Comparative Performance of Three Eye-Tracking Devices in Detection of Mild Traumatic Brain Injury in Acute Versus Chronic Subject Populations

John King[*,†]; *Chantele Friend*[‡]; *Dong Zhang*[§]; *Walter Carr*[†]

ABSTRACT

Introduction:

Presently, traumatic brain injury (TBI) triage in field settings relies on symptom-based screening tools such as the updated Military Acute Concussion Evaluation. Objective eye-tracking may provide an alternative means of neurotrauma screening due to sensitivity to neurotrauma brain-health changes. Previously, the US Army Medical Research and Development Command Non-Invasive NeuroAssessment Devices (NINAD) Integrated Product Team identified 3 commercially available eye-tracking devices (SyncThink EYE-SYNC, Oculogica EyeBOX, NeuroKinetics IPAS) as meeting criteria toward being operationally effective in the detection of TBI in service members. We compared these devices to assess their relative performance in the classification of mild traumatic brain injury (mTBI) subjects versus normal healthy controls.

Materials and Methods:

Participants 18 to 45 years of age were assigned to Acute mTBI, Chronic mTBI, or Control group per study criteria. Each completed a TBI assessment protocol with all 3 devices counterbalanced across participants. Acute mTBI participants were tested within 72 hours following injury whereas time since last injury for the Chronic mTBI group ranged from months to years. Discriminant analysis was undertaken to determine device classification performance in separating TBI subjects from controls. Area Under the Curves (AUCs) were calculated and used to compare the accuracy of device performance. Device-related factors including data quality, the need to repeat tests, and technical issues experienced were aggregated for reporting.

Results:

A total of 63 participants were recruited as Acute mTBI subjects, 34 as Chronic mTBI subjects, and 119 participants without history of TBI as controls. To maximize outcomes, poorer quality data were excluded from analysis using specific criteria where possible. Final analysis utilized 49 (43 male/6 female, mean [$\bar{x}$] age = 24.3 years, SD [s] = 5.1) Acute mTBI subjects, and 34 (33 male/1 female, $\bar{x}$ age = 38.8 years, s = 3.9) Chronic mTBI subjects were age- and gender-matched as closely as possible with Control subjects. AUCs obtained with 80% of total dataset ranged from 0.690 to 0.950 for the Acute Group and from 0.753 to 0.811 for the Chronic mTBI group. Validation with the remaining 20% of dataset produced AUCs ranging from 0.600 to 0.750 for Acute mTBI group and 0.490 to 0.571 for the Chronic mTBI group.

Conclusions:

Potential eye-tracking detection of mTBI, per training model outcomes, ranged from acceptable to excellent for the Acute mTBI group; however, it was less consistent for the Chronic mTBI group. The self-imposed targeted performance (AUC of 0.850) appears achievable, but further device improvements and research are necessary. Discriminant analysis models differed for the Acute versus Chronic mTBI groups, suggesting performance differences in eye-tracking. Although eye-tracking demonstrated sensitivity in the Chronic group, a more rigorous and/or longitudinal study design is required

[*]Oak Ridge Institute for Science and Education, Oak Ridge, TN 37830, USA

[†]Center for Military Psychology and Neuroscience, Walter Reed Army Institute of Research, Silver Spring, MD 20910, USA

[‡]Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD 20817, USA

[§]Department of Mathematics, Computer Science, and Digital Forensics, Bloomsburg University of Pennsylvania, Bloomsburg, PA 17815, USA

The information in this manuscript was previously presented as a poster at the Military Health System Research Symposium held in Kissimmee, FL, USA (August 14–17, 2023).

Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70–25. This research was supported in part by an appointment to the Department of Defense (DOD) Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and the DOD. ORISE is managed by ORAU under DOE contract number DE-SC0014664. The views expressed in this presentation are the private views of the authors, and are not to be construed as official or as reflecting true view of the Department of the Army, Defense Health Agency, Department of Defense, ORAU/ORISE, or any other U.S. Government agency. This work was prepared under Contract HT0014-22-C-0016 with DHA Contracting Office (CO-NCR) HT0014 and, therefore, is defined as U.S. Government work under Title 17 U.S.C.§101. Per Title 17 U.S.C.§105, copyright protection is not available for any work of the U.S. Government.

Corresponding author: John King, USA (john.e.king328.ctr@health.mil).

doi:https://doi.org/10.1093/milmed/usae205

to evaluate this observation. mTBI injuries were not controlled for this study, potentially reducing eye-tracking assessment sensitivity. Overall, these findings indicate that while eye-tracking remains a viable means of mTBI screening, device-specific variability in data quality, length of testing, and ease of use must be addressed to achieve NINAD objectives and DoD implementation.

INTRODUCTION

Presently, symptom-based screening tools such as the updated version of the Military Acute Concussion Evaluation (MACE2) are utilized in field settings to determine the need for traumatic brain injury referral with plans for implementation across clinical settings as a diagnostic tool. The use of subjective screening tools is unreliable at times, and it has been reported that the sensitivity and specificity of the MACE2 decline when performed more than 12 hours post-injury.[1,2] Objective metrics, such as the use of eye-tracking technology, may provide alternative means of neurotrauma screening due to eye-tracking sensitivity to changes in the health of the neural pathways required for normal ocular motor function, especially in response to neurotrauma.[3–6] This potential use of eye-tracking was suggested in 2009 by Heitger et al. who compared 36 mild closed head injury subjects with post-concussion syndrome to age-matched controls and reported that "eye movements showed additional dysfunction in motor/visuospatial areas, response inhibition, visual attention and subcortical function" for subjects in the injury group.[7] Since then, eye-tracking assessments have been reported as having good sensitivity to disruptions to the neural pathways required for normal ocular motor function that can detect acute and sub-acute neurotrauma, and products are now receiving FDA approval for use in concussion diagnosis.[8–10] In a study comparing healthy football players with no sports-related contact for several months versus non-athletic peers, Kocher demonstrated that eye-tracking not only correctly classified players versus controls with an observed Area Under the Curve (AUC) of 0.984, but also appeared to be sensitive to chronic effects of sports-related impacts that were previously undetected.[11]

Several years ago, the US Army Medical Research and Development Command Non-Invasive NeuroAssessment Devices (NINAD) Integrated Product Team identified 3 commercially available eye-tracking devices (SyncThink EYE-SYNC, Oculogica EyeBOX, NeuroKinetics IPAS) as meeting criteria toward being operationally effective in the detection of TBI in service members. While other commercially available devices at the time essentially met the criteria established by NINAD, they were excluded for reasons such as requirements that data be uploaded to a cloud-based storage system which raised operational security concerns.

Although all 3 of these devices utilize eye-tracking to detect brain-health issues, they are unique in their form factor, test execution, and metrics (Fig. 1). The EYE-SYNC device is based on the Oculus Rift virtual reality goggles that the subject holds in place (Fig. 1A) for the duration of the assessment. Prior to testing, a calibration process is used to determine which eye provides better eye-tracking and testing proceeds with the selected eye. Through the goggles, the subject views a red dot that moves in a circular pattern and is instructed to follow that dot with their eyes. At the bottom left of Fig. 1A is a trace pattern of the subject's eye as it follows the dot, within which there are 8 black dots. These dots represent check points where the system captures eye-position relative to the moving target dot, the sum of which are all represented as an error cloud on the bottom right of Fig. 1A, which forms the basis of the primary measures for the EYE-SYNC device. Figure 1B shows the EyeBOX device in use as a subject sits quietly with their chin on a rest while looking at a screen on which a small video box moves around the edges of the screen in a pre-determined pattern. A major difference with this device is that the eyes are tracked by external cameras located below the screen as opposed to enclosed goggles. While the subject watches a short video as it moves around the screen, the system tracks both eyes simultaneously and measures the conjugacy of the eye movements throughout the testing. The lower portion of Fig. 1B shows the traces of each individual eye as recorded throughout the testing, from which calculations of conjugacy are developed. The third and final device, IPAS, is shown in use in Fig. 1C. The IPAS is similar to the EYE-SYNC in that it uses an enclosed goggle form factor; however, the IPAS is held in place with adjustable straps. The IPAS is like the EyeBOX in that it records both eyes independently. The IPAS is unique from the other 2 devices in that it can perform a wide variety of eye-tracking tests, which are customizable. The entire test battery of the IPAS takes approximately 30 minutes to complete; however, the actual test time depends on test selection. It should be noted that all these technologies were the most up-to-date versions at the time of study initiation and that each of these devices has been updated since then.

The purpose of this project was to compare these 3 devices in a head-to-head format to assess their relative performance in the ability to detect mild traumatic brain injury (mTBI) cases and distinguish from normal, healthy individuals. As time since injury may influence device performance, comparisons were performed using both acute and chronic mTBI subject populations.

METHODS

Participants diagnosed with acute mTBI at the Womack Army Medical Center (WAMC) within 72 hours of their injury were recruited and assigned to the Acute mTBI Group. Details of the injury and mTBI diagnosis were not shared with the research team; that is, the research team was simply informed of the mTBI diagnosis and that the subject was willing to participate in the study. Chronic mTBI participants were

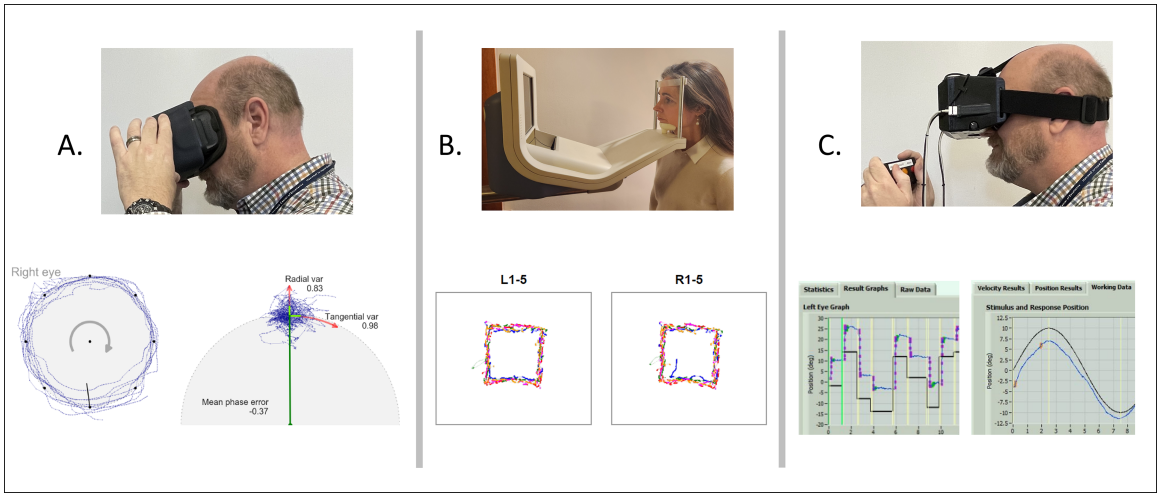

FIGURE 1. Images of 3 test devices used in study and representative data output. (A) EYE-SYNC: Three metrics are typically produced, test time about 2 minutes. (B) EyeBOX: Single metric (BOX score) is produced; test time is <4 minutes. (C) IPAS: Customizable test battery with multiple test capabilities, each with numerous variables and metrics. Complete test time about 30 minutes.

recruited from the Intensive Outpatient Program (IOP) at the National Intrepid Center of Excellence (NICoE) where a diagnosis of mTBI and comorbid psychological health issues was required for participation in the IOP. These participants were typically long-term post-injury (typically months to years) and assigned to the Chronic mTBI Group. Control participants without the history of mTBI as confirmed by intake interview were recruited from both Walter Reed National Military Medical Center and Fort Liberty, N.C. The Sport Concussion Assessment Tool version 3 (SCAT 3) is a self-reported symptom questionnaire consisting of 22 Likert scale questions ranging from 0 (No symptoms) to 6 (Severe symptoms) and was completed by all participants prior to study participation to screen eligibility criteria (i.e., unreported medical conditions). All participants were between 18 and 45 years old to minimize age-related oculomotor effects that occur after age 45. Participants from all 3 groups completed a TBI assessment protocol with all 3 previously discussed commercially available eye-tracking devices (EYE-SYNC, version 0.5.1 with Oculus positional driver version 1.0.9.0 and Oculus Runtime version 0.5.0.1-release-49,138; EyeBOX, version 2.124; IPAS, I-Portal version 3.5, VEST version 7.9), counterbalanced across participants. Discriminant analysis was performed using IBM SPSS Statistics version 29.0.0.0 to determine a classification model for each device that was able to differentiate between TBI subjects and controls. 80% of the dataset was used to build the classification model and the remaining 20% of the dataset was withheld and used to validate the model. The AUC results for each device were then calculated using the classification output for each respective device and used to compare the overall accuracy of their classifications. This analysis approach was necessary as each device utilizes proprietary tests and algorithms, and thus direct comparison of device raw data was not possible or appropriate. To evaluate device-related factors (e.g., technical difficulties) that might influence performance, research team notes were reviewed from subjects at WAMC in terms of data quality, the need to repeat tests, and technical issues experienced.

RESULTS

Overall, 63 participants were recruited as Acute mTBI subjects, 34 as Chronic mTBI subjects, and 119 participants without a history of TBI were recruited as control subjects. The data quality collected with the 3 devices was evaluated via internal device criteria and/or subject matter expert review as follows. For the EYE-SYNC device, a value of "0" for the FixationValid metric was used to determine the presence of poor-quality data, although other indicators of data quality including Test Error and EyeType were available. The TestError variable consisted of system-generated warning messages that ranged from minor (e.g., "13% of data was reported missing") to severe ("Only 0 valid fixation points") and approximately half of the collected data included a negative TestError report. The EyeType variable reports which of the 2 eyes was selected for data acquisition, and if data could not be obtained from either eye, "NeitherEye" was reported. In approximately 90% of the cases where "NeitherEye" was reported, the FixationValid metric value was "0," showing good consistency between the 2 metrics. For simplicity and to increase the number of subjects for which data could be used, the more conservative FixationValid metric was chosen as the determinant of EYE-SYNC data quality. Evaluating data quality was much simpler with the EYEBOX as it generated a quality score metric for each test. The recommendation obtained from Oculogica indicated that data obtained with quality scores higher than "6" are acceptable for use. The comprehensive test nature of the IPAS system complicates the ability to judge data quality due to the number of tests, the sophisticated nature of the tests and their analyses,

and the lack of clear data quality indicators for certain tests. As such, the IPAS requires an experienced user to manually review all the data obtained to determine if the data quality is acceptable. For this effort, poor data quality for the IPAS was hence determined by the presence of an invalid calibration and/or the presence of multiple tests that are deemed to be uninterpretable.

Using the data quality assessment criteria as described, the proportion of Acute mTBI data deemed to be poor-quality ranged from 6% (EyeBOX) to 56% (EYE-SYNC). The need to repeat testing was lowest for the EyeBOX (6%) followed by EYE-SYNC (21%) and IPAS (61%). It should be noted that the repeat rate for the IPAS is artificially inflated as it was counted as a repeat if any of the 18 tests within the test battery required a repeat whereas the other 2 devices consisted of a single test. Technical issues were defined as hardware or software issues that had to be addressed before data collection could be completed, typically requiring a reboot of the entire system. Using these criteria, the EyeBOX system failed less than 1% of the time, followed by EYE-SYNC (5%) and IPAS (15%).

To obtain the best-possible performance in detection of mTBI, poor quality data were excluded from analysis using the aforementioned criteria where possible; however, this was not possible with the EYE-SYNC device due to the high rate of poor-quality data. Hence, the dataset utilized for the final analysis required considerable retention of poor-quality data for the EYE-SYNC device to ensure adequate sample size. Given this, efforts were made to retain the best possible quality data for the EYE-SYNC device. The final analysis utilized 49 (43 male/6 female, mean [$\bar{x}$] age = 24.3 years, SD [s] = 5.1) Acute mTBI subjects and 34 (33 male/1 female, $\bar{x}$ age = 38.8 years, s = 3.9) Chronic mTBI subjects who were age- and gender-matched as closely as possible with equal numbers of Control subjects 49 (41 male/8 female, $\bar{x}$ age = 24.4, s = 5.0) and 34 (31 male/ 3 female, $\bar{x}$ age = 38.2 years, s = 3.9), respectively. For subjects assigned to the Acute mTBI group, all reported mTBI diagnosis within past 72 hours and their SCAT 3 scores ranged from 7 to 95 ($\bar{x}$ = 29.5, s = 19.4). Subjects assigned to the Chronic mTBI group were similar with 100% reporting history of mTBI and SCAT 3 scores ranging from 4 to 87 ($\bar{x}$ = 31.9, s = 20.7). For the Control subjects, 2 reported previous history of mTBI—one 11 years prior from sports, and one over a year prior during a motor vehicle accident. Both of these Control subjects reporting previous mTBI reported complete recovery, consistent with lack of reported symptoms on the SCAT 3 and intake survey. Reported symptoms on SCAT 3 were similar for Control subjects paired with the Acute mTBI group ($\bar{x}$ = 0.8, s = 1.7) and those paired with the Chronic mTBI group ($\bar{x}$ = 1.8, s = 3.1).

Figure 2 shows the AUC obtained for the 3 devices using 80% of the entire dataset with AUC values for the acute mTBI subjects on the left and AUC values for the Chronic mTBI subjects on the right. For the Acute mTBI group, in order from largest to smallest, the AUC values were 0.950 (EyeBOX), 0.845 (IPAS), and 0.690 (EYE-SYNC). For the Chronic mTBI group, again from largest to smallest, the AUCs obtained were 0.811 (IPAS), 0.796 (EyeBOX), and 0.753 (EYE-SYNC). Validation using the holdout 20% of the dataset demonstrated poor to fair performance of discriminant analysis classification for the Acute mTBI subjects, with AUCs ranging from 0.600 to 0.750. For the Chronic mTBI subjects, validation results indicated weaker performance of discriminant analysis classification, with AUCs ranging from 0.490 to 0.571 (Table I).

To confirm that devices provided significant classification performance versus simple guessing (0.5 probability of correctly guessing sensitivity and specificity), proportions analysis was performed for each device with the following P values obtained: IPAS, $P < 0.001$; EyeBOX, $P = 0.013$; and EYE-SYNC, $P = 0.066$. These results, which compare discriminant analysis performance versus chance classification, indicate that all 3 devices are performing as expected.

A summary of the data used to build discriminant analysis models for each device is presented in Table II. For all devices, the number of variables contributing to the discriminant analysis was less than the total number of available variables, especially for the EyeBOX and IPAS devices. For the EYE-SYNC device, the mean radial error variable did not contribute to the analysis for either mTBI group. For EyeBOX, an interesting finding is that the primary clinical measure, BOX score, did not contribute to either group's classification results. For IPAS, the Anti-Saccades, Auditory Reaction Time, and Horizontal Smooth Pursuit (0.01 Hz) tests contributed to classification results for both groups, albeit with different variable representations. For the EyeBOX and IPAS devices, the number of variables contributing to the discriminant analysis for the Chronic mTBI group were significantly less than for the Acute mTBI group. In fact, for the IPAS, only one-sixth of the tests in the overall test battery was useful for the analysis of the Chronic mTBI group.

Device-Specific Observations Were as Follows:

EYE-SYNC

Due to the large number of subjects with poor-quality data despite repeated testing, it was not possible to generate a sufficiently large data set free of poor-quality data for comparison with the other devices. A frequent issue reported by users was difficulty in obtaining clear eye-tracking as the EYE-SYNC device does not have a built-in means of adjusting camera placement on the pupils (other than manually moving the goggles around the subject's face). Often, one of the eyes would not track well during calibration, so the goggles would be adjusted to improve tracking for the affected eye; however, the repositioning of the goggles often resulted in the decline of the eye tracking performance of the opposite eye that had previously performed acceptably. Examiners

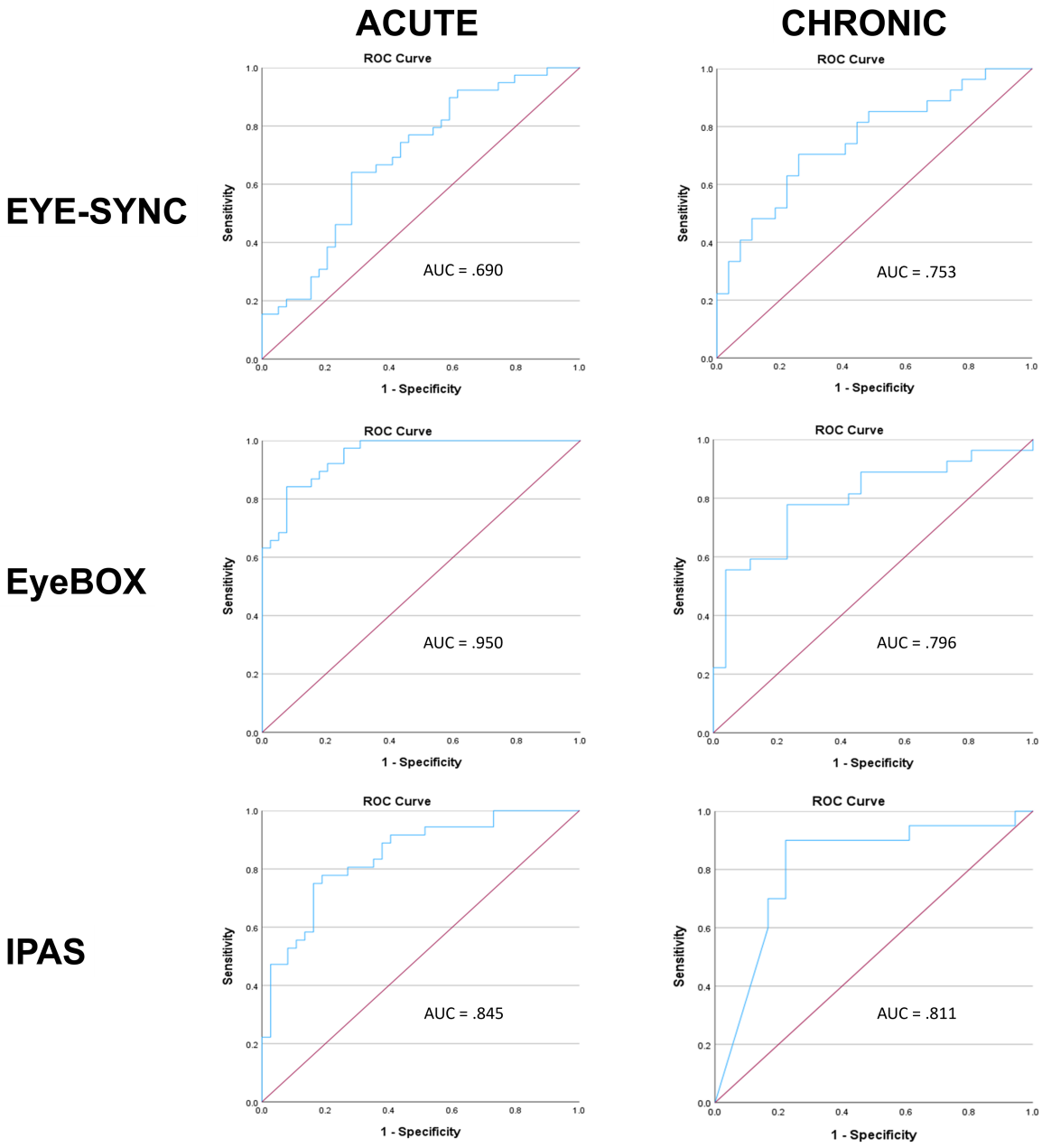

FIGURE 2. For all devices, the best possible outcomes for specified dataset were utilized as determined via discriminant classification. Values within each receiver operating characteristic (ROC) Curve indicate the respective area under the curve (AUC) calculated for that curve. ROC curves on the left represent data from acute mild traumatic brain injury (mTBI) subjects and curves on the right from chronic mTBI subjects.

were permitted 2 attempts to obtain acceptable quality eye-tracking performance; however, they reported failure to do so in many cases. Subject discomfort was frequently noted when attempting to adjust the goggle position to improve eye-tracking.

EyeBOX

EyeBOX data had the lowest incidence of poor-quality data and required the least repeat testing of all devices. There were very few issues reported by investigators regarding eye-tracking or technical errors making this system the most-user friendly of the 3. This device is the largest and least portable at present, with designs for a portable system in the works.

IPAS

The IPAS system includes an exhaustive battery of tests that may improve sensitivity, but at the cost of time, as it requires the most testing time at approximately 25 minutes. However, the test battery is customizable and could be shortened with

TABLE I. Discriminant Analysis Classification Summary for Each Device by Subject Condition

		Discriminant analysis classification (80% data set—model training)			Model validation (20% data set—withheld from training data set)		
Comparison	**Device**	**True positive rate**	**True negative rate**	**AUC**	**True positive rate**	**True negative rate**	**AUC**
Acute mTBI versus Controls	**EYE-SYNC**	46%	72%	0.690	60%	90%	0.700
	EyeBOX	87%	85%	0.950	70%	80%	0.750
	IPAS	78%	81%	0.845	30%	90%	0.600
Chronic mTBI versus Controls	**EYE-SYNC**	63%	78%	0.753	14%	86%	0.490
	EyeBOX	63%	77%	0.796	14%	100%	0.571
	IPAS	70%	78%	0.811	43%	43%	0.571

Discriminant analysis classification model training results using 80% of the data set are presented here for both Acute and Chronic mTBI groups versus matched controls. AUC values provide an overall measure of accuracy of classifying as mTBI or Control and True Positive/Negative Rates provide more detail on classification accuracy. The remaining 20% of the data set was used to validate classification results. These results indicate good performance of eye-tracking in detecting mTBI in the Acute group and weaker performance for the Chronic mTBI group. Abbreviations: AUC, area under the curve; mTBI, mild traumatic brain injury.

TABLE II. Variables Included in Discriminant Analyses

Device	**Total available data**	**Data included for acute mTBI analysis**	**Data included in chronic mTBI analysis**
EYE-SYNC	7 variables (1 test)	5 variables	5 variables
EyeBOX	102 variables (1 test)	42 variables	14 variables
IPAS	162 variables (18 tests)	24 variables (11 tests)	14 variables (3 tests)

For all 3 devices, the number of variables included in discriminant analysis was less than the total number of variables available. For the EyeBox and IPAS devices, the number of variables and tests (IPAS) entered into the discriminant model was significantly less for the Chronic mTBI group than for the Acute mTBI group, suggesting differences in eye-tracking performance for the 2 groups. Abbreviation: mTBI, mild traumatic brain injury.

the removal of less sensitive tests. Over 60% of subjects required at least 1 test to be repeated, further increasing testing time. A common issue reported by users was the difficulty in maintaining clear eye-tracking, which required constant vigilance and technical adjustment of parameters during testing. Participant discomfort was a noted complaint due to the size and weight of the goggle system. This device requires the most training, tester experience, and post-testing data analysis to obtain quality data and acceptable detection performance of mTBI.

DISCUSSION

While there is not a defined threshold for what is considered a good AUC score, it has been suggested that AUC values between 0.800 and 0.900 are considered excellent.[12] It was decided to adopt an AUC of 0.850 as the targeted objective since this threshold will be generally accepted as sufficient performance. This target was obtained by the training models for the EyeBOX (AUC = 0.950) and nearly so for the IPAS (AUC = 0.845) devices; however, this required the removal of poor-quality data and extensive post-testing analysis for the IPAS. The weak validation results are likely influenced by the small sample size ($n = 20$ for the Acute and $n = 14$ for the Chronic groups, respectively). Preliminary analysis of eye-tracking data collected from various tactical training environments suggests that the sensitivity of eye-tracking measures appears to be differentially sensitive to types of training exposures. As the injury mechanisms for the mTBI groups were uncontrolled, this likely resulted in decreased overall model performance due to the presence of a non-homogeneous subject population. On another note, it is interesting that the AUCs were generally larger for the Acute mTBI group, which would be expected when testing subjects closer to the point of injury. To reiterate, these models were developed from relatively small groups and are thus intended as proof of concepts and not to be used as working models. While the validation results are not as strong as desired, the training model AUCs suggest that the target of 0.850 is within reach especially considering that evolving data from current field research is demonstrating reproducible evidence of strong sensitivity of eye-tracking to brain health changes to blast exposures in different military training populations.

Findings indicated that data quality can sometimes be improved by repeating tests to remove artifacts (e.g., eyeblinks) or correct system-related pupil tracking issues, but increased testing time may reduce the likelihood of use in operational settings. Consequently, this experience indicates that factors such as device administrator training, administrator experience level, and device-specific technical issues contribute to overall device performance.

It is interesting that for the 2 devices with better performance (EyeBOX and IPAS), the critical variables utilized by discriminant analysis are quite different for the Acute versus Chronic mTBI groups, providing support for the potential presence of differences in eye-tracking performance for these 2 groups. This indicates that while an eye-tracking performance signal appears to be present in both groups following mTBI, the affected eye-tracking parameters are different. While it was not evaluated as part of the current study,

this difference in eye-tracking performance between the 2 groups should be taken into consideration before integrating eye-tracking into injury tracking protocols.

The injuries producing the mTBIs for both the Acute and Chronic subjects were not controlled, introducing variance into the model that may not exist when focusing the use of eye-tracking tests to a specific training population (e.g., Airborne). The Chronic mTBI group was represented by a mix of blunt force trauma and/or blast-related exposures yet demonstrated differences in eye-tracking performance from the Control group. This is suggestive that eye-tracking may be sensitive to both types of events; however, the current study design is insufficient for evaluation of such.

It was interesting that eye-tracking demonstrated sensitivity to mTBI in the chronic mTBI patient population at NICoE despite most participants being several months or even years post-injury. This may indicate that a signal from the previous mTBI is still present that eye-tracking is sensitive albeit at a sub-clinical level. As much of the patient population at NICoE reports multiple injuries and exposures during their career, it is possible that the signal detected by eye-tracking represents a cumulative result of repetitive injuries as reported by Kocher.[11] The classification of subjects into acute and chronic groups was only intended as a cursory examination of short- versus long-term effects of TBI on eye-tracking. That is, the interest of this classification was not intended to evaluate the effects of time on the performance of eye tracking, but rather as a description of injury state. It should be noted that these observations, while quite interesting, cannot be confirmed at present as a more rigorous and/or longitudinal study design is required to better control other variables and factors that may influence the eye-tracking performance in the mTBI population. Overall, these findings indicate that while eye-tracking remains a viable means of mTBI screening, device-specific variability in data quality, length of testing, and ease of use must be addressed to achieve NINAD objectives and DoD implementation.

CONCLUSIONS

1) The potential detection of mTBI with eye-tracking, per training model outcomes, ranged from acceptable to excellent for the Acute mTBI group; however, performance was not as consistent for the Chronic mTBI group.
2) The self-imposed targeted performance (AUC of 0.850) appears achievable with these data, but further understanding of the mTBI population and device improvements are necessary.
3) Data quality, participant comfort, and technical issues impact performance for some devices and increase test time due to repeat testing.
4) Results suggest eye-tracking ability to detect sequalae from mTBI may continue for some time post-injury, possibly when chronic symptoms are present.
5) Device-dependent variability in data quality, length of testing, and ease of use must be considered for NINAD objectives and DoD implementation.

ACKNOWLEDGMENTS

NICoE and TBICoE both contributed vital institutional resources necessary for execution of this study. Joint Program Committee 6 (JPC-6) provided resources necessary to obtain all devices used in this study. We would like to extend our deep appreciation to Dr M. Victoria Ingram and Mr Jacques Arrieux for their extensive support in collecting data and providing subject matter expertise toward this project.

INSTITUTIONAL REVIEW BOARD (HUMAN SUBJECTS)

Institutional review approval for this research was obtained by Walter Reed National Military Medical Center IRB.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC)

Not applicable.

INDIVIDUAL AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the article as follows: study conception and design: JK, WC; data collection: JK, CF, JA, MVI; analysis and interpretation of results: JK, DZ; draft manuscript preparation: JK, JA, MVI, WC. All authors reviewed the results and approved the final version of the manuscript.

INSTITUTIONAL CLEARANCE

Institutional clearance obtained.

SUPPLEMENT SPONSORSHIP

This article appears as part of the supplement "Proceedings of the 2023 Military Health System Research Symposium," sponsored by Assistant Secretary of Defense for Health Affairs.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

A large portion of the research data was obtained from clinical patients, and thus cannot be shared to protect privacy of individuals. The consent forms do not provide the research team with authority to share their data.

REFERENCES

1. Coldren RL, Kelly MP, Parish RV, Dretsch M, Russell ML: Evaluation of the military acute concussion evaluation for use in combat operations more than 12 hours after injury. Mil Med 2010; 175(7): 477–81. 10.7205/milmed-d-09-00258
2. Pryweller JR, Baughman BC, Frasier SD, et al: Performance on the DANA brief cognitive test correlates with MACE cognitive score and may be a new tool to diagnose concussion. Front Neurol 2020; 11: 1–7. 10.3389/fneur.2020.00839
3. Balaban C, Hoffer ME, Szczupak M, et al: Oculomotor, vestibular, and reaction time tests in mild traumatic brain injury. PLoS One 2016; 11(9): 1–11. 10.1371/journal.pone.0162168
4. King JE, Pape MM, Kodosky PN: Vestibular test patterns in the NICoE intensive outpatient program patient population. Mil Med 2018; 183(suppl_1): 237–44. 10.1093/milmed/usx170

5. Maruta J, Suh M, Niogi SN, Mukherjee P, Ghajar J: Visual tracking synchronization as a metric for concussion screening. J Head Trauma Rehabil 2010; 25(4): 293–305. 10.1097/HTR.0b013e3181e67936
6. Samadani U, Ritlop R, Reyes M, et al: Eye tracking detects disconjugate eye movements associated with structural traumatic brain injury and concussion. J Neurotrauma 2015; 32(8): 548–56. 10.1089/neu.2014.3687
7. Heitger MH, Jones RD, Macleod AD, Snell DL, Frampton CM, Anderson TJ: Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. Brain 2009; 132(Pt 10): 2850–70. 10.1093/brain/awp181
8. Kullmann A, Ashmore RC, Braverman A, et al: Portable eye-tracking as a reliable assessment of oculomotor, cognitive and reaction time function: normative data for 18-45 year old. PLoS One 2021; 16(11): 1–16. 10.1371/journal.pone.0260351
9. Maruta J, Spielman LA, Rajashekar U, Ghajar J: Association of visual tracking metrics with post-concussion symptomatology. Front Neurol 2018; 9: 1–8. 10.3389/fneur.2018.00611
10. Samadani U, Li M, Qian M, et al: Sensitivity and specificity of an eye movement tracking-based biomarker for concussion. Concussion 2016; 1(1): 1–6. 10.3389/fneur.2022.1039955
11. Kocher C: *Ocular Motility Following Sport-Related Impacts*. Bloomsburg University of Pennsylvania; 2015.
12. Hosmer DW, Lemeshow S, Sturdivant RX: *Applied Logistic Regression*. 3rd edition/ edn, Wiley series in probability and statistics. Wiley; 2013: 173–82.