Brief Report

Journal of Concussion
Volume 5: 1–9
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20597002211006585
journals.sagepub.com/home/ccn
SAGE

Prolonged concussion effects: Constellations of cognitive deficits detected up to year after injury

Daniel D Leeds[1], Annie Nguyen[1], Christopher D'Lauro[2], Jonathan C Jackson[3] and Brian R Johnson[4]

Abstract
Concussions are associated with an array of physical, emotional, cognitive, and sleep symptoms at multiple timescales. Cognitive recovery occurs relatively quickly – five-to-seven days on average. Yet, recent evidence suggests that some neurophysiological changes can be identified one year after a concussion. To that end, we examine more nuanced patterns in cognitive tests to determine whether cognitive abilities could identify a concussion within one-year post injury. A radial-basis (non-linear boundary) support vector machine classifier was trained to use cognitive performance measures to distinguish participants with no prior concussion from participants with prior concussion in the past year. After incorporating only 10 cognitive measures, or all 5 composite measures from the neurocognitive assessment (Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT)), over 90% accuracy was achieved in identifying both participants without prior concussions and participants with concussions in the past year, particularly when relying on non-linear patterns. Notably, classification accuracy stayed relatively constant between participants who had a concussion early or late in the one-year window. Thus, with substantial accuracy, a prior concussion can be identified using a non-linear combination of cognitive measures. Cognitive effects from concussion linger one-year post-injury, indicating the importance of continuing to follow concussion patients for many months after recovery and to take special note of constellations of cognitive abilities.

Keywords
Concussion, cognitive testing, machine learning

Date received: 23 January 2021; accepted: 11 March 2021

Introduction

Concussions are a growing public health issue. From 2006–2014, the Center for Disease Control (CDC) reported a 53% increase in visits to the emergency department, hospitalization, and death due to TBI-related injury.[1,2,3,4] To assist healthcare practitioners, a variety of assessments have been developed to measure the extent of disruption caused by the TBI injury.

Computer-based neurocognitive assessments have become common to detect concussion onset and to determine concussion recovery.[5] While typical cognitive recovery tends to be of short duration, long term effects of prior concussions have been reported well beyond typical recovery windows.[6,7] In the present work, we use a standard machine learning technique to identify the presence of concussion-based changes to standard cognitive factors up to one year after onset, which can aid in monitoring athlete's cognitive health.

[1]Computer and Information Sciences Department, Fordham University, Bronx, NY, USA
[2]Department of Behavioral Science and Leadership, United States Air Force Academy, USAF, CO, USA
[3]Sports Medicine Section 10th Medical Group, United States Air Force Academy, USAF, CO, USA
[4]Center for Military Psychiatry and Neuroscience, Walter Reed Army Institute of Research, Silver Spring, MD, USA

Corresponding author:
Daniel D Leeds, John Mulcahy Hall, Room 332, 441 East Fordham Road, Bronx, NY 10458, USA.
Email: dleeds@fordham.edu

Concussions are associated with an array of cognitive physical, emotional, and sleep symptoms[8] and these symptoms do not all heal at the same rate. To assess cognitive recovery, neurocognitive assessments are intended to provide objective data regarding cognitive performance in order to assist healthcare providers in making decisions on an athlete's brain health.[9,10]

Given the value of computer-based instruments for assessing baseline and post-concussion cognitive function, several states have passed laws mandating neuropsychological testing for various levels of athletes.[11] These tests tend to be relatively stable measures.[6] Further, they are self-sustaining; the administration of instruction, analysis of data, and interpretation of results are all provided[12] and can be administered simultaneously to large groups. Data collection is precise and objective and data storage/retrieval is easy.[10,13] Immediate Post-concussion Assessment and Cognitive Testing (ImPACT), in particular, is used in thousands of high schools and universities within the U.S.[14]

Typically, cognitive recovery from a concussion is reported to be of short duration. Average cognitive recovery from a concussion for high school and college athletes is five-to-seven days.[1] With that said, some concussion sufferers have lingering post-concussion symptoms. A prior concussion(s) predisposes patients to repeat concussion injury with longer recovery trajectories.[6] Further, the concussion inflammatory cascade may lead to elevated biomarkers at 30 days after concussion[7]; well beyond the typical recovery window. Thus, after apparent return to baseline cognitive performance, effects remain from concussion injury, including susceptibility to future concussions.

In the present study we examine more nuanced patterns in recorded cognitive tests to assess the effects of a concussion within one-year post injury. Rather than look forward and assess whether a specific symptom (s) or cognitive test score predicts a rapid or prolonged recovery, our goal was to look backward and determine whether test scores predict a prior concussion. We use support vector machines[3] to identify potentially complex rules for detection of prior concussions. A diverse collection of recent studies has used machine learning methods similarly to detect concussions based on symptoms and neuroimaging.[15–18]

At the U.S. Air Force Academy (hereafter referred to as "Academy" for short), the standard of care was adjusted from an annual neurocognitive baseline for all collegiate athletes to an annual neurocognitive baseline for all cadets. This adjustment was made, in part, because all cadets participate in athletics and military training events. Thus, an entire class of rising college sophomore cadets complete a baseline ImPACT as part of their pre-academic year annual medical evaluation. Our goal is to use their current baseline neurocognitive performance to determine whether we can accurately predict whether a prior concussion had occurred during the freshman year, i.e. the prior 12 months. Given the change to the standard of care this was a unique opportunity to detect whether factors remain elevated despite a complete clinically-defined concussion recovery in a large sample. In total there were 994 participants, of which 186 self-reported a prior concussion, and 55 cadets self-reported a recent concussion that occurred during the past 12 months, i.e. their freshman year at the Academy. Thus, this is a unique opportunity to detect factors that remain elevated despite a complete clinically-defined concussion recovery.

Methods

This following study protocol was deemed Exempt, Category IV by the U.S. Air Force Academy Institutional Review Board.

Participants

There were 994 participants. Of these, 217 identified themselves as female and 777 identified themselves as male. All participants were of the same academic year; they were rising sophomores. The average age was 19.30 years (range 18–23 years).

This year group was of interest because the prior freshman year at the Academy is particularly challenging. All freshman cadets are required to complete a military bootcamp – a six week physically demanding program that includes obstacle courses and pugil stick competition. During their freshman year cadets take a mandatory boxing class. Finally, the freshman year concludes with an additional intense military training event called "Recognition". Thus, freshman cadets participate in several military training events.

Apparatus

Neurocognitive performance was measured using the online full-version form of ImPACT which consists of six categories of cognitive measures.[19] This assessment takes approximately 25 minutes to complete. Each measure is briefly described below. Prior ImPACT results that cadets may have had from high school athletics were not included. Participants additionally self-reported the date of their last concussion, if any, as part of the survey accompanying the ImPACT assessment.

Cognitive measures. The Word Memory task evaluates attentional resources and verbal recognition memory. The task is to memorize and recall a set of words. The Design Memory task evaluates attentional processes and visual recognition memory. The task is to memorize and recall visual patterns. The X's and O's task assesses visual working memory, visual processing, and visual motor speed. The task is to memorize and recall the location of letters after a distractor task. The Symbol Match task assesses visual processing speed, learning, and memory. The task is to memorize and recall common symbols. The Color Match task assesses reaction time as well as impulse control/response inhibition. It is a variant of the Stroop effect.[20] An example task would be to select the word "green" when it is presented in green font color and ignore it when the color does not match the text. Finally, the Three Letters task assesses working memory and visual-motor response speed. The task is to memorize and recall random consonant letters after a distractor task. The ImPACT assessment recorded 35 cognitive performance measures and 5 composite scores for a total of 40 measures.

Procedure

Cadets completed a baseline ImPACT in June or July 2015. It was administered in the summer prior to their sophomore year. The assessment was administered in a computer lab.

Patient involvement

Patient and members of the public were not involved in the design, management, and conduct of this experiment.

Concussion history

A total of 800 cadets reported no prior concussion history; 194 cadets reported history of at least one prior concussion. There were eight data entry errors. (Cadets who reported no concussion history, but who entered a "last concussion date.") Of the remaining 186 cases of prior concussion, recent concussions, $N = 56$, were categorized as occurring within 12 months of the administered test date. Cadets whose latest concussions were more than 12 months prior to their test dates had sustained concussions spreading widely over time; there were not enough cadets in each year time window for classifier learning to be performed at the level of the one-year concussion group studied here. Thus, only recent "past 12 month" concussions were considered.

Analysis

We trained a support vector machine[3] classifier to distinguish participants with prior concussions from participants with no prior concussions based on ImPACT neurocognitive measures. Across all analyses, we used a 10-fold cross validation.

We used a support vector machine (SVM) with radial basis kernel, to capture a potentially non-linear boundary between participants with and without a prior concussion(s). Prior to learning, cognitive performance measures (features) were normalized to have zero mean and unit variance. The 'KernelScale' parameter was set to 'auto' which selects an appropriate scale factor using a heuristic procedure, following default function settings in the SVM package and standard practice in the field. All classification analyses were performed in MATLAB®.[21]

For 10-fold cross validation, the total data set for each analysis was split into ten subsets. During each "fold," one subset was used as a test set, and the other nine subsets were used to form the training set. As there were only 56 reported prior concussion injuries, the data was shuffled to ensure baseline ImPACT assessments from a minimum of five different prior concussion injuries were in each of the ten testing sets. This ensured a minimum number of cadets with prior concussion injury could be properly detected in each fold of analysis. A minimum of five baseline ImPACT assessments completed by cadets without prior concussion injury were included for each testing set as well.

Greedy forward feature selection was used to identify the best cognitive performance measures to distinguish participants with prior concussion from participants without reported injury.[22] The ImPACT assessment recorded 35 cognitive performance measures and 5 composite scores for a total of 40 measures. SVMs are trained on each of the 40 measures individually, resulting in 40 corresponding accuracy measures. The greedy forward feature selection method is used by selecting the best performing measure on the highest averaged "prior concussion" (subject with report of concussion in the past year) and "no prior concussion" (cadet with report of no prior concussion) prediction accuracy and was used as the first measure in classification. This process repeated for the remaining measures, adding the single measure that led to the best improvement in SVM performance when added to the list of measures found previously.

Data set resampling

The original data set is comprised of 56 prior concussion participants and 800 no prior concussion participants, omitting participants whose last concussion occurred over one year prior to impact. The substantially lopsided ratio of prior concussion participants to no prior concussion participants risks large errors in classifier learning, as the classifier will be biased to predict the more common group and ignore the less common group. For more effective machine learning, we generate synthetic new concussion participants to balance the two groups.

Synthetic Minority Oversampling Technique (SMOTE) increases the number of occurrences of the minority group by generating additional synthetic data points.[23] SMOTE, as defined in MATLAB, takes r feature vectors (in the present study, $r = 56$ concussion participants), each with dimension n (in the present study, $n = 40$ ImPACT cognitive measures). A random point is selected from among the feature vectors and k nearest neighbors are found for this point (in the present study, $k = 14$, consistent with the 14-fold increase in the concussion data set to set equal the number of concussion and non-concussion subjects). One of these nearest neighbors is randomly selected, and a line is drawn between these two selected feature vectors. A random location along this line is selected to define a new synthetic data point. This process is repeated $k–1$ times for each feature vector, increasing the number of data points by $(k–1) \times r$. SMOTE returns a new feature vector with dimension ($\grave{r}$, n) and its corresponding labels ($\grave{r}$, 1).

The "SMOTE Data" is comprised of 784 prior concussion participants and 784 no prior concussion participants. The 784 no prior concussion participants contain the 56 non-concussion participants included in the down-sampled original data set above, as well as 728 additional no prior concussion participants randomly taken from the original 800 no prior concussion subjects. When training and testing on the SMOTE Data set, both original and synthetic participants were used for learning in each fold, but only the original participants were used to compute testing accuracies. Due to the substantial imbalance in concussion and non-concussion group sizes, accuracies for prior concussion and no prior concussion groups were computed and reported separately, as in[24]; the two accuracies were averaged with equal weight. We further demonstrate these re-balanced accuracies are comparable to precision, recall, and F-score statistics.[25]

Results

The distribution of "recent" concussions that occurred in the past 12 months is presented in Figure 1. No concussions were reported in the most recent month, but every other previous month had at least two concussions. The peaks in Figure 1 at month four and month 12 are aligned with military training events that occur during those months that may exposure participants to a higher risk of concussion. For analyses, prior concussions were grouped into four three-month intervals. As seen in Figure 1, there were eight concussions reported within the first three months from the test date. Within 4–6 months of testing 16 concussions were reported. Within 7–9 months of testing there were 13 reported concussions. Finally, 19 concussions from ten to twelve months of testing, with the majority focused on the month furthest from testing.

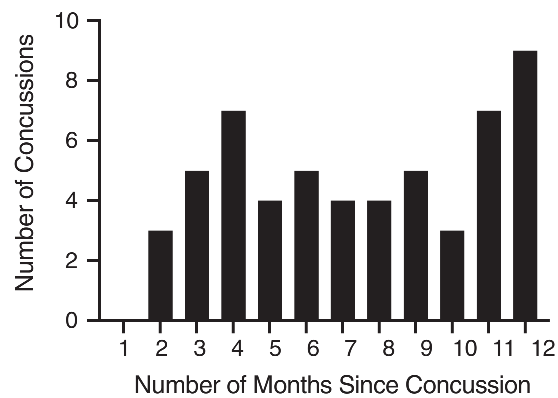

Figure 1. Distribution of reported concussions in 12 months prior to the administered test date. Most recent concussion recorded for each candidate.

The ImPACT assessment recorded 35 cognitive performance measures and 5 composite scores for a total of 40 measures.[19] The cognitive performance measures are organized into six categories: Word Memory, Design Memory, X's and O's, Symbol Match, Color Match, and Three Letters. Figure 2 shows high correlations (>0.6) among tests of the same category, particularly for Design Memory and Word Memory, with a smaller subset of high correlation measures for Three Letters. The Symbol Match and X's and O's Tasks had few correlations above 0.6 within task measures, though X's and O's correlated with several composite scores. The first performance measure for Color Match ("Total Correct") was identical for all cadets studied, leading to no correlation information for this measure (blue row/column in Color Match).

SMOTE data provides over 95% accuracy for both prior concussion and no-prior-concussion classification (Figure 3(b) to (d)). Only a subset of ImPACT cognitive measures were needed to determine concussion history. Learning on the SMOTE data set stabilizes after the first 5–10 features (cognitive measures) are added under greedy, composite-score only, and randomized learning conditions. Even using 5 features, mean classification accuracy reaches 70% with greedy feature selection (chance accuracy is 50%). Using greedy feature selection, the first 5 features are drawn from four distinct cognitive tasks (Table 1). The next five features (reaching 95% classification accuracy) add one additional task and one composite score (visual memory). Within the first 10 features, only two pairs of features have correlations above $r > 0.6$. Surprisingly, a randomized draw of features substantially outperforms

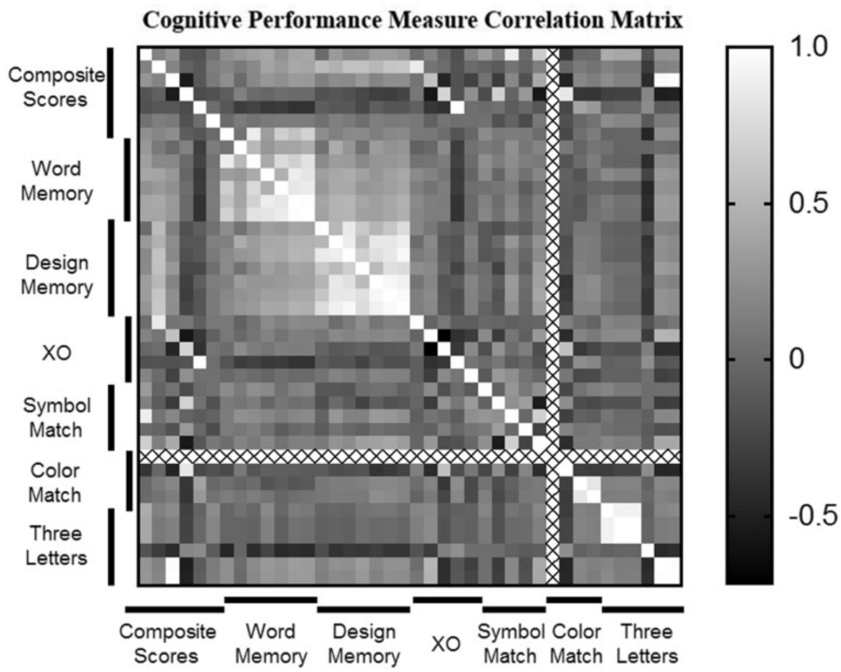

Figure 2. Correlation matrix among 40 cognitive measures ordered by high correlation.

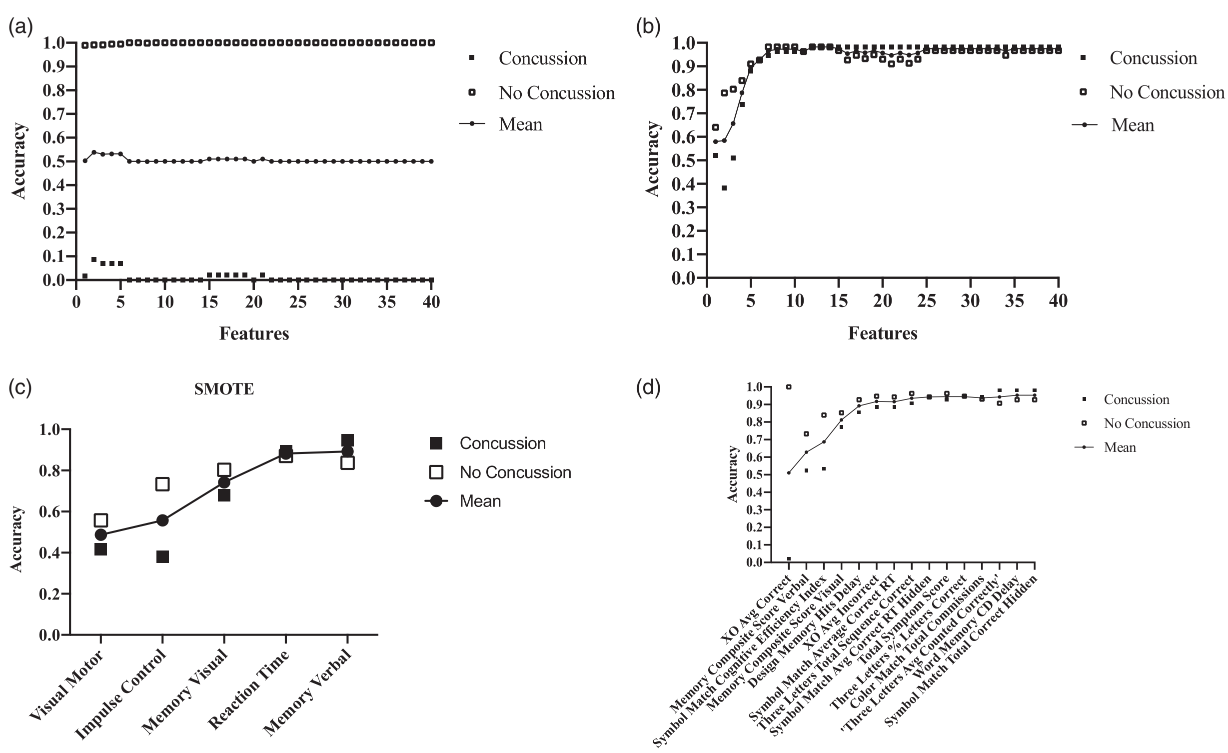

Figure 3. SVM prediction accuracies on prior concussion (black square filled) and no-prior-concussion (black square unfilled) cadets after each step of greedy feature selection using (a) original data set, (b) SMOTE data set, and (c) SMOTE data set using only composite scores. (d) Shows accuracies on SMOTE data set with one run of random feature selection.

Table 1. First 15 cognitive performance measures listed in order of selection using (left) greedy feature selection, (center) random feature selection, and (right) composite-only features from SMOTE data.

	Greedy		Random		Composite
1	"Design Mem DM Corr"	1	"XO Ave Corr"	1	"Vis Mot"
2	Symb Match Cog Effic Idx	2	"Mem Comp Score Verb"	2	"Impulse Ctrl"
3	"XO Total Corr Interfer"	3	Symb Match Cog Effic Idx	3	"Mem Vis"
4	"Color Match Tot Corr"	4	"Mem Compos Score Vis"	4	"React Time"
5	"Design Mem CD"	5	"Design Mem Hits Delay"	5	"Mem Verb"
6	"XO Ave Incorr"	6	"XO Ave Incorr"		
7	Symb Match Tot Corr Vis	7	"Symb Match Ave Corr RT"		
8	"Mem Comp Score Verb"	8	"3 Lett Tot Seq Corr"		
9	"Word Mem Hits"	9	"Symb Match Ave Corr RT Hidd"		
10	"XO Tot Corr Mem"	10	"Tot Symptom Score"		
11	"Mem Comp Score Vis"	11	"3 Lett % Lett Corr"		
12	"Word Mem Tot % Corr"	12	"Color Match Tot Commiss"		
13	"Design Mem Hits Delay"	13	"3 Lett Ave Count Corr"		
14	"Design Mem Hits"	14	"Word Mem CD Delay"		
15	"Design Mem Tot % Corr"	15	"Symb Match Tot Corr Hidd"		

greedy selection (Figure 3(d)). The first five randomly selected features achieve 89% accuracy, compared to 70% for greedy, and the first 10 randomly selected features achieve 95% accuracy, on par with greedy selection. The features span four tasks and the two memory composite scores (Table 1). Using only the five composite scores achieves 88% classification accuracy on SMOTE data (Figure 3(c)). In all three analyses, a moderate amount of feature diversity produces classification benefits.

Use of precision, recall, and F-score measures for concussion classification similarly show a strong classifier performance with only a few features (Table 2). All measures mirror group-wise accuracies in Figure 3, as complementary measures of classifier learning on the imbalanced data set.

Increasing the past-concussion data with additional synthetic data points substantially aids in training to predict concussion history (Figure 3(a) vs (b)), and prevents overfitting on the training data.

The majority of prior concussion-classification analyses pursued in this paper employ non-linear separators, using SVM's radial basis kernel. Equivalent experimentation with a linear separator SVM achieved 75% accuracy at best, using 23 features, and only 71% accuracy with 15 features. These findings indicate the presence of a non-linear classification boundary for concussions using cognitive measures.

Beyond simple "concussion" and "non-concussion" labels, the (radial basis kernel) SVM classifier provides a score for each cadet expressing the confidence of its classification. Larger magnitude positive numbers are more confidently concussion, larger magnitude negative numbers are more confidently non-concussion. We find a slight but insignificant increase in concussion classifier score for cadets with less recent concussions (Figure 4).

Discussion

Neurocognitive assessments help to provide a baseline for an athlete/military member in the event of concussion injury and can improve post-concussion assessment and care. They are recommended as part of a multifaceted approach to return-to-play determinations.[14] The purpose of this study was to use machine learning, specifically support vector machines, to investigate whether current neurocognitive performance could accurately predict a prior concussion that occurred within the past year, indicating statistically significant differences in cognitive abilities between individuals without a history of concussion and those with a history of recent concussion.

High school and college athletes tend to show cognitive recovery from a concussion within five-to-seven days.[1] However, we found that the cognitive effects of concussions remain and are detectable at least a year after injury, far beyond the typical two weeks of recovery. ImPACT's five composite measures are sufficient to achieve over 88% accuracy in determining if a subject did or did not have a concussion in the past year. Mixing ten composite and individual task measures achieves over 95% accuracy. Prior concussion effects remained roughly constant across the first twelve months after the event. There is significant difference in cognitive abilities between individuals without a history of concussion and those with a history of recent concussion.

This finding has several implications. First, our results indicate the importance of continuing to

Table 2. Comparison of Precision, Recall, and F-score metrics for SVM learning on SMOTE data set with subset of ImPACT features.

	Precision	Recall	F-score	Acc. concuss	Acc. Non-concuss	Acc. average
Greedy						
5 feat	0.91	0.86	0.89	0.86	0.89	0.88
10 feats	0.98	0.96	0.97	0.96	0.98	0.97
15 feats	0.97	0.98	0.98	0.98	0.97	0.98
20 feats	0.95	0.98	0.97	0.98	0.95	0.97
25 feats	0.97	0.98	0.98	0.98	0.97	0.98
Composite-only						
1 feat	0.46	0.4	0.43	0.4	0.54	0.47
2 feats	0.54	0.35	0.42	0.35	0.77	0.56
3 feats	0.78	0.7	0.74	0.7	0.79	0.74
4 feats	0.89	0.89	0.89	0.89	0.87	0.88
5 feats	0.85	0.93	0.89	0.93	0.82	0.87

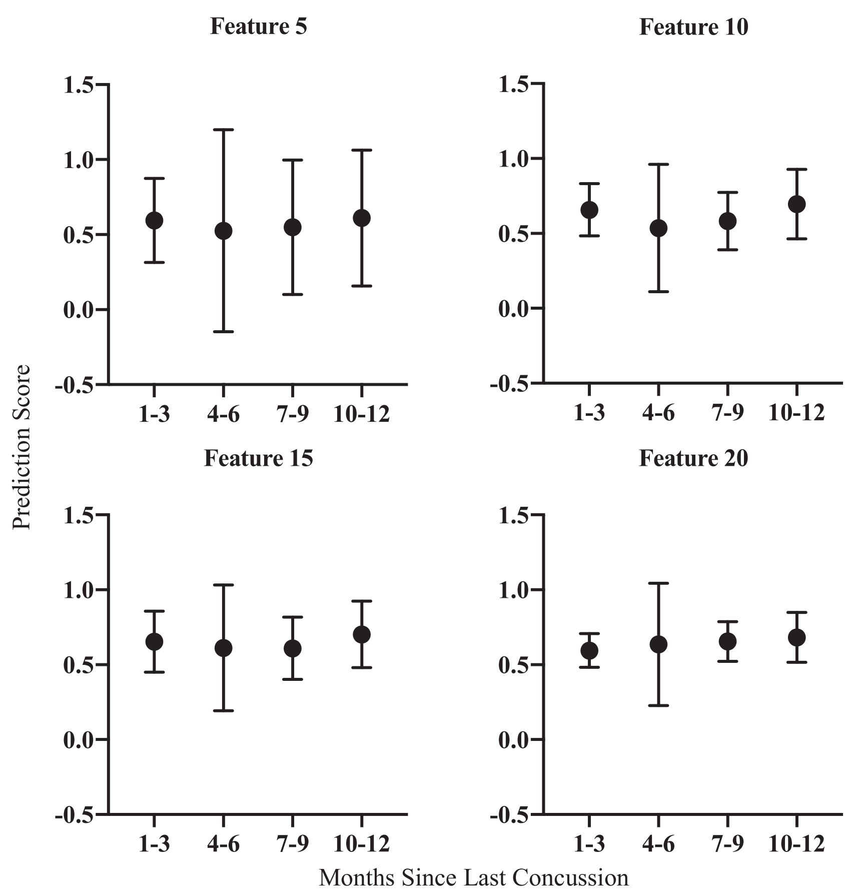

Figure 4. Prediction score mean and standard deviation using the first 5, 10, 15, and 20 features chosen through greedy feature selection.

follow concussions patients for many months after recovery. Given the effect on cognition, this may be particularly true for student-athletes. Special note must be taken of all elements of cognition including academic performance, and further evaluation and treatment may be advisable.

Further, additional research is needed to determine how long post-concussion physiology effects cognitive performance. This question is also being asked in emerging MRI research that indicates that some aspects of brain physiology remain changed one year after a concussion.[2] Thus, more longitudinal research is needed to determine whether cognitive effects remain active for multiple years beyond the concussive incident. Past studies[6] have established that history of prior concussions may worsen the effects of additional concussion injury.

Additionally, this finding needs to be generalized to different age populations. Most participants in this study were late adolescents. It is unknown whether these findings would generalize to pediatric, younger adolescent, and mature populations. Thus, more lifecycle research is needed to determine the lingering cognitive effects on different aged populations.

Also, the long-term effects of concussion are variable and complex. Effects do not co-vary linearly across all subjects with prior concussion, but follow one of a set of trends captured by our non-linear separator. As patients are followed post-concussion, it will be important to recognize their long-term recovery may not fall into a single track. Future work is needed to understand the nature of these multiple sets of cognitive trends and the links between these trends.

In the light of non-disclosure of concussion reported in the athlete and military personnel literature,[26] a secondary application of these findings may be the retrospective diagnosis of an undisclosed concussion. Some concussion sufferers are unaware of their injury or may choose not to self-report a potential concussion for a variety of reasons.[27] However, this was not a tested hypothesis and thus it would be a misapplication of these findings. A better application would be longer monitoring and follow-up care of concussion sufferers.

Expansions to the methodology in the present work would provide valuable additional perspective on the long-term cognitive impacts of concussion. Concussion identification was computed using support vector machine (SVM) classifiers using default settings from MATLAB.[21] Exploration of additional variants of SVM and of other classifiers would provide valuable additional perspective on what cognitive patterns continue to be expressed in the months after concussion "recovery." Additional investigation of SMOTE parameters on learned concussion classification would provide similar valuable insights.

Expansions to the data set in the present work would provide valuable additional perceptive. Concussion effects are studied only in the first year after onset; future access to a broader data set will be valuable to allow investigation of cognitive effects in additional years and decades after onset. "Concussion" occurrence is determined by participant self-report in the present data set. Future study of data where identification of concussions is provided by professionals, and accompanied by mechanism of injury, will provide valuable insight into distinctions in the evolution of cognition after concussion. Additionally, the impact of multiple concussions is not considered in the present study, despite its known influence on short-term recovery. Future study removing this confound promises to improve the already strong SVM ability to predict recent concussion history based on cognition.

In summary, cognitive recovery from the acute effects of a concussion tends to be within one week,[1] but concussions have lingering effects that may be detectable long after an athlete is cleared to returned-to-play or a Soldier is cleared to return-to-duty. The purpose of this study was to determine whether a machine learning model could accurately predict whether a recent concussion had occurred based on current cognitive performance. A support vector machine analysis, a specific sub-type of machine learning model, generated a prediction model that had greater than 90% accuracy for both prior concussion and no prior concussion classification. ImPACT composite scores visual-motor speed and impulse control were instrumental to increasing predictive accuracy.

Authorship contributions

Daniel D Leeds directed data analysis and co-wrote manuscript; Annie Nguyen led data analysis, collated tables and figures, and wrote manuscript initial drafts; Christopher D'Lauro advised on research and edited manuscript; Jonathan C Jackson collected cadet data, advised on interpretation of results, and edited manuscript; Brian R Johnson directed research and co-wrote manuscript.

Disclaimers

Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the author, and are not to be construed as official, or as reflecting true views of the Department of the Army or the Department of Defense.

Acknowledgements

We are grateful to our colleagues Col (ret.) Darren Campbell and Lt Col (ret.) Gerald McGinty for their participation with data collection.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Brian R Johnson https://orcid.org/0000-0001-5754-3223

References

1. Williams RM, Puetz TW, Giza CC, et al. Concussion recovery time among high school and collegiate athletes: a systematic review and meta-analysis. *Sports Med* 2015; 45: 893–903.
2. Churchill NW, Hutchison MG, Graham SJ, et al. Mapping brain recovery after concussion. *Neurology* 2019; 93: e1980–e1992.
3. Bosner BE, Guyon IM and Vapnik VN. A training algorithm for optimal margin classifiers. In: *Proceedings of the 5th annual ACM workshop on computational learning theory* (ed D Haussler), 1992, pp.144–152. Pittsburgh, PA: ACM Press.
4. Centers for Disease Control and Prevention. *Surveillance report of traumatic brain injury-related emergency department visits, hospitalizations, and deaths—United States, 2014*. Atlanta, GA: Centers for Disease Control and Prevention, U.S. Department of Health and Human Services, 2019.
5. McCrory P, Meeuwisse W, Dvorak J, et al. Consensus statement on concussion in sport – the 5th international conference on concussion in sport held in Berlin, October 2016. *Brit J Sport Med* 2017; 51: 838–847.
6. Levin HS, Eisenberg HM and Benton AL (eds) *Mild head injury*. Oxford, UK: Oxford University Press, 1989.
7. Lovell MR and Collins MW. New developments in the evaluation of sports-related concussion. *Curr Sports Med Rep* 2002; 1: 287–292.
8. Education Commission of the States. Education Commission of the States, www.ecs.org (accessed 18 March 2021).
9. Elbin RJ, Schatz P and Covassin T. One-year test-retest reliability of the online version of ImPACT in high school athletes. *Am J Sports Med* 2011; 39: 2319–2324.
10. Mayers LB and Redick TS. Clinical utility of ImPACT assessment for postconcussion return-to-play counseling. Psychometric issues. *J Clin Exp Neuropsychol* 2012; 34: 235–242.
11. Allen BJ and Gfeller JD. The immediate post-concussion assessment and cognitive testing battery and traditional neuropsychological measures: a construct and concurrent validity study. *Brain Inj* 2011; 25: 179–191.
12. Covassin T, Elbin RJ, Stiller-Ostrowski JL, et al. Immediate post-concussion assessment and cognitive testing (ImPACT) practices of sports medicine professionals. *J Athl Train* 2009; 44: 639–644.
13. Asken BM, McCrea MA, Clugston JR, et al. "Playing through it": delayed reporting and removal from athletic activity after concussion predicts prolonged recovery. *J Athl Train* 2016; 51: 329–335.
14. Giza CC and Hovda DA. The new neurometabolic cascade of concussion. *Neurosurgery* 2014; 4: S24–S33.
15. Bergeron MF, Landset S, Maugans TA, et al. Machine learning in modeling high school sport concussion symptom resolve. *Med Sci Sports Exerc* 2019; 51: 1362–1371.
16. Daley M, Dekaban G, Bartha R, et al. Metabolomics profiling of concussion in adolescent male hockey players: a novel diagnostic method. *Metabolomics* 2016; 12: 185.
17. Visscher RM, Feddermann-Demont N, Romano F, et al. Artificial intelligence for understanding concussion: retrospective cluster analysis on the balance and vestibular diagnostic data of concussion patients. *Plos One* 2019; 14: e0214525.
18. Boshra R, Dhindsa K, Boursalie O, et al. From group-level statistics to single-subject prediction: machine learning detection of concussion in retired athletes. *IEEE Trans Neural Syst Rehabil Eng* 2019; 27: 1492–1501.
19. ImPACT Version 2.0 Clinical User's Manual, 2004 Retrieved September 27, 2018, from http://www.impact-test.com/clients.htm.
20. Stroop JR. Studies of interference in serial verbal reactions. *J Exp Psychol* 1935; 18: 643–662.
21. MATLAB 2018a, The MathWorks, Inc., Natick, Massachusetts, United States.
22. John GH, Kohavi R and Pfleger K. Irrelevant features and the subset selection problem. In: *Machine learning proceedings*, New Brunswick, NJ, 1994, pp.121–129.
23. Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 2002; 16: 321–357.
24. Kubat M, Holte R and Matwin S. Machine learning for the detection of oil spills in satellite radar images. *Mach Learn* 1998; 30: 195–215.
25. Weiss GM. Mining with rarity: a unifying framework. *Sigkdd Explor Newsl* 2004; 6: 7–19.
26. Foster CA, D'Lauro C and Johnson BR. Pilots and athletes: different concerns, similar concussion non-disclosure. *PLoS One* 2019; 14: e0215030.
27. Baugh CM, Kroshus E, Daneshvar DH, et al. Perceived coach support and concussion symptom-reporting: differences between freshmen and non-freshmen college football players. *J Law Med Ethics* 2014; 42: 314–322.