

Appendix Tables 1-6

Kessler, R.C et al. Predicting suicides after outpatient mental health visits in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS).

Appendix Table 1. List and brief descriptions of administrative datasets included in the Army STARRS Historical Administrative Data Systems (HADS)

Database Acronym	Description
AFMETS	ARMED FORCES MEDICAL EXAMINER TRACKING SYSTEM (AFMETS): Variables include manner of death and cause of death, including self-inflicted.
MDR	MEDICAL DATA REPOSITORY (MDR): This database contains information about medical, dental, pharmaceutical, and ancillary claims data for both in network and purchased care as well as both inpatient and outpatient treatment. Data are collected on both Army personnel and their beneficiaries.
DCIPS	DEFENSE CASUALTY INFORMATION PROCESSING SYSTEM (DCIPS): Casualty data system for Army, Navy, Air Force, and Marines. It contains information on all casualties including deceased active duty members (and cause of death).
DMDC/Casualty	DEFENSE MANPOWER DATA CENTER (DMDC) / CASUALTY: Provides information on war deaths and active duty deaths, wounds, injuries, and illnesses.
DMDC/CTS	DEFENSE MANPOWER DATA CENTER (DMDC) / CONTINGENCY TRACKING SYSTEM (CTS): Collection of activation, mobilization, and deployment data. Provides information to DoD decision makers and includes a CTS Deployment File used for tracking the location of deployed personnel.
DMDC/Master Personnel & DMDC/Transaction files	DEFENSE MANPOWER DATA CENTER (DMDC) / MASTER PERSONNEL & TRANSACTION FILES: The Active Duty Master File provides an inventory of all individuals on active duty (excluding reservists on active duty for training) at a point in time. It is a standardized and centralized database of present and past members of the active duty force. Personal data elements include social security number, education level, home of record, date of birth, marital status, number of dependents, race, ethnic group, and name. Military data elements include Service, pay grade, Armed Forces Qualification Test percentile (enlisted only), source of commission (officers only), military primary duty and secondary occupation, Unit Identification Code, months of service, duty location, Estimated Termination of Service date, basic active service date, date of current rank, pay entry base date, foreign language ability, and major command code.
DODSER	DEPARTMENT OF DEFENSE SUICIDE EVENT REPORT (DODSER): Provides risk and protective factor information for suicide events. This file will contain unsuccessful attempts and completed suicide cases.
TMDS	THEATER MEDICAL DATA STORE (TMDS): Used to track, analyze, view and manage Soldier medical treatment information recorded on the battlefield. Features of TMDS: accessibility and visibility of service members' deployed medical records, outpatient and inpatient treatment records created in theater facilities, treatment records from other applications, reports on movement of patients, patient status and injury/illnesses.
DMDC/DEERS	DEFENSE MANPOWER DATA CENTER (DMDC) / DEFENSE ENROLLMENT ELIGIBILITY REPORTING SYSTEM (DEERS): A DoD PDR containing personnel data and data related to DoD affiliation, benefits, employment, and pay.
MEDPROS/Readiness	MEDICAL PROTECTION SYSTEM (MEDPROS) / READINESS: The comprehensive Medical Readiness data includes all medical and dental readiness requirements IAW AR 600-8-101. They include immunizations, permanent physical profiles/duty limitations, eyeglasses/inserts, blood type, medical warning tags, personal deployment meds, pregnancy screening, DNA, HIV and dental status among other data elements.
DMDC/Perstempo	DEFENSE MANPOWER DATA CENTER (DMDC) / PERSONNEL TEMPO: Provides information on soldier movement. Perstempo is a Congressionally mandated definition of time in which servicemembers engaged in official duties spend off-duty time away from housing associated with their garrison duty (time away from home station). The definition also applies to contingency operations such as OIF and OEF.
ACR	ARMY CENTRAL REGISTRY (ACR): This is a victim-based registry documenting spouse and child abuse events involving Army service members.
AWD	ARMY WAIVER DATA (AWD): Information on waivers granted to new recruits to join the Army. The waivers are provided for four primary categories: 1) dependence (married with over two children), 2) conduct, 3) drug and alcohol, and 4) medical.
DAMIS	DRUG AND ALCOHOL MANAGEMENT INFORMATION SYSTEM (DAMIS): Includes data from the Army drug and alcohol prevention training and substance abuse program.
DMDC/MEPCOM	DEFENSE MANPOWER DATA CENTER (DMDC) / MILITARY ENTRANCE PROCESSING COMMAND (MEPCOM): Contains transaction records on all individuals processed and examined at Military Entrance Processing Stations (MEPS) since July 1970. Examination data prior to FY 1976 is incomplete. Data elements in the USMEPCOM Files are divided into three groups: Personal, Medical, and Administrative. The primary user is the Office of the Undersecretary of Defense (Personnel and Readiness) where ASVAB data provides the basis for aptitude-related studies.
DMDC/Payroll	DEFENSE MANPOWER DATA CENTER (DMDC) / Payroll: Contains elements with information such as: Demographics, Special and Incentive Pays, Basic Pay and Allowances. This file is also used to answer pay-related questions for the General Accounting Office, Military Services research centers, and other government agencies.

DTMS	DIGITAL TRAINING MANAGEMENT SYSTEM (DTMS): Contains information on training requirements and standards, training planned and scheduled by Army units, and training accomplished by individuals and units in the Army.
MEDPROS/DHA	MEDICAL PROTECTION SYSTEM (MEDPROS) / DEPLOYMENT HEALTH ASSESSMENT (DHA): The DHA cycle begins when a Soldier completes the DD Form 2795 (Pre-Deployment Health Assessment) prior to deployment. Soldiers complete the DD Form 2796 Post Deployment Health Assessment (PDHA) prior to redeployment/demobilization. This triggers their PDHRA eligibility. At 90 days after completion of the DD Form 2796, Soldiers become eligible to participate in a PDHRA screening. The targeted window for PDHRA is 90-180 days after deployment. Once the Soldier is screened, follow-up appointments and treatment will be scheduled as needed.
MEDPROS/PHA	MEDICAL PROTECTION SYSTEM (MEDPROS) / PERIODIC HEALTH ASSESSMENT (PHA): The PHA is an annual physical and consists of two parts. Part one is a self-assessment completed by the Soldier; Part two is the provider assessment. PHA results consist of the health assessment, height, weight, PULHES, and potential for deployability within six months. These items will be entered into MEDPROS at the point of service. The PHA is intended to improve Individual Medical Readiness and to evaluate deployability.
RRPS	RISK AND REDUCTION PROGRAM SYSTEM (RRPS): The RRPS collects data at the unit level that may show that members of the unit are at a higher risk. The Risk Reduction Program collects the number of incidents of high risk behavior by battalion level or separate company.
SADMS	SEXUAL ASSAULT DATA MANAGEMENT SYSTEM (SADMS): Centralized repository of relevant data regarding the entire lifecycle of sexual assault cases, involving victims and/or alleged offenders.
TRAC2ES	TRANSCOM REGULATING AND COMMAND AND CONTROL EVACUATION SYSTEM (TRAC2ES): A tracking system for all medical transfers across the world for all DOD services.
ASMIS-R	ARMY SAFETY MANAGEMENT INFORMATION SYSTEM-REVISED (ASMIS-R): Contains incident information data. Components consist of the reporting, analysis, and tracking of data, as well as, efforts to provide preventative measures via risk assessments. ReportIT provides for the collection, verification, and timely dissemination of incident investigation and safety inspection data. AnalyzeIT serves as the business intelligence suite designed to identify trends and problem areas via data mining techniques and ad hoc analysis. TrackIT is designed for monitoring recommended controls and logging hazards. PreventIT encompasses risk assessment and guidance tools for units, battalions, commanders, civilians, and contractors.
CIMS/AC12	CRIMINAL INVESTIGATION DIVISION INFORMATION MANAGEMENT SYSTEM (CIMS) / AUTOMATED CRIMINAL INVESTIGATION/CRIMINAL INTELLIGENCE (ACI2): Event level crime record database. Separate system from ASCRC; Entity IDs are associated with offenses.
CIMS/ASCRC	CRIMINAL INVESTIGATION DIVISION INFORMATION MANAGEMENT SYSTEM (CIMS) / AUTOMATED SYSTEM CRIME RECORD CENTER (ASCRC): Receives, maintains, accounts for, disseminates information from, and disposes of Army crime records; retrieves and correlates data and statistics from the records and provides to authorized recipients; coordinates automation of crime records data and information; and serves as functional proponent for the electronic imaging of crime record.
COPS/VRS	CENTRALIZED OPERATIONS POLICE SUITE (COPS) / VEHICLE REGISTRATION SYSTEM (VRS): Provides information on privately owned vehicles of civilians and military personnel utilized at military installations.
PDCAPS	PHYSICAL DISABILITY CASE PROCESSING SYSTEM (PDCAPS): Contains information on fitness and applicable disability benefits of Soldiers with duty related impairments, including clinical information of all medical conditions the Soldier manifests to determine whether each condition meets medical retention standards.
ACMIS	ARMY COURT MARTIAL INFORMATION SYSTEM (ACMIS): Contains data on Soldiers (including Officers) arraigned on court-martial charges at either a Special or General Court-Martial. It also contains some basic information concerning them such as their rank, their GT score (if available), their SSNs, and their birthdays. The database includes such information as where the Soldier was tried, which General Court-Martial Jurisdiction convened the case, the offenses charged, how the accused plead, the findings in the case as well as the sentence (if any). It also contains information concerning cases required to proceed through the appellate court process and outcome in the appellate court system. It includes the location of the offenses and trial, and tracks the dates of certain events in the court-martial process. It does not contain any information concerning whether or not an accused has committed suicide.
ANAM	AUTOMATED NEUROPSYCHOLOGICAL ASSESSMENT METRICS (ANAM): This is an Army-developed and Army-owned, objective, computer-based test of human performance and neuropsychological functioning used to assess the cognitive performance of all active duty, reserve, and National Guard Soldiers.
ATRRS	ARMY TRAINING AND REQUIREMENTS RESOURCE SYSTEM (ATRRS): This is a system of record for management of personnel input to training and is the repository for training requirements, programs, personnel data and training costs.
COPS/ MPRS	CENTRALIZED OPERATIONS POLICE SUITE (COPS) / MILITARY POLICE REPORTING SYSTEM (MPRS): Centralized database that contains subsystems supporting the Military Police Corps Army wide. Contains information included in law enforcement reports.
COPS/Violations	CENTRALIZED OPERATIONS POLICE SUITE (COPS) / VIOLATIONS: Violation information contained in MPRS

system (e.g., incident, time of incident).

COPS/ACIS	CENTRALIZED OPERATIONS POLICE SUITE (COPS) / ARMY CORRECTIONAL INFORMATION SYSTEM (ACIS): This System has information for the confinement facility on new prisoners, currently confined prisoners, departed prisoners, and a monthly summary report of confinement facility activity data.
SFT	SOLDIER FITNESS TRACKER (SFT): Formerly known as the Comprehensive Soldier Fitness Global Assessment Tool (CSF-GAT), this source provides baseline information on strength in four areas: emotional, social, spiritual and family. Also provides an opportunity to track self-development and growth in these areas over time. The Chief of Staff of the Army is requiring all Soldiers to take the GAT by May 2010. The object of this program is to enhance performance and build resilience.
EORS	EQUAL OPPORTUNITY REPORTING SYSTEM (EORS): Database is used to collect, record, and maintain racial, ethnic group, and gender data and statistics needed to support the Army EO Program, to include AAP reporting requirements. No health related data. Contains information such as: rank, race, complaint type, category, gender, timeline case start to end, resolution.
ITAPDB	INTEGRATED TOTAL ARMY PERSONNEL DATABASE (ITAPDB): Contains consolidated human resource database providing the Army and DOD with a single repository for personnel information for all components of the Army.
WWAS	WOUNDED WARRIOR ACCOUNTABILITY SYSTEM (WWAS): Integrated applications that support the accurate, timely, and effective tracking and management of Warfighters in the Wounded Warrior Lifecycle.

Appendix Table 2. International Classification of Diseases, Ninth Revision Clinical Modification (ICD-9-CM) codes used to identify mental disorders

Diagnoses	ICD-9-CM codes
Adjustment disorder	309.29, .3, .4, .82, .83, .89, .9
Dysthymic disorder/Neurasthenia/Depression NOS	296.82, .90, .99 300.4, .5 309.0, .1 311, .0, .1 313.1
Major depression	296.2, .20, .21, .22, .23, .24, .25, .26, .3, .30, .31, .32, .33, .34, .35, .36
Bipolar disorder	296.00, .01, .02, .03, .04, .05, .06, .10, .11, .12, .13, .14, .15, .16, .40, .41, .42, .43, .44, .45, .46, .50, .51, .52, .53, .54, .55, .56, .60, .61, .62, .63, .64, .65, .66, .7, .80, .81, .89 301.13
Anxiety state/Anxiety disorder	300, .00, .01, .02, .09, .20, .21, .22, .23, .29, .3 309.21, .24, .28 313.0, .21, .22, .23
PTSD	309.81
ADHD/Learning disorders	314.0, .00, .01, .1, .2, .8, .9 315.00, .01, .02, .09, .1, .2, .3, .31, .32, .34, .39, .4, .5, .8, .9
Conduct disorder/Oppositional defiant disorder	301.7 312.4, .8, .81, .82, .89, .9 313.81 V62.83
Eating disorders	307.1, .50, .51, .59
Other impulse control disorders	312.00, .01, .02, .03, .10, .11, .12, .13, .20, .21, .22, .23, .3, .30, .31, .32, .33, .34, .35, .39
Alcohol induced mental disorders/Alcohol dependence/Alcohol abuse	291.0, .1, .2, .3, .4, .5, .8, .81, .82, .89, .9 303.00, .01, .02, .03, .9, .90, .91, .92, .93 305, .0, .00, .01, .02, .03
Drug induced mental disorders	292
Non-dependent drug abuse	305.2, .20, .21, .22, .23, .3, .30, .31, .32, .33, .4, .40, .41, .42, .43, .5, .50, .51, .52, .53, .6, .60, .61, .62, .63, .7, .70, .71, .72, .73, .8, .80, .81, .82, .83, .9, .90, .91, .92, .93
Drug dependence	304
Tobacco use disorder	305.1, .10, .11, .12, .13
Personality disorders	301.0, .1, .10, .11, .12, .20, .21, .22, .3, .4, .50, .51, .59, .6, .8, .80, .81, .82, .83, .84, .89, .9
Non-affective psychosis	295.00, .01, .02, .03, .04, .05, .10, .11, .12, .13, .14, .15, .20, .21, .22, .23, .24, .25, .30, .31, .32, .33, .34, .35, .40, .41, .42, .43, .44, .45, .50, .51, .52, .53, .54, .60, .61, .62, .63, .64, .65, .70, .71, .72, .73, .74, .75, .80, .81, .82, .83, .84, .85,

	.90, .91, .92, .93, .94, .95 297.0, .1, .2, .3, .8, .9 298.0, .1, .2, .3, .4, .8, .9, .90
Somatoform/Dissociative disorders	300.10, .11, .12, .13, .14, .15, .16, .19, .6, .7, .80, .81, .82, .89 306.0, .1, .2, .3, .4, .50, .51, .52, .53, .59, .6, .7, .8, .9 307.54, .80, .81, .89
Organic mental disorders	290.0, .10, .11, .12, .13, .20, .21, .3, .40, .41, .42, .43, .8, .9 293.0, .1, .81, .82, .83, .84, .89, .89, .9 294.0, .1, .10, .11, .8, .9 307.20, .21, .22, .23, .3 310.0, .8, .9 317 318.0, .1, 2 319
Sexual disorders	302, .0, .1, .2, .3, .4, .50, .51, .52, .53, .6, .70, .71, .72, .73, .74, .75, .76, .79, .81, .82, .83, .84, .85, .89, .9
Sleep disorders	307.4, .40, .41, .42, .43, .44, .45, .46, .47, .48, .49
Postconcussion syndrome	310.2
Other mental disorders/Mental illness	292.85 299.00, .01, .10, .80, .81, .90, .91 300.9 307.0, .52, .53, .6, .7, .9 309.22 310.1 313.3, .82, .89, .9 316
Suicidal ideation	V62.84
Self-Damaging behavior	V69.8
Symptoms, signs, and ill-defined conditions, mental	797 798, .0, .1, .2, .9 799, .0, .01, .02, .1, .2, .21, .22, .23, .24, .25, .29, .3, .4, .8, .81, .82, .89, .9
Prior history of mental disorders	V11.0, .1, .2, .3, .8, .80, .9, .90 V66.3 V67.3
Indicator of impulsivity and risky behavior	V69.2, .3
Stressors/Adversities	V40.0, .00, .1, .2, .20, .3, .30, .9, .90 V61, .0, .01, .02, .03, .04, .05, .06, .07, .08, .09, .2, .20, .21, .22, .23, .24, .29, .3, .4, .41, .42, .49, .8, .9 V62, .0, .1, .1 0, .2, .20, .21, .22, .29, .3, .4, .5, .8, 80, .81, .810, .811, .812, .82, .89, .9, .90 V69.4, .5, .9
Marital problems	V61.1, .10, .11, .12

Traumatic stress

308, .0, .1, .2, .3, .4, .9

Abbreviations: NOS, Not Otherwise Specified; PTSD, posttraumatic stress disorder; ADHD, Attention Deficit-Hyperactivity Disorder.

Appendix Table 3. Medication classification¹**Psychotropic medications**

Antianxiety Agent

- Antihistamine Type
- Benzodiazepines
- Non-Benzodiazepine
- GABA Analog

Antidepressants

- MAOI SARI NaSSA NDRI or Other
- SNRI
- SSRI
- TCA and related

Sedative-hypnotic

- Benzodiazepines
- Antidepressant Type
- Others

CNS Stimulant - Armodafinil/Modafinil

ADHD Therapy

Chemical Dependency

- Narcotic or Alcohol Treatment Agents
- Smoking Deterrents and Combinations

Migraine Therapy

- Serotonin Agonists
- All other

Antianxiety and Anticonvulsant Agent – Benzodiazepines

Anticonvulsant

Bipolar Therapy Agents

Antipsychotics (Neuroleptics)

Eating Disorder Therapy

Cognitive Disorder Therapy

PTSD Therapy – Prazosin

Sedative-Hypnotic and Anticonvulsant Agent – Barbiturates

Non-narcotic and narcotic pain relievers

Analgesic – Non-narcotic

- Analgesic or Antipyretic Non-Narcotic and Combinations
- NSAID Analgesics
- Salicylate Analgesic and Salicylate Combinations

Analgesic – Narcotic

- Long-acting
- Short-acting Schedule II

Other medications

Dermatological – Isotretinoin

Abbreviations: GABA, Gamma-Aminobutyric Acid; MAOI, Monoamine Oxidase Inhibitor; SARI, Serotonin Antagonist and Reuptake Inhibitor; NaSSA, Noradrenergic and Specific Serotonergic Antidepressant; NDRI, Norepinephrine-Dopamine Reuptake Inhibitor; SNRI, Serotonin–Norepinephrine Reuptake Inhibitor; SSRI, Selective Serotonin Reuptake Inhibitor; TCAs, Tricyclic antidepressants; CNS, Central Nervous System; ADHD, attention deficit-hyperactivity disorder; NSAID, Nonsteroidal Anti-Inflammatory Drug. PTSD, posttraumatic stress disorder

¹Source: First Databank (FDB) Enhanced Therapeutic Classification System™. Reprinted with permission by First Databank, Inc. All rights reserved. ©2013. www.fdbhealth.com

Appendix Table 4. Overview of the Historical Administrative Data System (HADS) variables used to predict suicide death in the 26 weeks after specialty mental health visits among non-deployed male soldiers, organized by broad conceptual category¹

Temporal controls (18 variables)

Year, season, and weeks since the most recent outpatient visit²

Clinical factors during the index outpatient visit (46 variables)

We included information on type of treatment provider seen and primary diagnosis of the index visit (mental vs. physical disorder). We also distinguished 23 categories of mental disorder diagnoses for the index visit largely focused on aggregated ICD-9-CM codes (e.g., ADHD/learning disorders [ICD-9-CM 314.0-315.9]) and 8 additional categories of behavioral stressors (e.g., marital problems, other stressors/adversities, suicidal ideation and self-damaging behavior) (See Appendix Table 2 for a list of ICD-9-CM codes included in the 31 mental disorders). We also created several summary measures representing broader categories of the 31 diagnoses (e.g., any impulse control disorder; any substance, count of total mental disorders). Further, we distinguished physical disorders considered important for suicide: traumatic brain injury (TBI); other severe injuries (amputations, burns, sensory losses, paralysis); pain diagnoses; and sleep diagnoses (distinguishing dyssomnias and parasomnias).

We assumed there to be no missing data for all clinical factor variables. Of the 46 index visit variables, only 2 had significant bivariate associations with the outcome and were thus included in the machine learning models.

Hundreds of clinical variables were created to distinguish soldiers with medical encounters and prescriptions occurring in the year prior to the index outpatient visit. Outpatient visit variables were created for all soldiers, while inpatient admission variables were created for the subset of soldiers who had a mental disorder hospitalization in the year prior to the index outpatient visit. Parallel variables were created using three recall periods prior to the index outpatient visit: past month, past 3 months, and past 12 months. We again distinguished 23 categories of mental disorder diagnoses largely focused on aggregated ICD-9-CM codes (e.g., ADHD/learning disorders [ICD-9-CM 314.0-315.9]), 8 additional categories of behavioral stressors (e.g., marital problems, other stressors/adversities, suicidal ideation and self-damaging behavior), and a summary measure of any of these 31 diagnoses (see Appendix Table 2). These variables included yes/no variables for any visit or admission diagnoses and parallel variables coded 0-4 for number of visits and number of days in the hospital (0=None, 1=1-2, 2=3-5, 3=6-10, 4=11+). Visits and admissions for physical disorders were classified into a single physical category that included the 17 major ICD-9-CM categories (e.g., diseases of the circulatory system [ICD-9-CM 520-579]). However, we also distinguished between four particular physical disorders of interest: traumatic brain injury (TBI); other severely traumatic injuries (amputations, burns, sensory losses, paralysis); pain diagnoses; and sleep diagnoses (distinguishing dyssomnias and parasomnias). Separate summary variables were also created to reflect total number of visits and admissions. In regard to medication and polypharmacy clinical factors, National Drug Code (NDC) psychotropic medication codes were collapsed into 15 categories (e.g., antianxiety, antidepressant, antipsychotic) and 25 sub-categories (e.g., SSRI, SNRI, TCA) using the First Databank (FDB) Enhanced Therapeutic Classification System™ (www.fdbhealth.com) (Appendix Table 3) and both dichotomized (yes-no; any prescription) and categorical (number of times prescribed) variables were created. Categorical variables were coded 0-4 for number of filled prescriptions (0=None, 1=1-30, 2=31-60, 3=61-90, 4=91+). In addition to the above, variables were also created to represent (i) the ratio of number of outpatient visits per month in service, (ii) history of suicide attempts since joining the Army, and (iii) most recent score on the mental and physical components of the Military Physical Profile Serial System (i.e., PULHES).

We assumed there to be no missing data for all clinical factor variables. Of the 782 prior clinical factor variables available for soldiers with a prior inpatient hospitalization, 244 had significant bivariate associations with the outcome and were thus included in the machine learning models. Of the 536 prior clinical factor variables available for soldiers with no previous inpatient hospitalization, 178 had significant bivariate associations with the outcome and were thus included in the machine learning models.

Prior clinical factors (782 variables among those with a prior inpatient hospitalization; 536 variables among those with no prior inpatient hospitalization)

Socio-demographics (20 variables)

Socio-demographic variables included age, marital status, race-ethnicity, religion, number of dependents, and education.

Missing socio-demographic data was handled using rational imputations (see Appendix Table 6 for additional information). All 20 socio-demographic variables were included in the machine learning models.

Army career (27 variables)

Army career variables included age-at-enlistment, Armed Forces Qualification Test (AFQT) scores at enlistment, time-in-service, military occupation, rank, demotion in the past 12 months, command (the major organizational subdivision of the Army), and current deployment either to a combat zone or in direct support outside the combat zone of Operation Enduring Freedom (Afghanistan) or Operation Iraqi Freedom, total number of deployments, and past 12 month demotion.

Missing Army career data was handled using rational imputations (see Appendix Table 7 for additional information). All 27 Army career variables were included in the machine learning models.

Prior crime (67 variables)

Prior crime variables included measures representing official criminal records of perpetration and victimization. Crime records were coded using the National Corrections Reporting Program classification system and variables were created to represent major physical violent crime, minor physical violent crime, major sexual violent crime, sexual and non-sexual verbal violent crime, 5 categories of non-violent crime (reckless endangerment; possession/use of illegal substance; weapons offence; drunkenness/disorderly conduct/vagrancy; other non-violent crime), and any crime. Parallel variables were created to represent both the yes-no occurrence *and* the number of times each respondent was a perpetrator of each crime in the past 12 and 24 months. The same victimization variables were created except that the five non-violent crime categories were collapsed into a single independent variable of non-violent crime victimization. We also included among the prior crime variables two variables pertaining to failed drug tests (in the past 3 months or past 12 months).

We assumed there to be no missing data for all crime variables. Of the 67 prior crime variables, only 2 had significant bivariate associations with the outcome and were thus included in the machine learning models.

Contextual factors (39 variables)

A series of unit-level and weapons-registry variables were created as potential contextual risk factors for suicide. Unit-level variables were focused on *parent* unit characteristics (i.e., at the level of the battalion). Unit-level measures included how long the soldier had been in his/her current duty unit, whether or not the soldier's current duty unit was his/her assigned unit, the number of months in the past 12 the soldier was on duty in a unit that was not his/her assigned unit, how many different duty units the soldier had been assigned to in the past 12 months, the median length of time in the Army (in months) of all non-commissioned officers (E5-E9) and warrant/commissioned officers in the soldiers' current duty unit over the past 3 and 12 months, the median length of time in the current unit of all lower enlisted soldiers (E1-E4), non-commissioned officers, and warrant/commissioned officers in the soldier's current duty unit over the past 3 and 12 months, and the median length ever deployed of all lower enlisted soldiers (E1-E4), non-commissioned officers, and warrant/commissioned officers in the soldier's current duty unit over the past 3 and 12 months. Access-to-weapons variables included both dichotomous and count variables for the number of guns, pistols, rifles, shotguns, knives, swords, bows, and any other weapon registered by soldiers for legal possession over the past 4 years.

We assumed there to be no missing data for all contextual factor variables. Of the 39 contextual variables, none had significant bivariate associations with the outcome and all were thus excluded from the machine learning models.

¹As HADS variables were not collected for research, there was a larger amount of missing-inconsistent data than would be typical for a research dataset. However, HADS datasets were updated monthly and missing values in one month almost always appeared in earlier and/or later months, allowing rational imputations using nearest neighbor methods. The few remaining missing values were resolved by assigning missing values to the median, mode, or omitting cases from analysis.

²All models controlled for the year (2004, 2005, 2006, 2007, 2008, 2009), season (Spring, Summer, Fall, Winter) of the observation, weeks since the most recent outpatient visit, primary diagnosis of the most recent visit (mental vs. physical), and type of treatment provider seen at the most recent visit, hence why these time-related variables are labeled "core" variables.

Appendix Table 5. Percent of missing socio-demographic or imputed Army career data in the population (N=110,851,165 person-months; 1,660,835 soldiers)¹

Variable	Population person-months missing <i>before</i> rational imputation, % ¹	Population person-months missing <i>after</i> rational imputation, %	Description of how missing data was handled after rational imputations
Age	0.2	0.0	No further imputations pursued; soldiers with missing person-months after rational imputation were omitted
Marital status	None	None	There was no missing data for marital status
Gender	0.0	0.0	Remaining missing person-months were coded to male
Race/Ethnicity	1.3	0.0	Remaining missing person-months were coded to non-Hispanic White
Religion	8.4	2.4	Remaining missing person-months were coded to the "None/Unknown" category
Number of dependents	None	None	There was no missing data for number of dependents
Education	1.8	0.2	Remaining missing person-months were coded to having completed high school (enlisted soldiers) or finished College (officers)
Age of Enlistment	0.6	0.4	Remaining missing person-months were coded to the first month the soldier had administrative data available.
Deployment	None	None	There was no missing data for deployment status
Command	0.2	0.0	Remaining missing person-months were coded into the "AMC/Other/Unknown" category
Rank	0.0	None	There was no missing data for rank after rational imputations.
Time in service	0.6	0.4	Time in service was calculated by counting the number of months in service between date of enlistment and date of discharge. As described above, remaining missing person-months for age of enlistment were coded to the first month the soldier had administrative data available. There were no missing data for date of discharge.
Armed Forces Qualifications Test (AFQT) score	14.7	3.7	Remaining missing person-months were coded to the 50th percentile
Demotion	None	None	There was no missing data for demotion

¹ Rational imputations refer to filling in missing person months with backward/forward imputations and/or taking information from additional Army and Department of Defense administrative data sources.

Appendix Table 6. Person-level projection of visit-level predicted values based on the models in Tables 2-3

	Est	(se)
I. Basic statistics		
A. Total sample		
Percent of soldiers in the population who made any specialty visit	12.1	(0.1)
Mean number of visits among soldiers with any visits	6.1	(0.1)
% of person years that are within the highest-risk time interval (5 weeks) of a top 5% risk visit	0.1	(0.0)
% of person years that are within the high-risk time interval (26 weeks) of a top 5% risk visit	0.2	(0.0)
% of person years that are within the highest-risk time interval (5 weeks) of a top 15% risk visit	0.4	(0.0)
% of person years that are within the high risk-time interval (26 weeks) of a top 15% risk visit	0.7	(0.0)
B. Visits in the top 5% of risk		
Percent of patients with at least 1 visit in the top 5% of risk	4.2	(0.2)
Mean number of visits in the top 5% of risk /any	7.0	(0.2)
Mean weeks in the highest-risk time interval (5 weeks) after such visits/any	10.3	(0.4)
Mean weeks in the high-risk time interval (26 weeks) after such visits/any	18.6	(0.6)
C. Visits in any of the top 15% of risk		
Percent of patients with at least 1 visit in any of the top 15% of risk	12.5	(0.3)
Mean number of visits in the top 15% of risk /any	7.3	(0.3)
Mean weeks in the highest-risk time interval (5 weeks) after such visits/any	11.2	(0.2)
Mean weeks in the high-risk time interval (26 weeks) after such visits/any	20.7	(0.3)
II. Highest-risk time interval (5 weeks) for visits in the top 5% of risk		
A. Theoretical maximum		
Number of person-years	1,770	(--) ¹
B. Observed		
Number of person-years	573	(8.5) ²
III. High-risk time interval (26 weeks) for visits in the top 5% of risk		
A. Theoretical maximum		
Number of person-years	9,204	(--) ³
B. Observed		
Number of person-years	1,103	(12.0) ⁴
IV. Highest-risk time interval (5 weeks) for visits in the top 15% of risk		
A. Theoretical maximum		
Number of person-years	5,310	(--) ⁵
B. Observed		
Number of person-years	1,871	(15.5) ⁶
V. High-risk time interval (26 weeks) for visits in the top 15% of risk		
A. Theoretical maximum		
Number of person-years	27,613 ⁷	
B. Observed		
Number of person-years	3,657	(21.5) ⁸

¹12.1169% of soldiers made specialty visits, with a mean of 6.07 visits for each of these soldiers x 500,574 soldiers x .05 of these visits in the top 5% of risk = 18,408 visits in top 5% of risk in the population x 5 weeks of highest-risk time interval/52 weeks) = THEORETICAL UPPER LIMIT of person years in the highest-risk time interval in the top 5% of risk. This would mean you'd be monitoring 1,770 top 5% of risk patients during their highest –risk time interval for every day throughout the year.

² 0.1145% of person years are in the highest-risk time interval after a top 5% of risk visit x 500,574 = 573 person years

³18,408 visits in the top 5% of risk x 26 weeks of high-risk time interval/52 weeks= 18,408/2 = 9,204 person years)

⁴0.2203% of person years are in the high-risk time interval after the top 5% of risk visit x 500,574 = 1103 person years

⁵12.1169% of soldiers made specialty visits, with a mean of 6.07 visits for each of these soldiers x 500,574 soldiers x .15 of these visits in the top 15% of risk = 55,225 visits in top 15% of risk in the population x 5 weeks of highest-risk time interval/52 weeks = 5,310

⁶ 0.3738% of person years are in the highest-risk time interval after a top 15% risk visit x 500,574 = 1871

⁷ 55,225 visits in the top 15% of risk x 26 weeks of high-risk time interval/52 weeks=55,225/2=27,613

⁸ 0.7305% of person years are in the high-risk time interval after a top 15% risk visit x 500,574 = 3657