nature portfolio

Corresponding author(s): Claire Houtsma

Last updated by author(s): Jul 1, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☒	☐	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	SAS studio statistical software was used to analyze the data housed on secure DoD servers.
Data analysis	Code used to produce the statistical output can be located here: https://osf.io/csm59/. Because the data used in this study is not made publicly available by DoD, we cannot post all of the source code used to create the 1700+ variables used in this analysis. However we can provide detailed appendices about each of the variables and code used to produce statistical output. Data management and calculations of prevalence and AU-ROC were conducted with SAS version 9.4. All remaining analyses were conducted in R version 3.6.3. R packages used for SL algorithms are listed in Supplementary Table 4. Net benefit was calculated using a customized function.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The STARRS Historical Administrative Data Study (HADS) data used in this paper are not available for public release. DoD clearance is required to access this data. However, the Army STARRS and STARRS-LS survey data are project data (not used here) are available to qualified researchers as restricted datasets through the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan (https://www.icpsr.umich.edu/web/ICPSR/studies/35197). Access to restricted datasets requires the applicant and their institution to establish confidential agreements with ICPSR. Researchers interested in this should submit applications via ICPSR's online Restricted Contracting System. Guidelines can be found under the data and documentation tab at the above URL.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	This study used the gender data available in U.S. Army and U.S. Department of Defense administrative records. Gender is based on soldier self-report, which might be different from sex at birth. In addition to including gender as a predictor in machine learning models, this study conducted fairness analyses to examine whether the models performed similarly among women and men.
Reporting on race, ethnicity, or other socially relevant groupings	This study used race-ethnicity data available in U.S. Army and U.S. Department of Defense administrative records. Race-ethnicity is based on soldier self-report using categories specified by the Army. During analysis race-ethnicity categories were collapsed: 19.9% Non-Hispanic Black; 62.8% Non-Hispanic White; 11.2% Hispanic; 4.2% Asian; 1.0% Native American/Alaskan Native; and 0.4% Other.
Population characteristics	See below.
Recruitment	This study analyzed deidentified historical data from U.S. Army and U.S. Department of Defense administrative data systems, focusing on Regular Army soldiers who separated from uniformed service between 2010 and 2019.
Ethics oversight	The study protocol was approved by the research ethics committees of Harvard Medical School and the Henry Jackson Foundation with a waiver of informed consent due to data de-identification.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This is a quantitative study of administrative data from a longitudinal cohort of U.S. Regular Army soldiers, in which machine learning methods were used to predict death by firearms suicide following soldiers' separation from uniformed service over a 10-year period.
Research sample	This study analyzed deidentified historical data from U.S. Army and U.S. Department of Defense administrative data systems, focusing on the n=800,579 Regular Army soldier separations from uniformed service between January 1, 2010 through November 30, 2019. The sample, of which 39.3% are age 24 or younger at separation, 84.9% are male, 62.8% are Non-Hispanic White (19.9% Non-Hispanic Black; 11.2% Hispanic; 4.2% Asian; 1.0% Native American/Alaskan Native; and 0.4% other), 69.8% are high school-educated, and 55.7% were currently married (38.8% never married), is representative of separated Regular U.S. Army soldiers during the study period.
Sampling strategy	The analysis focused on separations from Regular U.S. Army soldiers in calendar years 2010-2019. A total of n=800,579 Regular Army soldier separations occurred between 2010-2019. A total of n=798,818 soldiers separated in these n=800,579 separations. The sample was analyzed as a person-month survival analysis dataset, with up to 120 person-months per separation (n=48,483,236 total person-months). The sample was divided into a random 70% of the total person-months (i.e., n=33,938,265 person-months from n=559,378 soldiers) in which the models were developed (referred to subsequently as the training sample); a random 10% (i.e., n=4,848,324 person-

months from n=79,997 soldiers) in which the models were calibrated (referred to subsequently as the calibration sample); and a remaining 20% (i.e., n=9,696,647 person-months from the remaining n=159,993 soldiers) in which model performance was evaluated (referred to subsequently as the test sample).

We created a case-control subsample of the training sample for computational efficiency that included all n=1,445 person-months with death by suicide coded 1 on the outcome plus a probability sample of 20 times as many (i.e., n=28,900) person-months selected out of the remaining n=33,936,820 person-months without death by suicide, matched on months since separation and on calendar month and year. Case person-months were replicated 20 times and each were assigned a weight of 1/20. Each of the n=28,900 control person-months was assigned a weight equal to 1 divided by their probability of selection. Weights were then normalized such that they summed to the sample size.

We created a case-control subsample of the calibration sample for computational efficiency that included all n=213 person-months with death by suicide coded 1 on the outcome plus a probability sample of 20 times as many (i.e., n=4,260) person-months selected out of the remaining n=4,848,111 person-months without death by suicide, matched on months since separation and on calendar month and year. Case person-months were assigned a weight of 1. Each of the n=4,260 control person-months was assigned a weight equal to 1 divided by their probability of selection. Weights were then normalized such that they summed to the sample size.

We created a case-control subsample of the test sample for computational efficiency that included all n=426 person-months with death by suicide coded 1 on the outcome plus a probability sample of 20 times as many (i.e., n=8,520) person-months selected out of the remaining n=9,696,221 person-months without death by suicide, matched on months since separation and on calendar month and year. Case person-months were assigned a weight of 1. Each of the n=8,520 control person-months was assigned a weight equal to 1 divided by their probability of selection. Weights were then normalized such that they summed to the sample size.

Data collection

This study analyzed deidentified historical administrative data from the Study to Assess Risk and Resilience in Servicemembers-Longitudinal Study (STARRS-LS), which includes an integrated database of U.S. Army and U.S. Department of Defense administrative data systems. Study researchers were not blinded to the outcome as suicides and suicide attempts are administratively recorded in administrative (DoDser) and death NDI data, all of which had to be cleaned from source data into the analytic file. Active duty soldiers, whether suicidal or not, were randomized to either the 70% training, 10% calibration, or 20% test samples.

Timing

The U.S. Army and U.S. Department of Defense administrative data examined in this study included regular Army soldiers who separated from uniformed service between January 1, 2010 and November 30, 2019. Predictor variables were calculated using the 5 years prior to the date of separation. Thus, data used to build predictors for these soldiers went back until January 1, 2005.

Data exclusions

No data were excluded from the analyses.

Non-participation

No participants were involved in this study, as this is exempt research on strictly de-identified observational DoD data.

Randomization

As noted above, the sample was divided into a random 70% of the total person-months (i.e., n=33,938,265 person-months from n=559,378 soldiers) in which the models were developed (i.e., the training sample); a random 10% (i.e., n=4,848,324 person-months from n=79,997 soldiers) in which the models were calibrated (i.e., the calibration sample); and a remaining 20% (i.e., n=9,696,647 person-months from the remaining n=159,993 soldiers) in which model performance was evaluated (i.e., the test sample). These three subsamples were constructed to have the same prevalence of death by suicide.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A