Please cite the Published Version

Identifying probable Post-Traumatic Stress Disorder: Applying supervised machine learning to data from a UK military cohort

Background: Early identification of probable post-traumatic stress disorder (PTSD) can lead to early intervention and treatment.

Aims: This study aimed to evaluate supervised machine learning (ML) classifiers for the identification of *probable* PTSD in those who are serving, or have recently served in the United Kingdom (UK) Armed Forces.

Methods: Supervised ML classification techniques were applied to a military cohort of 13,690 serving and ex-serving UK Armed Forces personnel to identify probable PTSD based on self-reported service exposures and a range of validated self-report measures. Data were collected between 2003 and 2009.

Results: The predictive performance of supervised ML classifiers to detect cases of probable PTSD were encouraging when compared to a validated measure, demonstrating a capability of supervised ML to detect cases of probable PTSD. It was possible to identify which variables contributed to the performance, including alcohol misuse, gender and deployment status. A satisfactory sensitivity was obtained across a range of supervised ML classifiers but specificity was low, indicating a potential for false positive diagnoses.

Conclusions: Detection of probable PTSD based on self-reported measurement data is feasible, may greatly reduce the burden on public health and improve operational efficiencies by enabling early intervention, before manifestation of symptoms.

Introduction

The mental health of any solider influences their operational effectiveness, health and wellbeing (Payne et al., 2013; Rona et al., 2007). It has been shown to be an important factor in their retention and productivity while being a member of the Armed Forces (AF; Fear et al., 2010; Hotopf et al., 2006; Rona et al., 2012). Similarly, service leavers with poor mental health are at an increased risk of dysfunction and social exclusion (Fear et al., 2010; Hotopf et al., 2006; A. Iversen et al., 2005). Exposure to traumatic events is not uncommon during military service and can contribute to the development of post-traumatic stress disorder (PTSD; Greenberg et al., 2011; Murphy et al., 2017). Previous research has shown between 7% and 30% of United Kingdom (UK) AF are likely to suffer from PTSD (Dickstein et al., 2015; Greenberg et al., 2011).

Early identification and treatment of PTSD is vital to prevent manifestation of chronic symptoms (Bryant, 2003; Karstoft et al., 2015; Kelly et al., 2012; Otis et al., 2003). Many studies have estimated the likelihood of PTSD using probabilistic and regression modelling (O'Donnell et al., 2007; Orcutt et al., 2004; Rona et al., 2012; Sterling et al., 2011), but due to large quantities and increased complexity of data collected more sophisticated modelling methods may be useful.

Machine learning (ML) classifiers generate sophisticated statistical models that describe behaviours and patterns in data at an abstract level (Koza et al., 1996; Samuel, 1959). There are multiple branches of ML; in this work we use supervised ML. Supervised ML is capable of automatically learning from data (usually a set of variables, which represent a single dimension of the data), making predictions on what has been observed (referred to as outcome in ML literature) (Gentleman et al., 2008). This is achieved by using labelled training data (referred as input data) to learn data distributions and patterns, and when presented with testing data (which does not contain a label), the ML predicts which label it believes is the outcome.

Supervised ML has been found to be helpful in the diagnosis and personalisation of physical conditions such as the monitoring of musculoskeletal impairment with age (Ahad et al., 2008; Leightley et al., 2017). In genetics, Khondoker et al. (2013) performed a comparison of ML methods for classifying Bipolar gene expression, MRI brain imaging and Electroencephalographic datasets. The authors found that ML performed well in multiple testing scenarios, when altering the parameters (used to define the complexity and structure) and obtained acceptable outcome results.

The application of ML in PTSD detection is more limited. Only one study (Karstoft et al., 2015) has used supervised ML for PTSD prediction to date. Karstoft et al. (2015) found that ML was capable of accurately detecting *probable* PTSD of a Danish military cohort. The term *probable* PTSD is

determined based on participants completing a self-report measure and is therefore not a clinical diagnosis, nevertheless it is a useful indicator of the likelihood of having PTSD (Blanchard et al., 1996; Rona et al., 2012). However, the interpretation of these findings is limited as this study used only one ML technique, and a small number of data points.

The move to supervised ML-based approaches offers many advantages in developing automated, real-time systems to identify those who may be at risk of a mental health condition without the need for continuous self-reporting (Andreu-Perez et al., 2015; Schneeweiss, 2014). The aim of this study is to use several supervised ML classifiers to predict probable PTSD in a large, representative, UK military cohort, with advancements over current methodologies and studies by including a greater number of participants, using the literature to determine variable selection and evaluating which variables which contribute most to the prediction.

Methods

Participants

This study used the King's Centre for Military Health Research (KCMHR) cohort, which is a longitudinal cohort of the UK AF (Fear et al., 2010; Hotopf et al., 2006). Data was collected in two phases. Phase 1 data collected in 2004-2006 from a sample of personnel deployed to the Iraq war during 2003 and a sample of personnel who did not deploy at that time. Phase 1 participants were then followed up in Phase 2 (2007-2009), where two new samples were introduced to represent the socio-demographics and deployment characteristics of the AF at that time. A comprehensive description of the methods and rational for the study has been previously reported (Fear et al., 2010; Hotopf et al., 2006). In total, 13,856 AF personnel provided data that was collected in either (or both) phases. The most recent record was used for each participant irrespective of phase (*e.g.* if a person is in both phases, only phase 2 data was used; if a person was in one phase, only that data was used).

KCMHR cohort received ethical approval from the Ministry of Defence research ethics committee and King's College Hospital local research ethics committee.

Self-Report Measure of Probable PTSD

Each participant was asked to complete the PTSD Checklist Civilian Version, (PCL-C; Blanchard et al., 1996), based on the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013), which consists of 17-items assessing five re-experiencing, seven avoidance and five hyper-arousal symptoms, was used to categorise cases of probable PTSD. The PCL-C was used in preference to the military version because it is less restrictive in populations that may have experienced traumatic events unrelated to deployment. There are a variety of different cut-offs used for the PCL-C, but a cut-off of 50 has been widely used in both UK (Hotopf et al., 2006) and United States (Hoge et al., 2004) military studies, it is therefore used in this study.

166 participants did not complete the PCL-C measure and were excluded, resulting in a study sample of *n*=13,690.

Model Inputs (Training)

To train each ML classifier, variables were extracted from the KCMHR cohort dataset; these were selected based on their availability in routinely collected AF data and use in similar studies (Karstoft et al., 2015).

The following is a list of descriptive data variables used: age (continuous), gender (male/female), children under 18 years (y/n), level of education attainment (categorical), marital status (categorical), smoking (y/n), consumes alcohol (y/n), service (*e.g.* RN, RM, Army, RAF; categorical), role (*e.g.* regular or reserve; categorical), combat role (y/n) and deployment zone: Afghanistan (y/n), Iraq (y/n), Bosnia &Kosovo (y/n), Bosnia-Herzegovina (y/n), Falklands (y/n), Gulf (y/n), Kosovo (y/n) and Northern Ireland (y/n).

A range of self-reported psychometric measures were included based on previous literature (A. C. Iversen et al., 2009; Jones et al., 2013; Kessler et al., 2014; O'Donnell et al., 2007; Rona et al., 2012). AF personnel completed the following questionnaires: General Health Questionnaire (GHQ-12, those scoring $\geq$ 4; binary, Goldberg et al., 1997), GHQ-12 case (y/n), Alcohol Use Disorders Identification Test score (AUDIT; Saunders et al., 1993) and AUDIT case (those scoring $\geq$ 16, binary; Fear et al., 2010).

Alongside descriptive and self-reported variables, a binary variable (0: no PTSD, 1: probable PTSD), was included for each participant (see previous section). This label allowed supervised ML to generate the training data and learn the patterns/differences between those with probable PTSD and those without.

Prediction (Testing)

Unlike model inputs (training), caseness of probable PTSD is *not* included in the testing data. Simply, for each prediction (outcome), classifiers are blinded to the correct result, where they make a prediction of the class. To assess if the outcome was correct, a comparison was made between PCL-C caseness (PCL-C score > 50) and the result of the classifier.

Evaluation

In this study, the following supervised ML classifiers were selected: Support Vector Machines (SVM; Cortes et al., 1995), Random Forests (RF; Breiman, 2001), Artificial Neural Networks (ANN; McCulloch et al., 1943) and Bagging (Breiman, 1996). SVM was selected as it had been used by Karstoft et al. (2015) in detecting probable PTSD. RF, ANN and Bagging were selected based on their popularity and proven reliability in health-related research (Khondoker et al., 2016; Leightley et al., 2017; Nitze et al., 2012).

This study employed a similar data processing approach to Karstoft et al. (2015). Prior to undertaking any analyses, variables were standardised to normalise the distribution. Any missing values in the dataset were imputed using *k*-nearest neighbours algorithm (Freidman et al., 1977; Karstoft et al.,

2015), where a missing value is replaced with a weighted mean of the k-nearest neighbours based on their Euclidean distance. This consideration is important to ensure that missing data does not bias the modelling process (Hastie et al., 1999). The proportion of missing data was <10% across the dataset.

Data resampling, variable importance and testing was performed using 100 iterations to consider variance and to assess the repeatability of the results. In each iteration, a random 80% of the data was selected for training and the remaining 20% used for testing. Since the outcome is unbalanced (3.95% are probable PTSD cases), each dataset (training and testing) contained an equal ratio of probable PTSD cases. This was performed due to the large number of non-PTSD cases (96.05%).

Parameter selection was undertaken using 10-fold cross-validation, a common *fine-tuning* technique (Stone, 1974). In this procedure, training data was randomly split into ten equally sized non-overlapping subsets with the same ratio of cases. The classifier is then trained on nine of these random subsets, and tested on the tenth subset. This process is repeated until all sub-sets have been used for testing. The outcome of this validation is a set of training parameters which were identified as the most 'suitable' to train the model to achieve the best sensitivity outcome. The most suitable parameters are used to train the final model using all the training data and evaluated on the testing dataset.

ML models are sensitive to small variations in data and it is important to be able to identify which variables contribute towards the prediction. During each iteration, stabilised variable selection using bootstrapping is employed (Bi et al., 2003). Variable selection is repeated with each iteration (defined previously). To determine the order in which variables contributed to the prediction, model outputs are evaluated, detailed description is provided in Guyon et al., 2003.

Success of each ML is reported as sensitivity, specificity and Matthews Correlation Coefficient (MCC) across all testing iterations.

Analyses and machine learning were performed using Matlab 2016a and the Image Processing Toolbox (The Mathworks Inc).

Results

The results were generated using an automated script, on a machine with the following specifications; Windows 10 operating system, i7 processor, Nvidia GeForce GTX 1050 Ti and 16GB of RAM. The total computational processing time was approximately 19 hours.

Sample demographics (*n*=13,690) are presented in Table 1.

Prediction of probable PTSD

The performance of each ML classifier is presented in Table 2. Overall performance varied between ML classifiers, RF (accuracy: 0.97, SD: 0.13, CI 0.97-0.98) provided the highest accuracy rate (alongside the largest SD); on the other hand, SVM (sensitivity: 0.70, SD: 0.08, CI 0.69-0.84) obtained the highest sensitivity rates.

Variable Importance

Of the 23 variables used for training, the five highest ranked variables over all iterations, based on the highest contribution to prediction were identified for each classifier and are presented in Table 3, with the most common variables including AUDIT score, AUDIT case and GHQ-12.

Discussion

This study examined the potential for using supervised ML to identify UK AF personnel who have probable PTSD. Overall, our results demonstrated an accuracy of 0.89 or greater for all ML classifiers evaluated. Sensitivity analyses identified that cases of probable PTSD (true positive) could be identified alongside a high specificity to detect those without a probable PTSD diagnosis. These results extend previously published works, demonstrating the capability and suitability of ML classifiers in detecting probable PTSD (Karstoft et al., 2015).

In the UK, there is a moral obligation and duty of care to service personnel, as enshrined in the Armed Forces Covenant (House of Commons Defence Committee, 2015) to its AF by the 'People of the United Kingdom' (The Ministry of Defence, 2015). This study could contribute towards meeting these obligations by providing an early diagnostic tool by predicting those who may be susceptible to a probable PTSD diagnosis, enabling early intervention and treatment; further ensuring military fitness and operational readiness (Murphy et al., 2017). Similarly, ML application in other mental health conditions (*e.g.* anxiety, depression) and occupational settings (*e.g.* police, firefighters, paramedics) should be further explored.

This study used the PCL-C to identify cases of probable PTSD; it has previously been validated against a clinical interview (Blanchard et al., 1996), and used in a number of military studies (Hoge et al., 2004; Jones et al., 2013; Rona et al., 2012). However, it has yet to be audited in a UK military sample. Stigma connected to mental health problems remains a key barrier that prevents serving and ex-serving personnel seeking help, which may result in an under-reporting and under-diagnosis of PTSD (Fear et al., 2012; Murphy et al., 2015). Solutions with a greater degree of automation could be used to identify probable PTSD without the need for AF personnel or veterans to self-disclose their health and wellbeing, improving diagnoses rates and reducing report bias. However, focus should be placed on improving help seeking, which is low, among the AF community to enable automated solutions to function correctly (Hines et al., 2014; Murphy et al., 2015; Sharp et al., 2015).

This study identified variables which contributed to the prediction of probable PTSD. It is encouraging that several of the variables – AUDIT score, consuming alcohol and age – have all previously been associated with PTSD and comorbid alcohol misuse in military cohorts (Jacobson, 2008; Wilk et al., 2010). Further, the presence of gender in two ML models may reflect that males (90% UK AF are male) within the UK AF are more likely to be defined as having probable PTSD than female counterparts based on clinical interview data (A. C. Iversen et al., 2009). On the other hand,

this has not be found when using self-reported data measured using the PCL-C (Fear et al., 2010; Hotopf et al., 2006).

The sensitivity of the classifiers in detecting probable PTSD were relatively low (range: 0.60-0.70), but in combination with high specificity (range: 0.90-0.98) this demonstrates the ability to accurately determine those without probable PTSD. The low sensitivity could reflect participants' alcohol consumption as, overall levels of hazardous alcohol consumption are high in the UK AF for both those with and without probable PTSD (Thandi et al., 2015). This could explain the low sensitivity observed, with AUDIT score and consuming alcohol both being present in the top 5 variables of importance.

Limitations

The results of this study must be considered in light of the following limitations.

First, this study utilised a previously collected dataset that was not collected for the purposes of building ML models. Therefore, biases in data collection could potentially impact ML modelling. This approach is not uncommon and previous studies using ML to examine mental health outcomes have utilised similar methodologies (Karstoft et al., 2015).

Second, self-report measures were utilised which may have introduced bias, though, these methods are commonly used in research studies investigating PTSD (Blanchard et al., 1996; Jones et al., 2013). However, one study has shown participants who complete the PCL-C report higher rates than clinical interview (McDonald et al., 2010). Future work should focus on introducing new modalities of data to present a more holistic view of the person's health and well-being – for example, 1) secondary data including the number of enemy engagements, type of engagement and intensity of engagement, 2) wearable technology (*e.g.* heart rate monitor) to monitor physical health, which has been shown to be worse in those with PTSD (Schry et al., 2015), 3) external data including General Practitioner data on AF personnel.

Third, the time of questionnaire completion and study phase was not considered; therefore, temporal factors could have impacted the findings. Further considerations should be given to the utilisation of more complex ML approaches (*e.g.* deep learning).

Finally, in this study, serving and ex-serving personnel were not differentiated in our models. Nevertheless, previous research has indicated that service leavers may be at increased risk of mental health disorders including PTSD, anxiety and depression (Pesonen et al., 2007; Rona et al., 2012).

Thus, additional research is needed to examine the potential usability of ML in predicting probable PTSD in these subgroups.

Conclusion

In the present study, we demonstrate that supervised ML methods can reliably identify probable PTSD from self-reported data from UK AF personnel. Detection of probable PTSD based on existing data is feasible, may reduce the burden on public health and improve operational efficiencies by enabling early intervention before chronic manifestation of symptoms. However, it is important to stress that it is not this study's intention to replace the clinical decision-making process or provide direct patient feedback. Further work is required to improve ML outcome using a larger cohort, comparison to an actual clinical diagnosis and increasing the number of variables which do not rely upon self-reporting. Nonetheless, this study has shown that, compared to traditional self-report questionnaire measures which often require continuous user engagement, there are clear advantages to supervised methods which use routinely collected data and can reliably perform retrospective analyses to determine probable PTSD.

<u>Declarations</u>

Ethical approval

King's Centre for Military Health Research Health and Welling cohort received ethical approval from the Ministry of Defence Research Ethics Committee (Reference: MoD(N) PREC 150/034 & MoDREC 0732/117) and King's College Hospital Local Research Ethics Committee (Protocol Number 11-03-219 & 07/Q0703/36). Participants consented for data to be used in secondary data analysis.

Consent for publication

None required.

Source code

Source code is available upon request from the corresponding author.

Table 1: Descriptive analyses of variables used in modelling with PTSD diagnoses. Probable PTSD and no PTSD determined by PCL-C caseness (score >50). n=13,690

#	Socio-demographic and military variables	Total sample, n (%)	Probable PTSD, n (%)	No PTSD, n (%)
1	Sex			
	Male	12,206 (89.16)	483 (3.96)	11,723 (96.04)
	Female	1,484 (10.84)	58 (3.91)	1,426 (96.09)
2	Service			
	Royal Navy	1,631 (11.91)	52 (3.19)	1,579 (96.81)
	Royal Marines	517 (3.78)	14 (2.71)	503 (97.29)
	Army	8,909 (65.08)	414 (4.65)	8,495 (95.35)
	Royal Air Force	2,633 (19.23)	61 (2.32)	2,572 (97.68)
3	Engagement Type			
	Regular	11,513 (84.10)	459 (3.99)	11,054 (96.01)
	Reserve	2,177 (15.90)	82 (3.77)	2,095 (96.23)
4	Combat Role			
	No	10,698 (78.14)	356 (3.33)	10,342 (96.67)
	Yes	2,992 (21.86)	185 (6.18)	2,807 (93.82)
5	Age (in years)			
	<25	2,547 (18.60)	133 (5.22)	2,414 (94.78)
	25-29	2,792 (20.39)	131 (4.69)	2,661 (95.31)
	30-34	2,319 (16.94)	89 (3.84)	2,230 (96.16)
	35-39	2,429 (17.74)	73 (3.01)	2,356 (96.99)
	40-44	1,847 (13.49)	72 (3.90)	1,775 (96.10)
	45-49	1,026 (7.49)	25 (2.44)	1,001 (97.56)
	50-54	453 (3.31)	8 (1.77)	445 (98.23)
	60+	277 (2.02)	10 (5.79)	267 (94.21)
6	Children under 18 years			
	No	5,087 (37.16)	196 (3.85)	4,891 (96.15)
	Yes	8,603 (62.84)	345 (4.01)	8,258 (95.99)
7	Educational attainment			
	No qualifications	1,554 (11.35)	85 (5.47)	1,469 (94.53)
	GCSE/NVQ level 1-2 or similar	4,972 (36.32)	237 (4.77)	4,735 (95.23)
	A-Level/NVQ level 3 or similar	4,494 (32.83)	173 (3.85)	4,321 (96.15)
	Degree/NVQ level 4-5	1,731 (12.64)	36 (2.08)	1,695 (97.92)
	Postgraduate qualifications	939 (6.86)	10 (1.06)	929 (98.94)
8	Marital Status			
	Married	7,058 (51.56)	218 (3.09)	6,840 (96.91)
	Living with partner	1,565 (11.43)	63 (4.03)	1,502 (95.97)
	Long term relationship	1,790 (13.08)	82 (4.58)	1,708 (95.42)
	Single	2,404 (17.56)	112 (4.66)	2,292 (95.34)
	Separated, Divorced and Widowed	873 (6.37)	66 (7.56)	807 (92.44)
9	Smoking			
	No	10,074 (73.59)	296 (2.94)	9,778 (97.06)

	Yes	3,616 (26.41)	245 (6.78)	3,371 (93.22)
10	Deployment – Afghanistan			
	No	10,291 (75.17)	438 (4.26)	9,853 (95.74)
	Yes	3,399 (24.83)	103 (3.03)	3,296 (96.97)
11	Deployment – Iraq			
	No	5,793 (42.32)	195 (3.37)	5,598 (96.63)
	Yes	7,897 (57.68)	346 (4.38)	7,551 (95.62)
12	Deployment – Bosnia & Kosovo			
	No	12,664 (92.51)	486 (3.84)	12,178 (96.16)
	Yes	1,026 (7.49)	55 (5.36)	971 (94.64)
13	Deployment – Bosnia-Herzegovina			
	No	11,347 (82.89)	450 (3.97)	10,897 (96.03)
	Yes	2,343 (17.11)	91 (3.88)	2,252 (96.12)
14	Deployment – Falklands			
	No	12,944 (94.55)	516 (3.99)	12,428 (96.01)
	Yes	746 (5.45)	25 (3.35)	721 (96.65)
15	Deployment – Gulf			
	No	12,109 (88.45)	477 (3.94)	11,632 (96.06)
	Yes	1,581 (11.55)	64 (4.05)	1,517 (95.95)
16	Deployment – Kosovo			
	No	11,552 (84.38)	452 (3.91)	11,100 (96.09)
	Yes	2,138 (15.62)	89 (4.16)	2,049 (95.84)
17	Northern Ireland			
	No	10,117 (73.90)	374 (3.70)	9,743 (96.30)
	Yes	3,573 (26.10)	167 (4.67)	3,406 (95.33)
18	General Health Questionnaire-12 Caseness			
	No Case	11,046 (80.69)	78 (0.71)	10,968 (99.29)
	Case	2,644 (19.31)	463 (17.51)	2,181 (82.49)
19	General Health Questionnaire-12 Scoring[1]			
	≤4	12,236 (89.38)	239 (1.95)	11,997 (98.05)
	5-8	993 (7.26)	126 (12.68)	867 (87.32)
	9-12	461 (3.36)	176 (38.17)	285 (61.83)
20	AUDIT caseness			
	No case	11,725 (85.65)	314 (2.68)	11,411 (97.32)
	Case	1,965 (14.35)	227 (11.55)	1,738 (88.45)
21	Consumes Alcohol			
	No	429 (3.13)	29 (6.76)	400 (93.24)
	Yes	13,261 (96.87)	512 (3.86)	12,749

Detecting PTSD with Machine Learning

22	AUDIT Score[1]			(96.14)
	≤7	5,618 (41.04)	142 (2.53)	5,476 (97.47)
	8-15	6,107 (44.61)	172 (2.82)	5,935 (97.18)
	16-19	1,050 (7.67)	62 (5.90)	988 (94.10)
	20-40	915 (6.68)	165 (18.03)	750 (81.97)

n, number; %, percentage.

[1]To improve readability, score is presented in groups. A full breakdown can be requested from the corresponding author.

Table 2: Overall results for each machine learning classifier with Standard deviation and 95% confidence intervals

	Accuracy (SD, 95% CI)	Sensitivity (SD, 95% CI)	Specificity (SD, 95% CI)	Matthews Coefficient Correlation (SD, 95% CI)
Support Vector Machines (SVM)	0.91 (0.10, 0.91-0.93)	0.70 (0.08, 0.69-0.84)	0.92 (0.05, 0.92-0.93)	0.74 (0.07, 0.69-0.78)
Random Forests (RF)	0.97 (0.13, 0.97-0.98)	0.60 (0.20, 0.59-0.85)	0.98 (0.07, 0.97-0.98)	0.64 (0.17, 0.53-0.75)
Artificial Neural Networks (ANN)	0.89 (0.10, 0.88-0.92)	0.61 (0.21, 0.48-0.74)	0.92 (0.08, 0.91-0.93)	0.45 (0.16, 0.35-0.56)
Bagging	0.95 (0.11, 0.95-0.96)	0.69 (0.12, 0.61-0.77)	0.96 (0.12, 0.96-0.97)	0.55 (0.10, 0.49-0.61)

Table 3: Variable ranking, based on contribution to prediction, for each machine learning classifier.

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
Support Vector Machines (SVM)	AUDIT Score	GHQ-12	Age	Consumes alcohol	Deployment to Iraq
Random Forests (RF)	Gender	AUDIT score	GHQ-12	Service Type	Deployment to Iraq
Artificial Neural Networks (ANN)	GHQ-12	AUDIT case	AUDIT score	Consumes alcohol	Gender
Bagging	Age	Consumes alcohol	Smoking	GHQ-12	AUDIT score

Bibliography

Ahad, M. A. R., Tan, J. K., Kim, H. S., & Ishikawa, S. (2008). Human activity recognition: Various paradigms. In *2008 International Conference on Control, Automation and Systems* (pp. 1896–1901). COEX, Seoul, Korea: IEEE. https://doi.org/10.1109/ICCAS.2008.4694407

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC.

Andreu-Perez, J., Poon, C. C. Y., Merrifield, R. D., Wong, S. T. C., & Yang, G.-Z. (2015). Big Data for Health. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1193–1208. https://doi.org/10.1109/JBHI.2015.2450362

Bi, J., Bennett, K., Embrechts, M., Breneman, C., & Song, M. (2003). Dimensionality Reduction via Sparse Support Vector Machines. *J. Mach. Learn. Res.*, *3*, 1229–1243. Retrieved from http://dl.acm.org/citation.cfm?id=944919.944971

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, *34*(8), 669–673. https://doi.org/10.1016/0005-7967(96)00033-2

Breiman, L. (1996). No Title. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1023/A:1018054314350

Breiman, L. (2001). No Title. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bryant, R. A. (2003). Early predictors of posttraumatic stress disorder. *Biological Psychiatry*, *53*(9), 789–795. https://doi.org/10.1016/S0006-3223(02)01895-4

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

Dickstein, B. D., Weathers, F. W., Angkaw, A. C., Nievergelt, C. M., Yurgil, K., Nash, W. P., … Litz, B. T. (2015). Diagnostic Utility of the Posttraumatic Stress Disorder (PTSD) Checklist for Identifying Full and Partial PTSD in Active-Duty Military. *Assessment*, *22*(3), 289–297. https://doi.org/10.1177/1073191114548683

Fear, N. T., Jones, M., Murphy, D., Hull, L., Iversen, A. C., Coker, B., … Wessely, S. (2010). What are the consequences of deployment to Iraq and Afghanistan on the mental health of the UK armed forces? A cohort study. *The Lancet*, *375*(9728), 1783–1797. https://doi.org/10.1016/S0140-6736(10)60672-1

Fear, N. T., Seddon, R., Jones, N., Greenberg, N., & Wessely, S. (2012). Does anonymity increase the reporting of mental health symptoms? *BMC Public Health*, *12*(1), 797. https://doi.org/10.1186/1471-2458-12-797

Freidman, J. H., Bentley, J. L., & Finkel, R. A. (1977). An Algorithm for Finding Best Matches in Logarithmic Expected Time. *ACM Transactions on Mathematical Software*, *3*(3), 209–226. https://doi.org/10.1145/355744.355745

Gentleman, R., Huber, W., & Carey, V. J. (2008). Supervised Machine Learning. In *Bioconductor Case Studies* (pp. 121–136). New York, NY: Springer New York. https://doi.org/10.1007/978-0-387-77240-0_9

Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, *27*(1), S0033291796004242. https://doi.org/10.1017/S0033291796004242

Greenberg, N., Langston, V., Iversen, A. C., & Wessely, S. (2011). The acceptability of "Trauma Risk Management" within the UK Armed Forces. *Occupational Medicine*, *61*(3), 184–189. https://doi.org/10.1093/occmed/kqr022

Guyon, I., & Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.*, *3*, 1157–1182. Retrieved from http://dl.acm.org/citation.cfm?id=944919.944968

Hastie, T., Tibshirani, R., Sherlock, G., Eisen, M., Brown, P., & Botstein, D. (1999). *Imputing missing data for gene expression arrays*.

Hines, L. A., Goodwin, L., Jones, M., Hull, L., Wessely, S., Fear, N. T., & Rona, R. J. (2014). Factors Affecting Help Seeking for Mental Health Problems After Deployment to Iraq and Afghanistan. *Psychiatric Services*, *65*(1), 98–105. https://doi.org/10.1176/appi.ps.004972012

Hoge, C. W., Castro, C. A., Messer, S. C., McGurk, D., Cotting, D. I., & Koffman, R. L. (2004). Combat Duty in Iraq and Afghanistan, Mental Health Problems, and Barriers to Care. *New England Journal of Medicine*, *351*(1), 13–22. https://doi.org/10.1056/NEJMoa040603

Hotopf, M., Hull, L., Fear, N. T., Browne, T., Horn, O., Iversen, A., … Wessely, S. (2006). The health of UK military personnel who deployed to the 2003 Iraq war: a cohort study. *The Lancet*, *367*(9524), 1731–1741. https://doi.org/10.1016/S0140-6736(06)68662-5

House of Commons Defence Committee. (2015). *The Armed Forces Covenant in Action Part 5: Military Casualties , a review of progress*. London. Retrieved from https://publications.parliament.uk/pa/cm201415/cmselect/cmdfence/527/52702.htm

Iversen, A. C., van Staden, L., Hughes, J. H., Browne, T., Hull, L., Hall, J., … Fear, N. T. (2009). The prevalence of common mental disorders and PTSD in the UK military: using data from a clinical interview-based study. *BMC Psychiatry*, *9*(1), 68. https://doi.org/10.1186/1471-244X-9-68

Iversen, A., Nikolaou, V., Greenberg, N., Unwin, C., Hull, L., Hotopf, M., … Wessely, S. (2005). What happens to British veterans when they leave the armed forces? *European Journal of Public Health*, *15*(2), 175–184. https://doi.org/10.1093/eurpub/cki128

Jacobson, I. G. (2008). Alcohol Use and Alcohol-Related Problems Before and After Military Combat Deployment. *JAMA*, *300*(6), 663. https://doi.org/10.1001/jama.300.6.663

Jones, M., Sundin, J., Goodwin, L., Hull, L., Fear, N. T., Wessely, S., & Rona, R. J. (2013). What explains post-traumatic stress disorder (PTSD) in UK service personnel: deployment or something else? *Psychological Medicine*, *43*(8), 1703–1712. https://doi.org/10.1017/S0033291712002619

Karstoft, K.-I., Statnikov, A., Andersen, S. B., Madsen, T., & Galatzer-Levy, I. R. (2015). Early identification of posttraumatic stress following military deployment: Application of machine learning methods to a prospective study of Danish soldiers. *Journal of Affective Disorders*, *184*, 170–175. https://doi.org/10.1016/j.jad.2015.05.057

Kelly, J., Gooding, P., Pratt, D., Ainsworth, J., Welford, M., & Tarrier, N. (2012). Intelligent real-time therapy: Harnessing the power of machine learning to optimise the delivery of momentary cognitive–behavioural interventions. *Journal of Mental Health*, *21*(4), 404–414. https://doi.org/10.3109/09638237.2011.638001

Kessler, R. C., Rose, S., Koenen, K. C., Karam, E. G., Stang, P. E., Stein, D. J., … Carmen Viana, M. (2014). How well can post-traumatic stress disorder be predicted from pre-trauma risk factors? An exploratory study in the WHO World Mental Health Surveys. *World Psychiatry*, *13*(3), 265–274. https://doi.org/10.1002/wps.20150

Khondoker, M., Dobson, R., Skirrow, C., Simmons, A., & Stahl, D. (2016). A comparison of machine learning methods for classification using simulation with multiple real data examples from mental health studies. *Statistical Methods in Medical Research*, *25*(5), 1804–1823. https://doi.org/10.1177/0962280213502437

Koza, J. R., Bennett, F. H., Andre, D., & Keane, M. A. (1996). Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. In *Artificial Intelligence in Design '96* (pp. 151–170). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-0279-4_9

Leightley, D., McPhee, J. S., & Yap, M. H. (2017). Automated Analysis and Quantification of Human Mobility Using a Depth Sensor. *IEEE Journal of Biomedical and Health Informatics*, *21*(4), 939–948. https://doi.org/10.1109/JBHI.2016.2558540

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115–133. https://doi.org/10.1007/BF02478259

McDonald, S. D., & Calhoun, P. S. (2010). The diagnostic accuracy of the PTSD Checklist: A critical review. *Clinical Psychology Review*, *30*(8), 976–987. https://doi.org/10.1016/j.cpr.2010.06.012

Murphy, D., Ashwick, R., Palmer, E., & Busuttil, W. (2017). Describing the profile of a population of UK veterans seeking support for mental health difficulties. *Journal of Mental Health*, *0*(0), 1–8. https://doi.org/10.1080/09638237.2017.1385739

Murphy, D., & Busuttil, W. (2015). PTSD, stigma and barriers to help-seeking within the UK Armed Forces. *Journal of the Royal Army Medical Corps*, *161*(4), 322–326. https://doi.org/10.1136/jramc-2014-000344

Nitze, I., Schulthess, U., & Asche, H. (2012). Comparison of Machine Learning Algorithms Random Forest, Artificial Neural Network and Support Vector Machine to Maximum Likelihood For Supervised Crop Type Classification. In *GEOBIA* (pp. 35–40). Rio de Janeiro.

O'Donnell, M. L., Elliott, P., Lau, W., & Creamer, M. (2007). PTSD symptom trajectories: From early to chronic response. *Behaviour Research and Therapy*, *45*(3), 601–606. https://doi.org/10.1016/j.brat.2006.03.015

Orcutt, H. K., Erickson, D. J., & Wolfe, J. (2004). The course of PTSD symptoms among Gulf War veterans: A growth mixture modeling approach. *Journal of Traumatic Stress*, *17*(3), 195–202. https://doi.org/10.1023/B:JOTS.0000029262.42865.c2

Otis, J. D., Keane, T. M., & Kerns, R. D. (2003). An examination of the relationship between chronic pain and post-traumatic stress disorder. *Journal of Rehabilitation Research and Development*, 397–405.

Payne, R. A., Abel, G. A., Guthrie, B., & Mercer, S. W. (2013). The effect of physical multimorbidity, mental health conditions and socioeconomic deprivation on unplanned admissions to hospital: a retrospective cohort study. *Canadian Medical Association Journal*, *185*(5), E221–E228. https://doi.org/10.1503/cmaj.121349

Pesonen, A.-K., Raikkonen, K., Heinonen, K., Kajantie, E., Forsen, T., & Eriksson, J. G. (2007).

Depressive Symptoms in Adults Separated from Their Parents as Children: A Natural Experiment during World War II. *American Journal of Epidemiology*, *166*(10), 1126–1133. https://doi.org/10.1093/aje/kwm254

Rona, R. J., Fear, N. T., Hull, L., Greenberg, N., Earnshaw, M., Hotopf, M., & Wessely, S. (2007). Mental health consequences of overstretch in the UK armed forces: first phase of a cohort study. *BMJ*, *335*(7620), 603–603. https://doi.org/10.1136/bmj.39274.585752.BE

Rona, R. J., Jones, M., Sundin, J., Goodwin, L., Hull, L., Wessely, S., & Fear, N. T. (2012). Predicting persistent posttraumatic stress disorder (PTSD) in UK military personnel who served in Iraq: A longitudinal study. *Journal of Psychiatric Research*, *46*(9), 1191–1198. https://doi.org/10.1016/j.jpsychires.2012.05.009

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. https://doi.org/10.1147/rd.33.0210

Saunders, J. B., Aasland, O. G., Babor, T. F., De La Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction*, *88*(6), 791–804. https://doi.org/10.1111/j.1360-0443.1993.tb02093.x

Schneeweiss, S. (2014). Learning from Big Health Care Data. *New England Journal of Medicine*, *370*(23), 2161–2163. https://doi.org/10.1056/NEJMp1401111

Schry, A. R., Rissling, M. B., Gentes, E. L., Beckham, J. C., Kudler, H. S., Straits-Tröster, K., & Calhoun, P. S. (2015). The Relationship Between Posttraumatic Stress Symptoms and Physical Health in a Survey of U.S. Veterans of the Iraq and Afghanistan Era. *Psychosomatics*, *56*(6), 674–684. https://doi.org/https://doi.org/10.1016/j.psym.2015.07.010

Sharp, M. L., Fear, N. T., Rona, R. J., Wessely, S., Greenberg, N., Jones, N., & Goodwin, L. (2015). Stigma as a barrier to seeking health care among military personnel with mental health problems. *Epidemiologic Reviews*, *37*(1), 144–162. https://doi.org/10.1093/epirev/mxu012

Sterling, M., Hendrikz, J., & Kenardy, J. (2011). Similar factors predict disability and posttraumatic stress disorder trajectories after whiplash injury. *Pain*, *152*(6), 1272–1278. https://doi.org/10.1016/j.pain.2011.01.056

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*, *36*(2), 111–147.

Thandi, G., Sundin, J., Ng-Knight, T., Jones, M., Hull, L., Jones, N., … Fear, N. T. (2015). Alcohol misuse in the United Kingdom Armed Forces: A longitudinal study. *Drug and Alcohol Dependence*, *156*, 78–83. https://doi.org/10.1016/j.drugalcdep.2015.08.033

The Ministry of Defence. (2015). *THE ARMED FORCES COVENANT*. *The Ministery of Defence*. Ministry of Defense. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/49469/the_armed_forces_covenant.pdf

Wilk, J. E., Bliese, P. D., Kim, P. Y., Thomas, J. L., McGurk, D., & Hoge, C. W. (2010). Relationship of combat experiences to alcohol misuse among U.S. soldiers returning from the Iraq war. *Drug and Alcohol Dependence*, *108*(1–2), 115–121. https://doi.org/10.1016/j.drugalcdep.2009.12.003