=12,156) among LS1 participants. The same post-stratification procedures were used as in LS1 to adjust for nonrandom loss to follow-up.

Results reported here combine data from LS1 and LS2 among respondents who were in the Regular Army at the time of their baseline Army STARRS survey, were in active service for at least 6 months, and reported no longer being in active service at the time of their focal LS survey(s). The term “no longer in active service” is defined as either: being separated from active service in the Army, Navy, Air Force, Marine Corps, Coast Guard, Reserve or National Guard Component (e.g., administrative or medical discharge; fulfilled service obligation; released from obligation; transferred to Individual Ready Reserve or Inactive National Guard Standby Reserve; retired after 20+ years of qualifying service; reached retired pay eligibility age; medical retirement); and in a Reserve or National Guard Component but no longer on orders or activated (i.e., released from active service in the Selected Reserve, Active Guard Reserve, etc.). In addition, we excluded from analysis the LS2 respondents who were already eligible for this analysis at LS1 and who reported a SA in the 12 months before LS1. The latter exclusion was imposed to avoid double-counting any single respondent as having an SA in the pooled LS1-LS2 analysis. This means that, by construction, none of the n=3,110 respondents who were both in the LS1 and the LS2 analysis sample reported a SA in the 12 months before LS1. The full sample included n=8,335 observations, composed of n=3935 at LS1 (n=3,405 separated from active-duty service and n=530 no longer on orders or activated in a Reserve or National Guard Component) and n=4,400 at LS2 (n=3,985 separated and n=615 no longer on orders or activated).

**Measures**

**Suicide attempts:** LS1 and LS2 both included a section on suicidality adopted from the Columbia-Suicide Severity Rating Scale (Posner et al., 2011). One of these questions asked respondents *Did you ever make a suicide attempt (i.e., purposefully hurt yourself with at least some intention to die) at any time since your last survey?* Respondents who said yes were then asked about the number of such attempts and the recency of (that attempt/their most recent attempt). We focus on SAs reported to occur within 12 months of the survey.

**Predictors:** A review of the literature identified 9 categories of predictors of SAs that included socio-demographics, Army career variables, mental disorders, self-injurious thoughts and behaviors, physical health problems, chronic stressors, adverse childhood experiences, other lifetime traumatic events, and personality characteristics (Nock et al., 2013) (Franklin et al., 2017; Holliday et al., 2020; Klonsky, May, & Saffer, 2016). We identified 137 baseline Army STARRS survey measures and 83 administrative variables that could be assessed with survey questions as indicators of these categories (Supplementary Table S1). Given our focus on predictors that could be assessed when respondents were still in active service, the survey measures used to predict SAs reported at LS1 were taken primarily from the 2011-2014 baseline Army STARRS surveys. Those surveys were administered between 4.5 and 7 years before LS1. The 2011-2014 baseline survey data were also used to predict SAs reported at LS2 among respondents who were no longer in active service as of LS1, with a time lag between 5 and 8.5 years between the earlier surveys and LS2. However, in the case of the n=402 LS2 respondents who were still in active service as of LS1 but not within 12 months of LS2, the survey predictors were taken from the LS1 survey, which was administered an average of 18 months before LS2. The influence of time since leaving or being released from active service and time between surveys on strength of model prediction was assessed by carrying out subgroup analyses based on this variable.

**Analysis methods**

Analysis was carried out November 2022 using machine learning (ML) methods to predict suicide attempts in the 12 months before each LS survey. Most studies that use ML to facilitate targeting of preventive interventions among people predicted to be at high risk either use a single algorithm (Jiang et al., 2021) or try several different algorithms and choose the one with the best cross-validated prediction accuracy (e.g., Saxe, Ma, Ren, & Aliferis, 2017). We instead used the Super Learner (SL) ensemble machine learning method. SL uses stacked generalization to pool across multiple algorithms by generating a weight for each algorithm in a user-specified collection (“ensemble”) via 10-fold cross-validation (10F-CV) to create a composite predicted outcome score that is guaranteed in expectation to perform at least as well as the best component algorithm in the ensemble. This optimization guarantee is according to a pre-specified criterion which, in our case, was non-negative least squares (minimizing MSE) (Polley, LeDell, Kennedy, Lendle, & van der Laan, 2015). Consistent with recommendations (LeDell, van der Laan, & Petersen, 2016), we used a diverse set of algorithms in the ensemble to capture nonlinearities and interactions and reduce risk of misspecification (Kennedy CJ, 2017).As detailed in Supplementary Table S2, these included several different linear algorithms (logistic, regularized, spline, and polynomial spline regressions) along with several different tree-based algorithms (individual trees, Bayesian Additive trees, boosting and bagging ensemble trees). Given the small sample size, hyperparameter tuning was achieved by including individual algorithms multiple times in the ensemble with different hyperparameter values and allowing Super Learner to weight relative importance across this range rather than using an external grid search or random search procedure.

Model results were validated by dividing the total sample into a 70% training sample and a 30% test sample, estimating the SL model in the training sample using 10F-CV, and then validating the model in the test sample. We attempted to reduce risk of over-fitting by excluding from the predictor set dichotomous variables with fewer than 10 observed cases of SA in the training sample. We also estimated SL models with the number of predictors restricted to the top 0 through 50 (in groups of 10) based on each feature selection method. This was done to determine the smallest number of predictors needed to reproduce overall model accuracy for future surveys that might be administered to soldiers before they leave or are released from active service. We estimated variable importance for the final model using the model-agnostic kernel SHAP (SHapley Additive exPlanations) method, which estimates the marginal contribution to overall model accuracy of each variable in a predictor set (Lundberg & Lee, 2017). A simple lasso sparse penalized regression model (Tibshirani, 1996) was used as a simple benchmark comparison to the more complex SL models.

 In addition to estimating area under the ROC curve (ROC-AUC) to evaluate model accuracy in the test sample, we divided that sample into 10 risk deciles based on predicted risk scores defined in the best model and calculated both conditional and cumulative *sensitivity* (SN;the proportion of SA within and across deciles of predicted risk) and both conditional and cumulative *positive predictive value* (PPV; prevalence of SA within and across deciles of predicted risk). Multiple feature selection methods were explored to make sure we included as many predictors as needed to optimize model performance but to avoid over-fitting.

These methods were applied initially to the 137 survey measures and 83 administrative measures that could be converted into survey measures assessed in surveys prior to leaving active service. As noted in the body of the text, some of these measures were multi-item scales. For example, the measure of post-traumatic stress disorder (PTSD) was the PTSD Checklist for DSM-4 (Weathers, Litz, Herman, Huska, & Keane, 1993). It is often found that most of the meaningful variance in such scales can be captured with a subset of the items (e.g., Zuromski et al., 2019). To do this, we estimated an initial SL model using all 137 of the survey scales and items and 83 administrative variables we could convert into survey questions. We then estimated predictor importance from that model and retained all scales and single items with nonzero SHAP values for the item-level SL. The subset of scales in this reduced predictor set were then disaggregated to items. And we expanded the set of traumatic events (both those experienced during deployments and others) from those that entered the initial model to all those in the baseline surveys, resulting in a total of 64 variables (57 survey variables plus the 7 administrative variables) for use in an item-level SL model (Supplementary Table S6). These 64 variables were used in estimating a subsequent item-level SL model as well as a lasso model.

We explored the implications of restricting the number of predictors in this set of 64 to create a more practical set of survey questions for the Soldiers for Life (SFL) survey and to reduce over-fitting. This set of restricted models were again estimated in the 70% training sample and tested in the 30% test sample. We found the test sample ROC-AUC was maximized by restricting the number of predictors to a total of 20. However, we also found that the lasso model, which had 17 predictors, performed about as well as the best SL model, leading us to focus further investigation on the lasso model because of its simpler scoring.

We estimated a robust Poisson regression model (Zou, 2004) with the predictors in the lasso model to evaluate predictor importance, standardizing each predictor to a mean of 0 and variance of 1 and transforming the model coefficients and 95% confidence intervals to describe relative risk of SA associated with a one standard deviation change in each of the predictors. It is important to note that the confidence intervals are only heuristic because they are based on a prior lasso model.

**Fairness**

We noted in the body of the paper that concerns have been raised about the issue of fairness in machine learning models based on evidence that the MI models developed for some practical purposes do not predict as accurately for minority classes (e.g., women, who make up only a small minority of all soldiers; or soldiers with nontraditional sexual orientations) as majority classes (Chen et al., 2022). This is sometimes an issue of the researchers not considering minority classes in selecting predictors, in which case the predictors used in the analysis might be more appropriate for majority classes than minority classes. We minimized biases of this sort by building a wide range of predictors into the STARRS surveys and using all available Army administrative data that could realistically be converted into self-report equivalents in building the model.

Another form of bias involves the possibility that ML algorithms that predict optimally for a total population might not predict optimally for minority classes. This issue can be addressed by using algorithms that allow for interactions of minority status with other predictors and building separate models for minority classes. We included algorithms that allowed for interactions in the ensemble used to develop our risk classifier. As we saw, though, models that included such interactions did not out-perform the simple lasso model. It is possible, though, that a larger sample would have yielded evidence for interactions. In addition, in large samples it is possible to develop separate models specifically for minority classes and test whether those models out-perform models which, like our lasso model, assume that the associations of predictors with outcomes are the same for all segments of the population. We plan to explore possibilities of this sort once we are in production, and we have up to 80,000 TSMs complete the survey each year.

Finally, it is important to note the existence of one other type of unfairness that is inherent to prediction models of any sort: the likelihood that sensitivity (SN; the proportion of soldiers who will go on to make a SA who are above the threshold selected for intervention in the risk classifier) will typically not be the same for every segment of the population for a fixed value of positive predictive value (PPV; the probability of SA among soldiers with risk classifier scores above the threshold); or, looking at it the opposite way, the likelihood that PPV will typically not be the same for every segment of the population for a fixed value of SN. Indeed, such subgroup inequivalences are likely unless both prevalence of the outcome and within-subgroup AUC-ROC are the same across subgroups. As a result, the issue comes down to this: Do we want to intervention with everyone who has a SA risk (i.e., PPV) of, say, at least 2% even though we know that that captures only 25% of female TSMs compared to 50% of all male TSMs? Or do we want to use different PPVs for women and men to get the same SN, but thereby requiring men to have higher risk than women to get the intervention? It is important to recognize that we can’t have it both ways. We want our model to provide the best possible discrimination (i.e., highest possible AUC-ROC) for both women and men (and Blacks vs. Whites, etc.). But if prediction strength of the best model is better for some subgroups than others and/or if prevalence differs across subgroups, we will be faced with this dilemma. The obvious way out is to give something extra to disadvantaged groups if one wants to avoid being unfair (e.g., set a lower intervention threshold for women than men), but noting that global optimality goes down when this is done. By the latter, we mean the following: If we intervene with 1,000 people and we use a more favorable threshold for one subgroup than another, the number of people out of the 1,000 who will make a SA otherwise will be lower than if we optimized for the total. Thi is not a technical issue to be addressed by better statistical modeling, but an ethical decision that policy makers need to make once models are develop and thresholds are being set.

|  |
| --- |
| **Supplementary Table S1. The 220 variablesa used to predict suicide attempts in the scale-level super learner modelb** |
|  |
| **I. Socio-demographics** (Dervic et al., 2004; Kposowa, 2000; Lawrence, Oquendo, & Stanley, 2016; Matarazzo et al., 2014; Milner, Hjelmeland, Arensman, & Leo, 2013; Naifeh et al., 2019; Nock et al., 2008; Nock et al., 2014; Phillips & Hempstead, 2017; Ursano et al., 2018) |
| 1. Age: Standardized continuous; Stabilized continuous; 22+ years old; 25+ years old; 28+ years old; 34+ years old
 |
| 1. Gender: Male; Female
 |
| 1. Race/ethnicity: Non-Hispanic white; Non-Hispanic black; Hispanic
 |
| 1. Nativity: Born in US; Number of parents born in US (Standardized continuous; Stabilized continuous; 1 parent born in US; Both parents born in US); Highest education of parents (High school or less, some college, college undergraduate degree or less, college graduate degree; High school or less; Some college or less; College undergraduate degree or less)
 |
| 1. Education: Less than high school or GED; High school or less; Some college or less; Highest education (Less than high school, high school diploma, some college, college or more)
 |
| 1. Relationship status & sexual orientation: Currently married; Engaged; Steadily dating; Dating but not steadily; Not dating; Previously married; Never married; Marriage ended in the year prior to leaving or being released from active service; Married 3+ times; Sexual orientation (Heterosexual/straight, gay/lesbian/bisexual)
 |
| 1. Children: Any dependents in the following age groups: Ages 0-5, 6-13, 0-13, 14-17, 0-17; Has 1+ dependents; Has 2+ dependents; Has 3+ dependents; Stabilized continuous number of dependents; Has 1+ biological child; Has 1+ stepchild; Has 1+ biological or stepchild
 |
| 1. Religiosity: Religiosity scale (Standardized sum of 3 items; Stabilized sum of 3 items; Score of 2+ on 0-11 scale; Score of 3+; Score of 5+; Score of 8+ on 0-11 scale); Fundamentalism index (Standardized sum of 5 items; Stabilized sum of 5 items; Score of 1+ on 0-5 scale; Score of 2+ on 0-5 scale)
 |
|  |
| **II. Army career** (Bryan et al., 2015; Millner et al., 2018; Mitchell, Gallaway, Millikan, & Bell, 2012; Nock et al., 2014; Reger et al., 2015; Ursano et al., 2017; Ursano et al., 2018) |
| 1. Age at Army enlistment: Standardized continuous; Stabilized continuous; 18+ years old; 19+ years old; 21+ years old; 24+ years old
 |
| 1. Rank & promotions: Junior enlisted; Junior or Senior enlisted; Warrant officer; Highest rank (Junior enlisted, Senior enlisted, Warrant officer, Commissioned officer)
 |
| 1. Command: FORSCOM; TRADOC; US Army Service Component Central Command, Northern Command, Southern Command, Europe & Africa Command, and Pacific Command; Special Operations Command; AMC/Other/Unknown; MEDCOM; Guard/Reserve
 |
| 1. MOS: Direct or indirect combat arms; Combat support; Combat service support/other support; Special operator
 |
| 1. Deployments: Currently deployed in support of GWOT prior to leaving or being released from active service; Previously deployed in support of GWOT; Never deployed in support of GWOT; Number of GWOT deployments (Standardized continuous; Stabilized continuous; 1+ deployments; 2+ deployments); Number of months since last GWOT deployment (Standardized continuous; Stabilized continuous; 14+ months; 39+ months)
 |
| 1. Time in service: Total months active/not active in any service component since Army enlistment (Standardized continuous; Stabilized continuous; 37+ months; 55+ months; 70+ months; 124+ months)
 |
| 1. Service characterization: Discharged Honorably or Under Honorable Conditions; Discharged Under Other Than Honorable Conditions, Bad Conduct, or Dishonorable; Unknown reason for discharge; Not discharged
 |
|  |
| **III. Mental disorders** (Bryan, Bryan, Roberge, Leifker, & Rozek, 2018; Cougle, Keough, Riccardi, & Sachs-Ericsson, 2009; Franklin et al., 2017; Holliday et al., 2020; Huang et al., 2018; Naifeh et al., 2019; Nock et al., 2008; Nock et al., 2014; Panagioti, Gooding, & Tarrier, 2009; Ratcliffe, Enns, Belik, & Sareen, 2008; Stanley, Rogers, Hanson, Gutierrez, & Joiner, 2019; Ursano et al., 2018) |
| * 1. Psychiatric comorbidities: Lifetime MDE, GAD, Panic, PTSD, Broad bipolar disorder, ADHD, IED, and Substance use disorder; Childhood conduct & ODD scale (Standardized sum of 12 items; Stabilized sum of 12 items; Score of 2+ on 0-48 scale; Score of 5+; Score of 9+; Score of 13+ on 0-48 scale)
 |
|  |
| **IV. Self-injurious thoughts and behaviors**(Bostwick, Pabbati, Geske, & McKean, 2016; Franklin et al., 2017; Klonsky, May, & Saffer, 2016; Nock et al., 2008; Nock et al., 2013; Ribeiro et al., 2016; Ursano et al., 2015) |
| 1. Suicidal events: Lifetime active/passive suicidal ideation; Active/passive ideation within 2 years of leaving or being released from active service; Lifetime suicidal plan; Lifetime suicide attempt; 1+ lifetime attempts; Attempt within 2 years prior to leaving or being released from active servicev
 |
|  |
| **V. Physical health problems** (Bryan & Clemans, 2013; Daugherty, Waltzman, Sarmiento, & Xu, 2019; Khazem & Anestis, 2019; Nock et al., 2013; Ratcliffe, Enns, Belik, & Sareen, 2008; Ryb et al., 2006; Stanley, Joiner, & Bryan, 2017; Ursano et al., 2018) |
| 1. Physical health problems: Lifetime TBIc; 30-day/6-month pain severityd (Standardized continuous; Stabilized continuous; Score of 2+ on 0-10 scale; Score of 3+; Score of 5+ on 0-10 scale)
 |
| 1. Accidents & disability: Involved in an accident in the 12 months prior to leaving or being released from active service
 |
|  |
| **VI. Chronic stressors** (Milner, Hjelmeland, Arensman, & Leo, 2013; Naifeh et al., 2019) |
| 1. 12-month stress severitye: Severity of stress in life overall (Standardized continuous; Stabilized continuous; At least mild stress; At least moderately severe stress); Stress severity scale (Standardized sum of 5 items; Stabilized sum of 5 items; Score of 1+ on 0-20 scale; Score of 3+; Score of 4+; Score of 7+ on 0-20 scale)
 |
|  |
| **VII. Adverse childhood experiences** (Afifi et al., 2016; Dube et al., 2001; Stein et al., 2018) |
| 1. Abuse & neglect: Family hit so hard it left bruises; Experienced physical abuse at home; Felt hated by family; Family said hurtful or insulting things (Standardized continuous; Stabilized continuous; Verbally hurt/insulted at least rarely; Verbally hurt/insulted at least sometimes); Experienced emotional abuse at home; Touched by someone in a sexual; Experienced sexual abuse at home; Had to do difficult/dangerous chores; Did not have adequate care/protection; Lacked food, clothing, and basic needs
 |
| 1. Other adverse childhood events: Family was on welfare; Family was homeless
 |
|  |
| **Supplementary Table S1 (continued). The 220 variablesa used to predict suicide attempts in the scale-level super learner modelb** |
|  |
| **VIII. Other lifetime traumatic events** (Bryan, Bryan, Roberge, Leifker, & Rozek, 2018; Bryan et al., 2015; Franklin et al., 2017; Naifeh et al., 2019; Nock et al., 2013; Panagioti, Gooding, & Tarrier, 2009; Reger et al., 2017; Stanley, Rogers, Hanson, Gutierrez, & Joiner, 2019; Stewart et al., 2019; Ursano et al., 2018) |
| 1. Stressful experiences during deployment: Went on combat patrols/dangerous duty; Had a close call; Wounded by the enemy; Fired rounds at the enemy/took enemy fire; Responsible for death of enemy combatant; Responsible for death of non-combatant; Responsible for death of US or ally personnel; Saved the life of a servicemember/civilian; Saw destroyed homes/villages or people begging for food; Exposed to wounded/dying people; Witnessed violence in the local population/mistreatment toward non-combatants; Physically assaulted; Sexually assaulted/raped; Hazed or bullied by unit members
 |
| 1. Other lifetime stressful experiences: Experienced a life-threatening illness/injury; Physically assaulted, sexually assaulted, or raped; Exposed to a natural disaster; Other experience that caused risk of injury/death; Close friend/relative accidentally died; Close friend/relative died in combat; Close friend/relative was murdered; Close friend/relative died by suicide; Total number of interpersonal losses (Standardized continuous; Stabilized continuous; 1+ interpersonal losses; 4+ interpersonal losses); Close friend/relative attempted suicide; Close friend/relative was assaulted; Close friend/relative was at risk of injury/death; Witnessed someone being seriously injured/killed; Discovered or handled a dead body
 |
| 1. Wounded Warrior/Warrior Transition Unit: Eligible/enrolled in Wounded Warrior program at any time prior to leaving or being released from active service; Eligible/enrolled in Warrior Transition Unit program at any time prior to leaving or being released from active service
 |
| 1. Crimef: Victim of each of the following crimes in the 12 months and 4 years prior to leaving or being released from active service: Major physical violence, Non-violent criminal offense, Any criminal offense; Perpetrator of any criminal offense in the 12 months and 4 years prior to leaving or being released from active service
 |
| 1. MEPS evaluations: 1+ previous attempts to enlistg
 |
|  |
| **IX. Personality characteristics** (Soloff, Lynch, Kelly, Malone, & Mann, 2000; Yen et al., 2021)  |
| 1. Negative affect scales: Antisocial traits (Standardized sum of 6 items; Stabilized sum of 6 items; Score of 1+ on 0-24 scale; Score of 4+ on 0-24 scale); Borderline traits (Standardized sum of 6 items; Stabilized sum of 6 items; Score of 1+ on 0-24 scale; Score of 3+; Score of 6+; Score of 9+ on 0-24 scale); Anger-irritability (Standardized sum of 9 items; Stabilized sum of 9 items; Score of 1+ on 0-36 scale; Score of 3+; Score of 6+; Score of 10+ on 0-36 scale)
 |
| 1. Resilience: Perceived psychological resilience scale (Standardized sum of 5 items; Stabilized sum of 5 items; Score of 10+ on 0-20 scale; Score of 13+; Score of 15+; Score of 18+ on 0-20 scale)
 |
|  |

Abbreviations: US; United States; GED, General Educational Development Test; FORSCOM, United States Army Forces Command; TRADOC, United States Army Training and Doctrine Command; AMC, United States Army Materiel Command; MEDCOM, United States Army Medical Command; MOS, Military Occupational Specialty; GWOT, Global War on Terror; MDE, major depressive episode; GAD, generalized anxiety disorder; PTSD, post-traumatic stress disorder; ADHD, attention deficit-hyperactivity disorder; IED, intermittent explosive disorder; ODD, oppositional defiant disorder; TBI, traumatic brain injury; MEPS, Military Entrance Processing Station.

aConsisting of 137 survey variables plus 83 administrative variables that we felt could be assessed with self-report survey questions. All variables were defined as of the time period prior to the respondent leaving or being released from active service. Detailed information about scoring and sources is reported elsewhere (Stanley et al., 2022; Zuromski et al., 2020).

bThe full Army Study to Assess Risk & Resilience in Servicemembers (Army STARRS) All Army Study (AAS), Pre-Post Deployment Study (PPDS), New Soldier Study (NSS), and STARRS-Longitudinal Study (LS) survey instruments are available at <https://starrs-ls.org/#/page/instruments>.

cTBI = A head, neck, or blast injury with loss of consciousness and/or posttraumatic amnesia for 30+ minutes.

dSeverity of pain on average in the past 30 days/6-months, rated on a 0-10 numeric rating scale where 0=no pain and 10=pain as bad as could be.

eRespondents were asked to rate severity of stress in 5 areas of life (financial situation/career, personal health, love life, relationships with family/health of loved ones, problems with their unit/being bullied by unit members) and in life overall, where 0=no stress, 1=mild, 2=moderate, 3=severe, and 4 = very severe stress.

fThe National Corrections Reporting Program (NCRP) (United States Department of Justice, Office of Justice Programs, & Bureau of Justice Statistics, 2009) was used to categorize criminal offenses. Major physical violence included murder, homicide, manslaughter, kidnapping, aggravated arson, aggravated assault, family-related aggravated assault, and robbery; Non-violent criminal offenses included reckless endangerment, burglary, simple arson, forgery, fraud, theft, property offenses, commercialized vice (illegal substances and other), possession/use/other violation of illegal substance, court-ordered violations, weapon offense, minor traffic offense, drunkenness, vagrancy, disorderly conduct, military specific non-violent offenses, family-related non-violent offenses, and other non-violent offenses (e.g., bribery, contributing to the delinquency of a minor, immigration violations); Any criminal offense included major physical violence and non-violent offenses along with minor physical violence (simple assault, family-related simple assault, and other physical violence), verbal violence (blackmail, extortion, intimidation, rioting, harassment (non-sexual), family-related emotional abuse, sexual verbal violence (sexual harassment), and sexual violence (rape, sodomy, sexual assault, family-related sexual violence).

gCount of previous MEPS records indicating attempts to enlist before date of enlistment in the Army.

|  |
| --- |
| **Supplementary Table S2. Algorithms used in the Super Learner machine learning analysis** |
|  |  |
| **Algorithm** | **Description** |
| **I. Super Learner** |  |
|  | 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 pre-specified criterion that performs at least as well as the best component algorithm. R package: *Superlearner.* (Polley et al., 2015; Van der Laan, Polley, & Hubbard, 2007) |
|  |  |
| **II. Learners in the Super Learner library** |  |
| A. Elastic Net | Elastic net is a regularization method that minimizes the problem of overlap among predictors by explicitly penalizing over-fitting with a composite penalty λ{MPP x Plasso + (1- MPP) X Pridge}mppmppmppmp, 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 towards 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. R package: *glmnet*. (Friedman, Hastie, & Tibshirani, 2010)Hyperparameters: alpha = (.042,.141,.489,.637,.865,.968),. |
|  |  |
|  B. Decision trees – bagging | Random Forest. Independent variables are partitioned (based on contiguous values) and stacked to build decision trees that are combined (ensemble) to create an aggregate “forest”. Random forest builds numerous trees in bootstrapped samples and generates an aggregate prediction by averaging across trees, thereby reducing over-fitting. R package: *ranger.* (Wright & Ziegler, 2017)Hyperparameters:max.depth = (1, 2, 3, 4, 5, 6), num.trees = (1000), mtry = (3,4,5,10,42,77,145,147), min.node.size = (2, 4, 5,6,11,13,22). |

|  |
| --- |
| **Supplementary Table S2 (continued). Algorithms used in the Super Learner machine learning analysis**  |
|  |  |
| **Algorithm** | **Description** |
| C. Decision trees – boosting  |  |
| Extreme Gradient Boosting  | A fast and efficient implementation of gradient boosting. R package: *xgboost*. (Chen & Guestrin, 2016)Hyperparameters: ntrees = (1000), max\_depth = (1, 2, 3, 4, 7), shrinkage = (.04,.23,.26,.27,.50,.51,.62,.77,.81,.92), gamma = (0.5,0.7,2.9,3.0,3.2,3.3,4.9,5.3,6.1,6.7,9.9), minobspernode = (1,5,8,10,13,16), early\_stopping\_rounds = 50, colsample\_bytree = (.12,.16,.18,.24,.29,.55,.68,.69,.72,.83,.86), colsample\_bynode = (.04,.05,.27,.49,.51,.55,.64,.65,.71,.76,.85,.89). |
|  |  |
|  D. Support Vector Machine | Support vector machines treat independent variables as dimensions in high dimensional space and attempts to identify the best hyperplane to separate the sample into classes (e.g., cases and non-cases). The goal is to find the hyperplane with the maximum margin between the two closest points in space. SVM captures linear associations, but alternate kernels can be used to capture nonlinearities. R package *WeightSVM* (Cortes & Vapnik, 1995)Hyperparameters: kernel = (‘radial’), cost = (1.2,1.4,2.0,2.2,3.2,3.3,4.6,5.0), gamma = (1.0,1.1,1.2,1.7,1.9). |
|  |  |
|  E. Neural networks | Connections between predictors and the outcome are modeled as a network. The predictors affect the outcome through intermediate layers. Weights are assigned to connections. Neural networks capture interactions and non-linear associations but have low interpretability. R package: *nnet.*Hyperparameters: size = (1,2,6), decay = (6.2,21.8,50.5). |
|  |  |

|  |
| --- |
| **Supplementary Table S3. Comparison of socio-demographic and Army career in the LS1 sample versus the population both before and after weighting** |
|  |
|  | **New Soldier Study** |  | **All Army Study/Pre-Post Deployment Study** |
|  |  |  | **LS1** **n=6,331** |  |  |  | **LS1****n=8,173** |
|  | **Populationa N=869,996** | **Unweighted** | **Weighted** |  | **Populationa N=8,108,371** | **Unweighted** | **Weighted** |
|  | **%** | **(SE)** | **%** | **(SE)** | **%** | **(SE)** |  | **%** | **(SE)** | **%** | **(SE)** | **%** | **(SE)** |
| **I. Socio-demographics** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Gender |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Male | 80.7 | (0.0) | 78.3 | (0.7) | 81.9 | (0.8) |  | 85.9 | (0.0) | 88.8 | (0.5) | 85.5 | (0.8) |
| Female | 19.3 | (0.0) | 21.6 | (0.7) | 18.1 | (0.8) |  | 14.1 | (0.0) | 11.2 | (0.5) | 14.5 | (0.8) |
| Age |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 18-22 | 76.2 | (0.0) | 68.6 | (0.7) | 74.6 | (0.7) |  | 18.2 | (0.0) | 15.2 | (0.4) | 16.7 | (0.8) |
| 23-27 | 16.7 | (0.0) | 21.0 | (0.6) | 17.5 | (0.6) |  | 28.3 | (0.0) | 31.3 | (0.6) | 29.1 | (0.8) |
| 28-33 | 6.1 | (0.0) | 9.0 | (0.4) | 6.9 | (0.4) |  | 26.8 | (0.0) | 30.2 | (0.5) | 29.0 | (0.8) |
| 34+ | 1.0 | (0.0) | 1.4 | (0.2) | 0.9 | (0.1) |  | 26.6 | (0.0) | 23.3 | (0.6) | 25.2 | (0.9) |
| Race-ethnicity |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Non-Hispanic White | 60.0 | (0.0) | 66.4 | (0.7) | 63.1 | (1.0) |  | 61.2 | (0.0) | 70.2 | (0.6) | 65.4 | (1.3) |
| Non-Hispanic Black | 20.6 | (0.0) | 16.2 | (0.5) | 18.5 | (0.7) |  | 20.5 | (0.0) | 11.4 | (0.4) | 16.0 | (0.8) |
| Hispanic | 13.2 | (0.0) | 11.5 | (0.5) | 12.5 | (0.7) |  | 11.2 | (0.0) | 11.8 | (0.4) | 11.6 | (0.7) |
| Other | 6.2 | (0.0) | 5.9 | (0.3) | 5.9 | (0.4) |  | 7.1 | (0.0) | 6.6 | (0.3) | 7.1 | (0.7) |
| Highest education |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Less than high school | 16.9 | (0.0) | 14.8 | (0.4) | 16.1 | (0.7) |  | 4.2 | (0.0) | 1.9 | (0.2) | 1.8 | (0.3) |
| Certificate, GED, ARNG | 5.7 | (0.0) | 3.6 | (0.2) | 4.2 | (0.4) |  | 8.5 | (0.0) | 6.9 | (0.3) | 6.7 | (0.6) |
| High school | 69.1 | (0.0) | 66.6 | (0.7) | 69.1 | (0.8) |  | 60.3 | (0.0) | 63.5 | (0.6) | 61.3 | (1.1) |
| Some college | 2.0 | (0.0) | 2.6 | (0.2) | 2.0 | (0.2) |  | 3.8 | (0.0) | 4.4 | (0.2) | 4.6 | (0.3) |
| College graduate | 6.3 | (0.0) | 12.4 | (0.5) | 8.6 | (0.4) |  | 23.0 | (0.0) | 23.3 | (0.7) | 25.6 | (1.3) |
| Marital status |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Never married | 87.8 | (0.0) | 84.8 | (0.4) | 85.2 | (0.5) |  | 32.1 | (0.0) | 30.3 | (0.6) | 33.6 | (1.0) |
| Previously married | 1.1 | (0.0) | 0.2 | (0.0) | 0.1 | (0.0) |  | 6.4 | (0.0) | 4.2 | (0.2) | 3.9 | (0.3) |
| Currently married | 11.0 | (0.0) | 15.1 | (0.4) | 14.7 | (0.5) |  | 61.5 | (0.0) | 65.4 | (0.6) | 62.5 | (1.0) |
| **II. Army careerb** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Rank |  |  |  |  |  |  |  |  |  |  |  |  |  |
| E1-E4 |  |  |  |  |  |  |  | 43.8 | (0.0) | 39.4 | (0.8) | 43.4 | (1.1) |
| E5-E9 |  |  |  |  |  |  |  | 38.1 | (0.0) | 43.6 | (0.6) | 38.3 | (1.1) |
| Warrant officer |  |  |  |  |  |  |  | 2.9 | (0.0) | 2.6 | (0.2) | 3.2 | (0.5) |
| Commissioned officer |  |  |  |  |  |  |  | 15.2 | (0.0) | 14.4 | (0.5) | 15.1 | (1.0) |

|  |
| --- |
| **Supplementary Table S3 (continued). Comparison of socio-demographic and Army career in the LS1 sample versus the population both before and after weighting** |
|  |
|  | **New Soldier Study** |  | **All Army Study/Pre-Post Deployment Study** |
|  |  |  | **LS1** **n=6,331** |  |  |  | **LS1****n=8,173** |
|  | **Populationa N=869,996** | **Unweighted** | **Weighted** |  | **Populationa N=8,108,371** | **Unweighted** | **Weighted** |
|  | **%** | **(SE)** | **%** | **(SE)** | **%** | **(SE)** |  | **%** | **(SE)** | **%** | **(SE)** | **%** | **(SE)** |
| Time in the Army |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1-24 months |  |  |  |  |  |  |  | 16.8 | (0.0) | 9.3 | (0.4) | 16.3 | (0.9) |
| 25-60 months |  |  |  |  |  |  |  | 26.6 | (0.0) | 32.1 | (0.8) | 26.2 | (0.9) |
| 61-120 months |  |  |  |  |  |  |  | 24.3 | (0.0) | 28.4 | (0.5) | 25.2 | (0.9) |
| 121+ months |  |  |  |  |  |  |  | 32.3 | (0.0) | 30.1 | (0.8) | 32.3 | (1.1) |
| Command |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FORSCOM |  |  |  |  |  |  |  | 51.1 | (0.0) | 64.3 | (1.1) | 47.4 | (1.6) |
| TRADOC |  |  |  |  |  |  |  | 9.8 | (0.0) | 3.6 | (0.4) | 10.0 | (1.3) |
| North/South America, Europe/Central/Africa, Pacific |  |  |  |  |  | 15.3 | (0.0) | 8.4 | (0.6) | 12.6 | (1.3) |
| Special Operations |  |  |  |  |  |  |  | 5.2 | (0.0) | 3.8 | (0.7) | 5.0 | (0.8) |
| MEDCOM |  |  |  |  |  |  |  | 6.8 | (0.0) | 4.6 | (0.5) | 7.3 | (0.9) |
| AMC/other/unknown |  |  |  |  |  |  |  | 11.7 | (0.0) | 3.8 | (0.4) | 8.1 | (0.9) |
| Guard/Reserve |  |  |  |  |  |  |  | 0.0 | (0.0) | 11.4 | (0.6) | 9.7 | (0.8) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Abbreviations: LS1, Study to Assess Risk & Resilience in Servicemembers-Longitudinal Study Wave 1; SE, standard error; GED, General Educational Development; ARNG, Army National Guard; FORSCOM, United States Army Forces Command; TRADOC, United States Army Training and Doctrine Command; North/South America, Europe/Central/Africa, Pacific, US Army Service Component Central Command, Northern Command, Southern Command, Europe & Africa Command, and Pacific Command; MEDCOM, United States Army Medical Command; AMC, United States Army Materiel Command.

aThe population for the New Soldier Study (NSS) survey was defined as all new soldiers in the Army over the years the NSS was administered (2011-2012), with records defined for person-months. The population for the All Army Study (AAS) survey was defined as all soldiers in duty units in the Army over the years the AAS was administered (2011-2013), again with records defined for person-months. The population for the Pre-Post Deployment Study (PPDS) survey was defined as all soldiers in Brigade Combat Teams over the year the PPDS baseline survey as administered (2012).

bArmy career characteristics were not examined for the NSS, as all NSS respondents were new soldiers.

|  |
| --- |
| **Supplementary Table S4. Comparison of socio-demographic and Army career in the LS2 sample versus the population both before and after weighting** |
|  |
|  | **New Soldier Study** |  | **All Army Study/Pre-Post Deployment Study** |
|  |  |  | **LS2** **n=5,172** |  |  |  | **LS2** **n=6,984** |
|  | **Populationa N=869,996** | **Unweighted** | **Weighted** |  | **Populationa N=8,108,371** | **Unweighted** | **Weighted** |
|  | **%** | **(SE)** | **%** | **(SE)** | **%** | **(SE)** |  | **%** | **(SE)** | **%** | **(SE)** | **%** | **(SE)** |
| **I. Socio-demographics** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Gender |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Male | 80.7 | (0.0) | 77.5 | (0.8) | 81.1 | (0.8) |  | 85.9 | (0.0) | 88.7 | (0.5) | 85.4 | (0.8) |
| Female | 19.3 | (0.0) | 22.5 | (0.8) | 18.9 | (0.8) |  | 14.1 | (0.0) | 11.3 | (0.5) | 14.6 | (0.8) |
| Age |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 18-22 | 76.2 | (0.0) | 67.2 | (0.7) | 73.5 | (0.7) |  | 18.2 | (0.0) | 14.6 | (0.5) | 16.7 | (0.7) |
| 23-27 | 16.7 | (0.0) | 21.9 | (0.6) | 18.2 | (0.7) |  | 28.3 | (0.0) | 31.0 | (0.6) | 28.6 | (0.9) |
| 28-33 | 6.1 | (0.0) | 9.4 | (0.5) | 7.3 | (0.4) |  | 26.8 | (0.0) | 30.3 | (0.5) | 29.2 | (0.9) |
| 34+ | 1.0 | (0.0) | 1.5 | (0.2) | 0.9 | (0.1) |  | 26.6 | (0.0) | 24.1 | (0.7) | 25.4 | (1.0) |
| Race-ethnicity |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Non-Hispanic White | 60.0 | (0.0) | 65.8 | (0.7) | 62.4 | (1.0) |  | 61.2 | (0.0) | 70.5 | (0.6) | 65.6 | (1.4) |
| Non-Hispanic Black | 20.6 | (0.0) | 16.4 | (0.6) | 18.6 | (0.8) |  | 20.5 | (0.0) | 11.0 | (0.4) | 15.8 | (0.8) |
| Hispanic | 13.2 | (0.0) | 11.6 | (0.5) | 12.6 | (0.7) |  | 11.2 | (0.0) | 11.8 | (0.4) | 11.6 | (0.8) |
| Other | 6.2 | (0.0) | 6.2 | (0.3) | 6.4 | (0.4) |  | 7.1 | (0.0) | 6.7 | (0.3) | 7.0 | (0.8) |
| Highest education |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Less than high school | 16.9 | (0.0) | 14.7 | (0.5) | 16.0 | (0.7) |  | 4.2 | (0.0) | 1.8 | (0.2) | 1.7 | (0.4) |
| Certificate, GED, ARNG | 5.7 | (0.0) | 3.2 | (0.2) | 3.5 | (0.4) |  | 8.5 | (0.0) | 6.7 | (0.3) | 6.4 | (0.6) |
| High school | 69.1 | (0.0) | 65.7 | (0.7) | 69.1 | (0.8) |  | 60.3 | (0.0) | 62.0 | (0.7) | 60.4 | (1.1) |
| Some college | 2.0 | (0.0) | 2.8 | (0.2) | 2.2 | (0.2) |  | 3.8 | (0.0) | 4.7 | (0.2) | 4.9 | (0.4) |
| College graduate | 6.3 | (0.0) | 13.6 | (0.6) | 9.2 | (0.5) |  | 23.0 | (0.0) | 24.8 | (0.7) | 26.5 | (1.3) |
| Marital status |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Never married | 87.8 | (0.0) | 84.2 | (0.5) | 85.0 | (0.6) |  | 32.1 | (0.0) | 29.8 | (0.6) | 34.1 | (1.2) |
| Previously married | 1.1 | (0.0) | 0.1 | (0.0) | 0.0 | (0.0) |  | 6.4 | (0.0) | 4.2 | (0.2) | 3.9 | (0.4) |
| Currently married | 11.0 | (0.0) | 15.6 | (0.5) | 15.0 | (0.6) |  | 61.5 | (0.0) | 66.0 | (0.6) | 62.0 | (1.2) |
| **II. Army careerb** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Rank |  |  |  |  |  |  |  |  |  |  |  |  |  |
| E1-E4 |  |  |  |  |  |  |  | 43.8 | (0.0) | 38.0 | (0.8) | 42.8 | (1.2) |
| E5-E9 |  |  |  |  |  |  |  | 38.1 | (0.0) | 43.9 | (0.6) | 38.2 | (1.0) |
| Warrant officer |  |  |  |  |  |  |  | 2.9 | (0.0) | 2.6 | (0.2) | 3.0 | (0.5) |
| Commissioned officer |  |  |  |  |  |  |  | 15.2 | (0.0) | 15.5 | (0.5) | 15.9 | (1.0) |

|  |
| --- |
| **Supplementary Table S4 (continued). Comparison of socio-demographic and Army career in the LS2 sample versus the population both before and after weighting** |
|  |
|  | **New Soldier Study** |  | **All Army Study/Pre-Post Deployment Study** |
|  |  |  | **LS2** **n=5,172** |  |  |  | **LS2** **n=6,984** |
|  | **Populationa N=869,996** | **Unweighted** | **Weighted** |  | **Populationa N=8,108,371** | **Unweighted** | **Weighted** |
|  | **%** | **(SE)** | **%** | **(SE)** | **%** | **(SE)** |  | **%** | **(SE)** | **%** | **(SE)** | **%** | **(SE)** |
| Time in the Army |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 1-24 months |  |  |  |  |  |  |  | 16.8 | (0.0) | 9.0 | (0.4) | 16.5 | (0.9) |
| 25-60 months |  |  |  |  |  |  |  | 26.6 | (0.0) | 31.6 | (0.8) | 25.6 | (0.8) |
| 61-120 months |  |  |  |  |  |  |  | 24.3 | (0.0) | 28.3 | (0.6) | 25.3 | (1.0) |
| 121+ months |  |  |  |  |  |  |  | 32.3 | (0.0) | 31.0 | (0.8) | 32.6 | (1.2) |
| Command |  |  |  |  |  |  |  |  |  |  |  |  |  |
| FORSCOM |  |  |  |  |  |  |  | 51.1 | (0.0) | 63.6 | (1.1) | 46.2 | (1.5) |
| TRADOC |  |  |  |  |  |  |  | 9.8 | (0.0) | 3.8 | (0.3) | 9.9 | (1.0) |
| North/South America, Europe/Central/Africa, Pacific |  |  |  |  |  | 15.3 | (0.0) | 8.6 | (0.6) | 12.7 | (1.2) |
| Special Operations |  |  |  |  |  |  |  | 5.2 | (0.0) | 3.8 | (0.7) | 5.2 | (0.8) |
| MEDCOM |  |  |  |  |  |  |  | 6.8 | (0.0) | 4.8 | (0.5) | 7.6 | (1.0) |
| AMC/other/unknown |  |  |  |  |  |  |  | 11.7 | (0.0) | 4.1 | (0.4) | 8.6 | (0.8) |
| Guard/Reserve |  |  |  |  |  |  |  | 0.0 | (0.0) | 11.3 | (0.6) | 9.7 | (0.9) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Abbreviations: LS2, Study to Assess Risk & Resilience in Servicemembers-Longitudinal Study Wave 2; SE, standard error; GED, General Educational Development; ARNG, Army National Guard; FORSCOM, United States Army Forces Command; TRADOC, United States Army Training and Doctrine Command; North/South America, Europe/Central/Africa, Pacific, US Army Service Component Central Command, Northern Command, Southern Command, Europe & Africa Command, and Pacific Command; MEDCOM, United States Army Medical Command; AMC, United States Army Materiel Command.

aThe population for the New Soldier Study (NSS) survey was defined as all new soldiers in the Army over the years the NSS was administered (2011-2012), with records defined for person-months. The population for the All Army Study (AAS) survey was defined as all soldiers in duty units in the Army over the years the AAS was administered (2011-2013), again with records defined for person-months. The population for the Pre-Post Deployment Study (PPDS) survey was defined as all soldiers in Brigade Combat Teams over the year the PPDS baseline survey as administered (2012).

bArmy career characteristics were not examined for the NSS, as all NSS respondents were new soldiers.

|  |
| --- |
| **Supplementary Table S5. Super Learner algorithm weights and ROC-AUCs in the test sample (n=2509) for the item-level model with 64 predictors** |
|  | **Weight** | **ROC-AUC** |
| **I. Linear algorithms** |  |  |
| Elastic Net (alpha = .865, screen = 30 lasso) | 4.5% | 80.8 |
| Support Vector Machine (kernel = radial, cost = 4.583, gamma = 1.213; screen = 30 lasso | 6.6 | 59.3 |
|  |  |  |
| **II. Tree-based algorithms**  |  |  |
| Ranger (num.trees = 1000, max.depth = 5, min.node.size = 10, mtry = 2; screen = 15 dbarts)  | 15.8 | 60.4 |
| Ranger (num.trees = 1000, max.depth = 5, min.node.size = 10, mtry = 2; screen = 15 dbarts)  | 11.3 | 60.3 |
| Ranger (num.trees = 1000, max.depth = 4, min.node.size = 4, mtry = 5; screen = 15 dbarts) | 37.1 | 80.9 |
| Ranger (num.trees = 1000, max.depth = 3, min.node.size = 4, mtry = 22; screen = 30 dbarts) | 12.7 | 72.5 |
| XGBoost (ntrees = 1000, max\_depth = 4, shrinkage = .265, gamma = 3.047, minob/node = 16, early\_stopping\_rounds = 50, colsample\_bytree = .181, colsample\_bynode = .76, screen = 30 dbarts) | 1.4 | 66.7 |
| XGBoost (ntrees = 1000, max\_depth = 2, shrinkage = .765, gamma = .738, minob/node = 19, early\_stopping\_rounds = 50, colsample\_bytree = .717, colsample\_bynode = .505, screen = 60 dbarts) | 3.2 | 68.9 |
| XGBoost (ntrees = 1000, max\_depth = 3, shrinkage = .509, gamma = .516, minob/node = 34, early\_stopping\_rounds = 50, colsample\_bytree = .121, colsample\_bynode = .893, screen = 60 dbarts) | 7.8 | 51.5 |
|  |  |  |
| **III. Total Super Learner ensemble** |  |  |
| Total | 100.0 | 80.8 |
|  |  |  |

Abbreviations: ROC-AUC, area under the receiver operating characteristic curve.

|  |
| --- |
| **Supplementary Table S6. The 64 predictors in the item-level super learner modela**  |
|  |
| **I. Socio-demographics** |
| 1. Age: 34+ years old at the time of leaving or being released from active service
 |
| 1. Relationship status & sexual orientation: Sexual orientation (Heterosexual/straight, gay/lesbian/bisexual)
 |
| 1. Children: Any dependents ages 0-5; Any dependents ages 6-13; Has 1+ biological child
 |
|  |
| **II. Army career** |
| 1. Deployments: 2+ GWOT deployments
 |
| 1. Time in service: Standardized continuous number of months active/not active in any service component since Army enlistment
 |
| 1. Service characterization: Discharged Honorably or Under Honorable Conditions
 |
|  |
| **III. Mental disordersb** |
| 1. Lifetime frequency of depression symptomsc: Felt sad or depressed; Discouraged about how things were going in life; Had little or no interest or pleasure in things; Felt down, no good, or worthless
 |
| 1. Lifetime frequency of substance use-related problemsd: Interfered with responsibilities at work, school, home, or on duty; Caused arguments or other serious problems with family, friends, neighbors, or unit members; Under the influence in risky/potentially harmful situations hurt, like when driving or using a weapon
 |
| 1. Lifetime frequency of childhood conduct/behavioral problemse: Bullied or threatened other kids; Started fights; Set fires; Lied or conned people; Ran away from home and stayed away overnight; Stayed out very late; Skipped school; Argued or talked back to adults; Disobeyed rules at home, school, or work; Refused to follow directions from adults; Blamed others for bad behavior or mistakes; Did mean things to pay people back
 |
| 1. Other lifetime psychiatric comorbidities: Ever had an episode of high mood; Number of anger attacks (Raw continuous)
 |
|  |
| **IV. Self-injurious thoughts and behaviors** |
| 1. Suicidal events: Ever had thoughts of killing self; Ever wanted to be dead or go to sleep and never wake up; Lifetime active/passive suicidal ideation; Active/passive ideation within 2 years of leaving or being released from active service; Lifetime suicidal plan; Lifetime suicide attempt; Attempt within 2 years prior to leaving or being released from active service
 |
|  |
| **V. Adverse childhood experiences** |
| 1. Abuse & neglect: Family said hurtful or insulting things (Standardized continuous)
 |
| 1. Other adverse childhood events: Family was homeless
 |
|  |
| **VI. Other lifetime traumatic events**  |
| 1. Stressful experiences during deployment: Went on combat patrols/dangerous duty; Had a close call; Wounded by the enemy; Fired rounds at the enemy/took enemy fire; Responsible for death of enemy combatant; Responsible for death of non-combatant; Responsible for death of US or ally personnel; Saved the life of a servicemember/civilian; Saw destroyed homes/villages or people begging for food; Exposed to wounded/dying people; Witnessed violence in the local population/mistreatment toward non-combatants; Any other highly stressful experience during deployment
 |
| 1. Other lifetime stressful experiences: Experienced a life-threatening illness/injury; Physically assaulted, sexually assaulted, or raped; Exposed to a natural disaster; In a life-threatening accident or other experience that caused risk of injury/death; Unexpected death of a close friend/relative; Any other highly stressful experience; Witnessed someone being seriously injured/killed; Discovered or handled a dead body
 |
| 1. Crime: Victim of any criminal offense in the 4 years prior to leaving or being released from active service
 |
|  |
| **VII. Personality characteristics**  |
| 1. Antisocial traitsf: Do things that are against the law; Often have to lie to get own way; Sometimes hit people so hard they get bruises or have to see a doctor; Sometimes do things that might indirectly harm other people; Feel justified in doing some things that other people might see as wrong
 |
|  |

Abbreviations: GWOT, Global War on Terror.

aConsisting of 57 survey variables plus 7 administrative variables that we felt could be assessed with self-report survey questions. All variables were defined as of the time period prior to the respondent leaving or being released from active service.

bThe psychiatric comorbidities that screened in during phase 1 included lifetime MDE, Broad bipolar disorder, IED, Substance use disorder, and Childhood conduct & ODD. We took the screening items that were asked of everyone for each disorder rather than the full scale due to skip logic in the instrument. The raw scores of each screening item were reverse scored so that higher values indicated greater frequency of symptoms. These reverse scored raw variables were included in the super learner.

cAll respondents were asked how often they experienced 4 depression-related symptoms during the worst month in their life. Response options were reverse scored so that 4 = all or almost all of the time, 3 = most of the time, 2 = some of the time, 1 = a little of the time, and 0 = none of the time.

dIf respondents endorsed the use of alcohol or drugs at any time in their life, they were asked often they experienced each problem related to their alcohol or drug use during the times in their life when they used alcohol or drugs most often. Each item was reverse scored so that 5 = every or nearly every day, 4 = 3-4 days a week, 3 = 1-2 days a week, 2 = 1-3 days a month, 1 = less than once a month, and 0 = never.

eRespondents were asked how often they had each experience up through the age of 17, with response options of very often, often, sometimes, rarely, or never. Response options were reversed scored, resulting in 12 variables ranging from 4-0.

fAll respondents were asked to rate how well each antisocial personality trait described them, with response options of exactly, a lot, somewhat, a little, and not at all. As in the psychiatric comorbidity scales, response options were reverse scored (4-0 range) and we only used the items that were asked of everyone across all surveys.

fRespondents were asked how well each personality trait described them, with response options of not at all, a little, somewhat, a lot, and exactly.

Not included in Phase 5 (in replicates 3-10, 34)

n=5,926d

Eligible for Phase 5

n=14,111d

Not selected for Phase 5 n=11,988

Selected for Phase 5 (End Game, Short Non-Response Interview) n=2,123

End game not complete, no indication of correct contact info, no questions answered n=1,109

End game not complete, indication of correct contact info, no questions answered n=467

End game not complete, interview started, 1+ questions answered

n=34

Deceased (from Phase 5, coded as non-sample)

n=6

Final Non-Interviews in Phase 5

Refusals: n=20

Short End Game Interviews complete in Phase 5

n=487

Not selected for Phases 3-4 n=16,909

Selected for Phases 3-4 n=26,516

Main study not complete, no indication of correct contact info, no questions answered

n=14,455

Main study not complete, indication of correct contact info, no questions answered

n=4,470

Main study not complete, but interview started, 1+ questions answered n=1,112

Final Non-Interviews in Phases 3-4

Refusals: n=248

Other Non-Iws: n=47

Interviews complete in Phases 3-4 (Harvard definition) n=6,184

Eligible for STARRS-LS1 n=72,387a

Not selected for Phase 1 contact n=20,424

Non-Sample Duplicate removed n=1

Selected for Phase 1 contact n=51,962b

Phases 1-2 Not Complete n=43,425

Final Non-Interviews in Phases 1-2

Refusals: n=6

Other Non-Iws: n=58

Non-Sample Deceased (from Phases 1-4)

n=149

Interviews complete in Phases 1-2 (Harvard definition) n=8,324c

Eligible for Phases 3-4 n=43,425c

Abbreviations: STARRS-LS1, Study to Assess Risk & Resilience in Servicemembers-Longitudinal Study Wave 1; Iws, interviews.

aAs noted in the text, LS1 was carried out in a probability sample of participants from the earlier Army STARRS survey. The design of the Army STARRS surveys is described elsewhere.2 The n=72,387 baseline respondents were weighted to post-stratify for a wide range of socio-demographic and Army career variables recorded in administrative systems for all the over 500,000 soldiers in active serviceas a point in time during the years of Army STARRS Recruitment. Non-response bias and weighting adjustments to correct for this bias in the baseline Army STARRS surveys are described elsewhere.3

bSee the text for a brief overview of stratification used for the Army STARRS-LS Wave 1 sample selection.

cAs in the original Army STARRS surveys,3 recruitment in the LS surveys was carried out in phases in which an attempt was made to survey all predesignated respondents in the first 2 phases and then subsampling was used to administer more intensive recruitment efforts only to a probability subsample of phases 1-2 non-respondents. Appropriate weights were used to adjust for this under-sampling of hard-to-recruit respondents at the end of data collection.

dPhase 5 was an attempt to obtain a small amount of information from a probability sample of non-respondents to facilitate model-based adjustment for non-response. This “endgame” short form survey administered a small number of marker questions about suicidal behaviors and other outcomes of special interest to the research team, but the endgame survey respondents were still treated as non-respondents in analysis. That is, their data were not used in the analysis sample but rather were used only to help improve the weights used in the sample of respondents who completed the full survey.

 **Supplementary Figure S1. Recruitment diagram for the STARRS-LS1 survey**

**Supplemental References**

End game not complete, no indication of correct contact info, no questions answered n=212

End game not complete, indication of correct contact info, no questions answered n=44

End game not complete, interview started, 1+ questions answered

n=4

Short End Game Interviews complete in Phase 5

n=19

Final Non-Interviews in Phase 5

Refusals: n=3

Deceased (from Phase 5, coded as non-sample)

n=0

Main study not complete, no indication of correct contact info, no questions answered

n=1,107

Main study not complete, indication of correct contact info, no questions answered n=983

Main study not complete, interview started, 1+ questions answered

n=180

Final Non-Interviews in Phases 3-4 Refusals: n=55

(10 withdrawals)

Other Non-Iws: n=4

Interviews complete in Phases 3-4

(Harvard definition)

n=5,365

Eligible for Phase 5 n=2,270

Not selected for Phase 5 n=1,808

Selected for Phase 5 (End Game, Short Non-Response Interview)

n=462

Eligible for Phases 3-4 n=7,694

Phases 1-2

Not Complete

n=7,694

Final Non-Interviews in Phases 1-2

Refusals: n=8

(6 withdrawals)

Other Non-Iws: n=5

Non-Sample

Deceased (from Phases 1-4) n=10

Eligible for STARRS-LS2 n=14,522b

Non-Sample

Wave 1 Wrong Respondent n=14

Phase 1

n=14,508

Interviews complete in Phases 1-2 (Harvard definition)

n=6,791

**Supplementary Figure S2. Recruitment diagram for the STARRS-LS2 surveya**

Abbreviations: STARRS-LS2, Study to Assess Risk & Resilience in Servicemembers-Longitudinal Study Wave 2; Iws, interviews.

aSee the footnotes in Supplementary Figure S1 for an overview of the phases used in all Army STARRS and STARRS-LS surveys and the text for a brief overview of stratification used for LS1/LS2 sample selection. These procedures are discussed in detail in earlier Army STARRS reports.2

bAll STARRS-LS Wave 1 respondents were eligible for LS2.

Afifi, T. O., Taillieu, T., Zamorski, M. A., Turner, S., Cheung, K., & Sareen, J. (2016). Association of child abuse exposure with suicidal ideation, suicide plans, and suicide attempts in military personnel and the general population in Canada. *JAMA Psychiatry, 73*(3), 229-238. doi:10.1001/jamapsychiatry.2015.2732

Bostwick, J. M., Pabbati, C., Geske, J. R., & McKean, A. J. (2016). Suicide Attempt as a Risk Factor for Completed Suicide: Even More Lethal Than We Knew. *American Journal of Psychiatry, 173*(11), 1094-1100. doi:10.1176/appi.ajp.2016.15070854

Bryan, C. J., Bryan, A. O., Roberge, E., Leifker, F. R., & Rozek, D. C. (2018). Moral injury, posttraumatic stress disorder, and suicidal behavior among National Guard personnel. *Psychological Trauma, 10*(1), 36-45. doi:10.1037/tra0000290

Bryan, C. J., & Clemans, T. A. (2013). Repetitive traumatic brain injury, psychological symptoms, and suicide risk in a clinical sample of deployed military personnel. *JAMA Psychiatry, 70*(7), 686-691. doi:10.1001/jamapsychiatry.2013.1093

Bryan, C. J., Griffith, J. E., Pace, B. T., Hinkson, K., Bryan, A. O., Clemans, T. A., . . . Imel, Z. E. (2015). Combat Exposure and Risk for Suicidal Thoughts and Behaviors Among Military Personnel and Veterans: A Systematic Review and Meta-Analysis. *Suicide and Life-Threatening Behavior, 45*(5), 633-649. doi:10.1111/sltb.12163

Chen, R. J., Chen, T. Y., Lipkova, J., Wang, J. J., Williamson, D. F. K., Lu, M. Y., . . . Mahmood, F. (2022). Algorithm Fairness in AI for Medicine and Healthcare. Retrieved from https://arxiv.org/abs/2110.00603

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System *International Conference on Knowledge Discovery and Data Mining* (pp. 785-794): ACM. Retrieved from https://dl.acm.org/doi/10.1145/2939672.2939785

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273-297. doi:10.1007/BF00994018

Cougle, J. R., Keough, M. E., Riccardi, C. J., & Sachs-Ericsson, N. (2009). Anxiety disorders and suicidality in the National Comorbidity Survey-Replication. *Journal of Psychiatric Research, 43*(9), 825-829. doi:10.1016/j.jpsychires.2008.12.004

Daugherty, J., Waltzman, D., Sarmiento, K., & Xu, L. (2019). Traumatic Brain Injury-Related Deaths by Race/Ethnicity, Sex, Intent, and Mechanism of Injury - United States, 2000-2017. *MMWR Morbidity and Mortality Weekly Report, 68*(46), 1050-1056. doi:10.15585/mmwr.mm6846a2

Dervic, K., Oquendo, M. A., Grunebaum, M. F., Ellis, S., Burke, A. K., & Mann, J. J. (2004). Religious affiliation and suicide attempt. *American Journal of Psychiatry, 161*(12), 2303-2308. doi:10.1176/appi.ajp.161.12.2303

Dube, S. R., Anda, R. F., Felitti, V. J., Chapman, D. P., Williamson, D. F., & Giles, W. H. (2001). Childhood abuse, household dysfunction, and the risk of attempted suicide throughout the life span: Findings from the Adverse Childhood Experiences Study. *JAMA, 286*(24), 3089-3096. doi:10.1001/jama.286.24.3089

Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., . . . Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin, 143*(2), 187-232. doi:10.1037/bul0000084

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software, 33*(1), 1-22.

Heeringa, S. G., Gebler, N., Colpe, L. J., Fullerton, C. S., Hwang, I., Kessler, R. C., . . . Ursano, R. J. (2013). Field procedures in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *International Journal of Methods in Psychiatric Research, 22*(4), 276-287. doi:10.1002/mpr.1400

Holliday, R., Borges, L. M., Stearns-Yoder, K. A., Hoffberg, A. S., Brenner, L. A., & Monteith, L. L. (2020). Posttraumatic stress disorder, suicidal ideation, and suicidal self-directed violence among U.S. Military personnel and veterans: A systematic review of the literature from 2010 to 2018. *Frontiers in Psychology, 11*, 998. doi:10.3389/fpsyg.2020.01998

Huang, K. L., Wei, H. T., Hsu, J. W., Bai, Y. M., Su, T. P., Li, C. T., . . . Chen, M. H. (2018). Risk of suicide attempts in adolescents and young adults with attention-deficit hyperactivity disorder: A nationwide longitudinal study. *British Journal of Psychiatry, 212*(4), 234-238. doi:10.1192/bjp.2018.8

Jiang, T., Rosellini, A. J., Horváth-Puhó, E., Shiner, B., Street, A. E., Lash, T. L., . . . Gradus, J. L. (2021). Using machine learning to predict suicide in the 30 days after discharge from psychiatric hospital in Denmark. *The British Journal of Psychiatry, 219*(2), 440-447. doi:10.1192/bjp.2021.19

Kennedy, C. J. (2017). Guide to SuperLearner. Retrieved from https://cran.r-project.org/web/packages/SuperLearner/vignettes/Guide-to-SuperLearner.html

Kessler, R. C., Colpe, L. J., Fullerton, C. S., Gebler, N., Naifeh, J. A., Nock, M. K., . . . Heeringa, S. G. (2013). Design of the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *International Journal of Methods in Psychiatric Research, 22*(4), 267-275. doi:10.1002/mpr.1401

Kessler, R. C., Heeringa, S. G., Colpe, L. J., Fullerton, C. S., Gebler, N., Hwang, I., . . . Ursano, R. J. (2013). 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, 22*(4), 288-302. doi:10.1002/mpr.1399

Khazem, L. R., & Anestis, M. D. (2019). Do physical disabilities differentiate between suicidal ideation and attempts? An examination within the lens of the ideation to action framework of suicide. *Journal of Clinical Psychology, 75*(4), 681–695. doi:10.1002/jclp.22735

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

Kposowa, A. J. (2000). Marital status and suicide in the National Longitudinal Mortality Study. *Journal of Epidemiology and Community Health, 54*(4), 254-261. doi:10.1136/jech.54.4.254

Lawrence, R. E., Oquendo, M. A., & Stanley, B. (2016). Religion and Suicide Risk: A Systematic Review. *Archives of Suicide Research, 20*(1), 1-21. doi:10.1080/13811118.2015.1004494

LeDell, E., van der Laan, M. J., & Petersen, M. (2016). AUC-Maximizing Ensembles through Metalearning. *The International Journal of Biostatistics, 12*(1), 203-218. doi:10.1515/ijb-2015-0035

Lundberg, S., & Lee, S.-I. (2017, December 4-9). A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), Red Hook, NY. Retrieved from https://papers.nips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.

Matarazzo, B. B., Barnes, S. M., Pease, J. L., Russell, L. M., Hanson, J. E., Soberay, K. A., . . . Gutierrez, P. M. (2014). Suicide risk among lesbian, gay, bisexual, and transgender military personnel and veterans: what does the literature tell us? *Suicide and Life-Threatening Behavior, 44*(2), 200-217. doi:10.1111/sltb.12073

Millner, A. J., Ursano, R. J., Hwang, I., King, A. J., Naifeh, J. A., Sampson, N. A., . . . Nock, M. K. (2018). Lifetime suicidal behaviors and career characteristics among U.S. Army soldiers: Results from the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *Suicide and Life-Threatening Behavior, 48*(2), 230-250. doi:10.1111/sltb.12363

Milner, A., Hjelmeland, H., Arensman, E., & Leo, D. (2013). Social-environmental factors and suicide mortality: A narrative review of over 200 articles. *Sociology Mind, 3*(137-148. doi:10.4236/sm.2013.32021

Mitchell, M. M., Gallaway, M. S., Millikan, A. M., & Bell, M. (2012). Interaction of combat exposure and unit cohesion in predicting suicide-related ideation among post-deployment soldiers. *Suicide and Life-Threatening Behavior, 42*(5), 486-494. doi:10.1111/j.1943-278X.2012.00106.x

Naifeh, J. A., Ursano, R. J., Kessler, R. C., Zaslavsky, A. M., Nock, M. K., Dempsey, C. L., . . . Stein, M. B. (2019). Transition to suicide attempt from recent suicide ideation in U.S. Army soldiers: Results from the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *Depression and Anxiety, 36*(5), 412-422. doi:10.1002/da.22870

Nock, M. K., Borges, G., Bromet, E. J., Alonso, J., Angermeyer, M., Beautrais, A., . . . Williams, D. (2008). Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *British Journal of Psychiatry, 192*(2), 98-105. doi:10.1192/bjp.bp.107.040113

Nock, M. K., Deming, C. A., Fullerton, C. S., Gilman, S. E., Goldenberg, M., Kessler, R. C., . . . Ursano, R. J. (2013). Suicide among soldiers: a review of psychosocial risk and protective factors. *Psychiatry, 76*(2), 97-125. doi:10.1521/psyc.2013.76.2.97

Nock, M. K., Stein, M. B., Heeringa, S. G., Ursano, R. J., Colpe, L. J., Fullerton, C. S., . . . Kessler, R. C. (2014). Prevalence and correlates of suicidal behavior among soldiers: results from the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *JAMA Psychiatry, 71*(5), 514-522. doi:10.1001/jamapsychiatry.2014.30

Panagioti, M., Gooding, P., & Tarrier, N. (2009). Post-traumatic stress disorder and suicidal behavior: A narrative review. *Clinical Psychology Review, 29*(6), 471-482. doi:10.1016/j.cpr.2009.05.001

Phillips, J. A., & Hempstead, K. (2017). Differences in U.S. suicide rates by educational attainment, 2000-2014. *American Journal of Preventive Medicine, 53*(4), e123-e130. doi:10.1016/j.amepre.2017.04.010

Polley, E., LeDell, E., Kennedy, C. J., Lendle, S., & van der Laan, M. J. (2015). SuperLearner: Super Learner Prediction (Version 2.0-28): CRAN. Retrieved from https://cran.r-project.org/web/packages/SuperLearner/index.html

Posner, K., Brown, G. K., Stanley, B., Brent, D. A., Yershova, K. V., Oquendo, M. A., . . . Mann, J. J. (2011). The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *American Journal of Psychiatry, 168*(12), 1266-1277. doi:10.1176/appi.ajp.2011.10111704

Ratcliffe, G. E., Enns, M. W., Belik, S. L., & Sareen, J. (2008). Chronic pain conditions and suicidal ideation and suicide attempts: an epidemiologic perspective. *The Clinical Journal of Pain, 24*(3), 204-210. doi:10.1097/AJP.0b013e31815ca2a3

Reger, M. A., Smolenski, D. J., Skopp, N. A., Metzger-Abamukong, M. J., Kang, H. K., Bullman, T. A., & Gahm, G. A. (2017). Suicide risk among wounded U.S. Service Members. *Suicide and Life-Threatening Behavior, 47*(2), 242–247. doi:10.1111/sltb.12282

Reger, M. A., Smolenski, D. J., Skopp, N. A., Metzger-Abamukang, M. J., Kang, H. K., Bullman, T. A., . . . Gahm, G. A. (2015). Risk of suicide among U.S. Military service members following Operation Enduring Freedom or Operation Iraqi Freedom deployment and separation from the U.S. Military. *JAMA Psychiatry, 72*(6), 561-569. doi:10.1001/jamapsychiatry.2014.3195

Ribeiro, J. D., Franklin, J. C., Fox, K. R., Bentley, K. H., Kleiman, E. M., Chang, B. P., . . . Nock, M. K. (2016). Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: a meta-analysis of longitudinal studies. *Psychological Medicine, 46*(2), 225-236. doi:10.1017/s0033291715001804

Ryb, G. E., Soderstrom, C. A., Kufera, J. A., & Dischinger, P. (2006). Longitudinal study of suicide after traumatic injury. *The Journal of Trauma, 61*(4), 799–804. doi:10.1097/01.ta.0000196763.14289.4e

Saxe, G. N., Ma, S., Ren, J., & Aliferis, C. (2017). Machine learning methods to predict child posttraumatic stress: a proof of concept study. *BMC psychiatry, 17*(1), 223-223. doi:10.1186/s12888-017-1384-1

Soloff, P. H., Lynch, K. G., Kelly, T. M., Malone, K. M., & Mann, J. J. (2000). Characteristics of suicide attempts of patients with major depressive episode and borderline personality disorder: A comparative study. *American Journal of Psychiatry, 157*(4), 601-608. doi:10.1176/appi.ajp.157.4.601

Stanley, I. H., Chu, C., Gildea, S. M., Hwang, I. H., King, A. J., Kennedy, C. J., . . . Kessler, R. C. (2022). Predicting suicide attempts among U.S. Army soldiers after leaving active duty using information available before leaving active duty: results from the Study to Assess Risk and Resilience in Servicemembers-Longitudinal Study (STARRS-LS). *Molecular Psychiatry, 27*(3), 1631-1639. doi:10.1038/s41380-021-01423-4

Stanley, I. H., Joiner, T. E., & Bryan, C. J. (2017). Mild traumatic brain injury and suicide risk among a clinical sample of deployed military personnel: Evidence for a serial mediation model of anger and depression. *Journal of Psychiatric Research, 84*, 161-168. doi:10.1016/j.jpsychires.2016.10.004

Stanley, I. H., Rogers, M. L., Hanson, J. E., Gutierrez, P. M., & Joiner, T. E. (2019). PTSD symptom clusters and suicide attempts among high-risk military service members: A three-month prospective investigation. *Journal of Consulting and Clinical Psychology, 87*(1), 67-78. doi:10.1037/ccp0000350

Stein, M. B., Campbell-Sills, L., Ursano, R. J., Rosellini, A. J., Colpe, L. J., He, F., . . . Kessler, R. C. (2018). Childhood maltreatment and lifetime suicidal behaviors among new soldiers in the US Army: Results From the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *Journal of Clinical Psychiatry, 79*(2), 16m10900. doi:10.4088/JCP.16m10900

Stewart, J. G., Shields, G. S., Esposito, E. C., Cosby, E. A., Allen, N. B., Slavich, G. M., . . . Auerbach, R. P. (2019). Life Stress and Suicide in Adolescents. *Journal of Abnormal Child Psychology, 47*(10), 1707-1722. doi:10.1007/s10802-019-00534-5

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological), 58*(1), 267-288. Retrieved from http://www.jstor.org/stable/2346178

United States Department of Justice, Office of Justice Programs, & Bureau of Justice Statistics. (2009). National Corrections Reporting Program. Retrieved from https://www.icpsr.umich.edu/web/NACJD/studies/30799#

Ursano, R. J., Colpe, L. J., Heeringa, S. G., Kessler, R. C., Schoenbaum, M., & Stein, M. B. (2014). The Army study to assess risk and resilience in servicemembers (Army STARRS). *Psychiatry, 77*(2), 107-119. doi:10.1521/psyc.2014.77.2.107

Ursano, R. J., Heeringa, S. G., Stein, M. B., Jain, S., Raman, R., Sun, X., . . . Kessler, R. C. (2015). Prevalence and correlates of suicidal behavior among new soldiers in the U.S. Army: Results from the Army Study to Assess Risk and Resilience in Servicemembers (ARMY STARRS). *Depression and Anxiety, 32*(1), 3-12. doi:10.1002/da.22317

Ursano, R. J., Kessler, R. C., Naifeh, J. A., Herberman Mash, H., Fullerton, C. S., Bliese, P. D., Zaslavsky, A. M., Ng, T. H. H., Aliaga, P. A., Wynn, G. H., Dinh, H. M., McCarroll, J. E., Sampson, N. A., Kao, T. C., Schoenbaum, M., Heeringa, S. G., Stein, M. B., & Army Study to Assess Risk and Resilience in Servicemembers (STARRS) Collaborators (2017). Risk of suicide attempt among soldiers in army units with a history of suicide attempts. *JAMA Psychiatry, 74*(9), 924–931. doi:10.1001/jamapsychiatry.2017.1925

Ursano, R. J., Kessler, R. C., Naifeh, J. A., Herberman Mash, H. B., Nock, M. K., Aliaga, P. A., . . . Stein, M. B. (2018). risk factors associated with attempted suicide among U.S. Army soldiers without a history of mental health diagnosis. *JAMA Psychiatry, 75*(10), 1022-1032. doi:10.1001/jamapsychiatry.2018.2069

van der Laan, M. J., Polley, E., & Hubbard, A. E. (2007). Super learner. *Statistical Applications in Genetics and Molecular Biology, 6*, 25. doi:10.2202/1544-6115.1309

Weathers, F., Litz, B., Herman, D., Huska, J. A., & Keane, T. (1993). The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility. *Paper Presented at the 9th Annual Convention of the International Society for Traumatic Stress Studies,* San Antonio, Texas.

Wright, M. N., & Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software, 77*(1), 1-17. doi:10.18637/jss.v077.i01

Yen, S., Peters, J. R., Nishar, S., Grilo, C. M., Sanislow, C. A., Shea, M. T., . . . Skodol, A. E. (2021). Association of borderline personality disorder criteria with suicide attempts: Findings from the collaborative longitudinal study of personality disorders over 10 years of follow-up. *JAMA Psychiatry, 78*(2), 187-194. doi:10.1001/jamapsychiatry.2020.3598

Zou, G. (2004). A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *American Journal of Epidemiology, 159*(7), 702-706. doi:10.1093/aje/kwh090

Zuromski, K. L., Bernecker, S. L., Chu, C., Wilks, C. R., Gutierrez, P. M., Joiner, T. E., . . . Kessler, R. C. (2020). Pre-deployment predictors of suicide attempt during and after combat deployment: Results from the Army Study to Assess Risk and Resilience in Servicemembers. *The Journal of Psychiatric Research, 121*, 214-221. doi:10.1016/j.jpsychires.2019.12.003

Zuromski, K. L., Ustun, B., Hwang, I., Keane, T. M., Marx, B. P., Stein, M. B., . . . Kessler, R. C. (2019). Developing an optimal short-form of the PTSD Checklist for DSM-5 (PCL-5). *Depression and Anxiety, 36*(9), 790-800. doi:10.1002/da.22942