Test Performance and Test-Retest Reliability of the Vestibular/Ocular Motor Screening and King-Devick Test in Adolescent Athletes During a Competitive Sport Season

Phillip R. Worts,*† MS, ATC, Philip Schatz,‡ PhD, and Scott O. Burkhart,§‖ PsyD
Investigation performed at Florida State University, Tallahassee, Florida, USA

Background: The Vestibular/Ocular Motor Screening (VOMS) and King-Devick (K-D) test are tools designed to assess ocular or vestibular function after a sport-related concussion.

Purpose: To determine the test-retest reliability and rate of false-positive results of the VOMS and K-D test in a healthy athlete sample.

Study Design: Cohort study (diagnosis); Level of evidence, 2.

Methods: Forty-five healthy high school student-athletes (mean age, 16.11 ± 1.43 years) completed self-reported demographics and medical history and were administered the VOMS and K-D test during rest on day 1 (baseline). The VOMS and K-D test were administered again once during rest (prepractice) and once within 5 minutes of removal from sport practice on day 2 (removal). The Borg rating of perceived exertion scale was administered at removal. Intraclass correlation coefficients were used to determine test-retest reliability on the K-D test and the average near point of convergence (NPC) distance on the VOMS. Level of agreement was used to examine VOMS symptom provocation over the 3 administration times. Multivariate base rates were used to determine the rate of false-positive results when simultaneously considering multiple clinical cutoffs.

Results: Test-retest reliability of total time on the K-D test (0.91 [95% CI, 0.86-0.95]) and NPC distance (0.91 [95% CI, 0.85-0.95]) was high across the 3 administration times. Level of agreement ranged from 48.9% to 88.9% across all 3 times for the VOMS items. Using established clinical cutoffs, false-positive results occurred in 2% of the sample using the VOMS at removal and 36% using the K-D test.

Conclusion: The VOMS displayed a false-positive rate of 2% in this high school student-athlete cohort. The K-D test's false-positive rate was 36% while maintaining a high level of test-retest reliability (0.91). Results from this study support future investigation of VOMS administration in an acutely injured high school athletic sample. Going forward, the VOMS may be more stable than other neurological and symptom report screening measures and less vulnerable to false-positive results than the K-D test.

Keywords: high school; concussion; ocular; vestibular; sports; sideline; acute

A sport-related concussion (SRC) is recognized as a highly individualized injury, and athlete-specific symptom presentation may include impairments in cognitive function, mental status, balance, gait, and vision, requiring a multimodal assessment to correctly diagnose the injury.[18] The recent literature suggests that ocular and vestibular disturbances may be detected by screening tools in approximately 70% of patients after a SRC, and subjective reporting of "visual problems" and "dizziness" in concussed athletes also supports the idea of underlying ocular or vestibular problems.[3,9,14] The King-Devick (K-D) test and the Vestibular/Ocular Motor Screening (VOMS) are clinical instruments used to assess ocular and/or vestibular function.[8,19] The K-D test examines eye movements and the saccade pathways,[9] and the VOMS examines the vestibulo-ocular and oculomotor systems.[19] The oculomotor and vestibulo-ocular pathways are distributed throughout

‖Address correspondence to Scott O. Burkhart, PsyD, Children's Health Andrews Institute for Orthopaedics and Sports Medicine, 7211 Preston Road, Suite 1400, Plano, TX 75024, USA (email: scott.burkhart@childrens.com).
*Department of Nutrition, Food and Exercise Sciences, College of Human Sciences, Florida State University, Tallahassee, Florida, USA.
†Tallahassee Orthopedic Clinic, Tallahassee, Florida, USA.
‡Department of Psychology, Saint Joseph's University, Philadelphia, Pennsylvania, USA.
§Children's Health Andrews Institute for Orthopaedics and Sports Medicine, Plano, Texas, USA.
One or more of the authors has declared the following potential conflict of interest or source of funding: P.S. serves as a consultant to ImPACT Applications and is a member of its scientific advisory board.

The American Journal of Sports Medicine
1–7
DOI: 10.1177/0363546518768750

the brain, making these systems highly susceptible to dysfunction after a SRC.[28] Therefore, the inclusion of ocular- and vestibular-related testing appears to be a critical component of the multimodal concussion evaluation. To date, both tools have demonstrated utility in identifying a SRC and appear to be reliable measures.[20,31]

Established clinical tools for identifying a SRC (such as the Sport Concussion Assessment Tool 2 [SCAT2], Balance Error Scoring System [BESS], and Immediate Post-Concussion Assessment and Cognitive Testing [ImPACT]) have displayed vulnerability when physical activity is introduced before testing. Factors such as hydration status, testing environment, presence of an orthopaedic injury, exercise participation, and time between strenuous exercise and evaluation have negatively affected the diagnostic utility of these clinical tools.[4,5,21,27,29] Of the previously mentioned vision tests (VOMS and K-D test), only the K-D test was evaluated for acute changes in test performance within the context of exercise.[8,9,11,13] However, these K-D test cohorts were almost entirely male and older than 18 years of age, thus making generalizability for adolescents and female patients difficult. Additionally, all of the previously mentioned studies reported sample sizes of less than 40. The purpose of the current study was to investigate the utility of the K-D test and VOMS in a -representative adolescent sample, examine test-retest reliability, provide base rates in healthy participants and examine test performance within the context of sport participation.

METHODS

Participants

Healthy (nonconcussed) student-athletes were recruited from local-area high schools during the 2015-2016 academic year. Participants were included if they were competing in a high school–sanctioned sport at the time of recruitment and were between 13 and 18 years old. Participants were excluded if they did not complete all 3 testing sessions, suffered a diagnosed concussion in the 9 months before baseline testing, suffered a diagnosed concussion while enrolled in the study, or reported a history of ocular disorders. All participants and legal guardians provided consent/assent/parental permission before participation. All methods were approved by the Florida State University Institutional Review Board.

Design

A cohort of 45 participants was administered the VOMS and K-D test at the beginning of their competitive season in a classroom setting (baseline), later during their competitive season in a classroom setting before practice (prepractice), and within 5 minutes of their removal from sport practice on the sidelines on the same day as prepractice testing (removal).

The participants provided self-reported medical information during the baseline session. Personal and/or family medical history, including but not limited to concussion history, migraines, motion sickness, ocular disorders, attention/hyperactivity disorders, and psychological conditions, was collected.

Instrumentation

Vestibular/Ocular Motor Screening. The VOMS is a brief clinical tool used to assess vestibular and ocular motor impairment through patient-reported symptom provocation across 7 items.[19] The items include smooth pursuit, horizontal saccade, vertical saccade, near point of convergence (NPC), horizontal vestibulo-ocular reflex (VOR), vertical VOR, and visual motion sensitivity (VMS). In addition to symptom provocation, average NPC distance across 3 trials (in cm) is included. Clinical cutoff scores were based on $\geq$2 symptom provocation on the VOR and VMS items and NPC distance $\geq$5 cm.[19]

K-D Test. The K-D test is a rapid, number-naming sideline screening test for a concussion that requires less than 2 minutes to administer.[8] When establishing baseline K-D test times, the participant is required to complete 2 error-free trials with 3 cards, and the quickest total time is used. For a concussion screening, the total time and errors are recorded using the same number of cards as baseline. If reading times are slower or any errors are committed, the student-athlete should follow up with his or her health care professional. The version 1 plastic cards were utilized for this study.

Borg Rating of Perceived Exertion (RPE) Scale

The Borg RPE scale is a method to determine perceived physiological and psychological stress.[1] The scale ranges from 6 (very, very light) to 20 (very, very hard). RPE was recorded during removal to quantify the physical intensity of sport practice.

Procedures

The participants were recruited at team-specific parent meetings, and written consent forms were completed before participation in the study. The trained research team, consisting of 4 athletic trainers, a neuropsychologist, and an honors psychology student, administered the VOMS and K-D test in accordance with the instructions provided with the clinical assessment tools. Of the 45 participants included in the study, 30 were administered the VOMS before the K-D test, and 15 were administered the K-D test before the VOMS. During the prepractice session, participants were required to complete 1 error-free trial. During the removal session, reading times and errors were recorded. Because there was a chance that testing might occur during a tactical practice rather than a scrimmage or traditional practice, the participants were asked to define if it was a contact or noncontact practice during removal. Any athletes who performed slower reading times or committed errors were not removed from play, but the researchers were in communication with the on-site staff athletic trainer to determine if the athlete experienced

a delayed presentation of signs or symptoms indicative of a concussion.

Statistical Analysis

Participant demographic information was reported using descriptive statistics. Test-retest reliability was calculated using intraclass correlation coefficients (ICCs). The ICC model performed was a "1-way"–type "consistency," and the averages across 3 measures were reported. ICCs were also generated for baseline to prepractice, baseline to removal, and prepractice to removal comparisons. We defined high reliability as ICC values $\geq$0.90, good reliability as 0.80-0.89, fair reliability as 0.70-0.79, and poor reliability as $<$0.70.[23] Reliable change indices (RCIs) were calculated to determine the estimate of the probability that the scores were not caused by measurement error.[22] The modified RCI formula also includes an adjustment for practice effects.[2] Test-retest reliability of the VOMS items could not be performed using ICCs because of a lack of variability in symptom provocation; therefore, the level of agreement was defined as any change in symptom provocation for each individual item across time points as described in a previous work.[31] Multivariate base rate analyses were used to simultaneously interpret times and errors on the K-D test and symptom provocation of the VOMS items and NPC distance during removal.[10] A series of chi-square tests were performed to determine if distributions of false-positive results were different between test administration order or practice type. Independent-samples t tests were performed when comparing the subset of patients with a false-positive test finding (VOMS or K-D test) to the normal performance sample for days between testing, minutes between prepractice and removal, RPE, and duration of practice. Abnormal test performance for the K-D test was defined as an increase in total time (in seconds) from baseline to prepractice, prepractice to removal, or any errors committed during the removal trial. Abnormal test performance for the VOMS was defined as $\geq$2 symptom provocation increase for the VOR and VMS items and an average NPC distance $\geq$5 cm across 3 trials. All analyses were conducted with the alpha level set at $P < .05$. Statistical analysis was performed using SPSS Version 23 (IBM) and Excel (Microsoft).

RESULTS

Demographic Data

The sample of healthy student-athletes consisted of 45 participants (24 male, 21 female) with a mean age of 16.11 $\pm$ 1.43 years. Seven sports were represented in the sample: cheerleading (n = 4, 8.9%), football (n = 13, 28.9%), girls' volleyball (n = 7, 15.6%), boys' soccer (n = 10, 22.2%), softball (n = 6, 13.3%), baseball (n = 1, 2.2%), and flag football (n = 4, 8.9%). The medical history of the participants is described in Table 1. One participant reported a history of both migraines and motion sickness. Forty-four participants denied a history of seizures, and 1 participant was unsure.

TABLE 1
Self-Reported Medical History[a]

	n (%)
Previous concussion (N = 45)	9 (20.0)
Migraine (N = 40)	12 (30.0)
Motion sickness (N = 43)	7 (16.3)
Attention-deficit disorder/attention-deficit hyperactivity disorder (N = 43)	4 (9.3)
Depression/anxiety (N = 40)	1 (2.5)
Learning disability (N = 35)	7 (20.0)

[a]N in each row indicates the number of participants responding to each question.

TABLE 2
VOMS Symptom Provocation and K-D Test[a]

	Baseline	Prepractice	Removal
VOMS			
Smooth pursuit	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.13 $\pm$ 0.69
Horizontal saccade	0.27 $\pm$ 0.72	0.13 $\pm$ 0.66	0.18 $\pm$ 0.96
Vertical saccade	0.53 $\pm$ 0.92	0.13 $\pm$ 0.76	0.24 $\pm$ 1.05
NPC	0.13 $\pm$ 0.63	0.20 $\pm$ 0.99	0.11 $\pm$ 1.01
Horizontal VOR	0.42 $\pm$ 1.10	0.44 $\pm$ 1.42	0.22 $\pm$ 1.46
Vertical VOR	0.47 $\pm$ 1.33	0.27 $\pm$ 0.99	0.31 $\pm$ 1.38
VMS	0.42 $\pm$ 1.12	0.36 $\pm$ 1.07	0.20 $\pm$ 1.38
NPC distance, cm	2.18 $\pm$ 3.46	2.64 $\pm$ 4.25	3.50 $\pm$ 6.02
K-D test			
Total time, s	44.57 $\pm$ 7.57	44.57 $\pm$ 8.49	40.99 $\pm$ 6.91

[a]Data are shown as mean score $\pm$ SD, unless otherwise indicated. K-D, King-Devick; NPC, near point of convergence; VMS, visual motion sensitivity; VOMS, Vestibular/Ocular Motor Screening; VOR, vestibulo-ocular reflex.

Test Performance

The fastest times for the K-D test at baseline, prepractice, and removal were 44.57 $\pm$ 7.57 seconds, 44.57 $\pm$ 8.49 seconds, and 40.99 $\pm$ 6.91 seconds, respectively. Both VOR items provoked the greatest increase in symptoms at baseline (+5) and prepractice (+7), but horizontal saccade saw the greatest increase (+7) at removal (Table 2).

Test-Retest Reliability

Total times for the K-D test displayed a wide range at baseline (31.92-64.59 seconds), prepractice (31.70-69.34 seconds), and removal (27.59-60.03 seconds). The range of VOMS symptom provocation was large (–4 to +7) across the 3 times, as was the NPC distance (0 to 28.8 cm). Reliability for the K-D test's total time was good to high (0.91 [95% CI, 0.86-0.95]) across 3 times, good to high (0.90 [95% CI, 0.82-0.94]) from baseline to prepractice, fair to high (0.87 [95% CI, 0.76-0.93]) from baseline to removal, and fair to high (0.85 [95% CI, 0.72-0.92]) from prepractice to removal. Level of agreement ranged from 57.8% to 100.0% across any 2 times and from 48.9% to 88.9% across all 3 times (Table 3). Reliability of the average NPC

TABLE 3
Level of Agreement for Symptom Provocation on the Vestibular/Ocular Motor Screening[a]

	Baseline to Prepractice	Baseline to Removal	Prepractice to Removal	All 3 Time Points
Smooth pursuit	45 (100.0)	41 (91.1)	40 (88.9)	40 (88.9)
Horizontal saccade	38 (84.4)	36 (80.0)	36 (80.0)	33 (73.3)
Vertical saccade	29 (64.4)	26 (57.8)	31 (68.9)	23 (51.1)
NPC	34 (75.6)	31 (68.9)	35 (77.8)	29 (64.4)
Horizontal VOR	28 (62.2)	29 (64.4)	31 (68.9)	24 (53.3)
Vertical VOR	27 (60.0)	26 (57.8)	33 (73.3)	22 (48.9)
VMS	29 (64.4)	31 (68.9)	32 (71.1)	26 (57.8)

[a]Data are shown as n (%). NPC, near point of convergence; VMS, visual motion sensitivity; VOR, vestibulo-ocular reflex.

TABLE 4
RCIs for Total Time on the King-Devick Test[a]

		Standard Error of Measure					
	r Value	Baseline	Prepractice	Removal	Standard Error of Difference	RCI	RCI With Practice Effects[b]
Baseline to prepractice	0.818	3.23	3.62		4.85	9.51	9.51
Baseline to removal	0.872	2.71		2.47	3.67	7.19	3.61
Prepractice to removal	0.838		3.42	2.78	4.41	8.64	5.06

[a]Measurements are in seconds. RCI, reliable change index.
[b]Practice effects were subtracted from the RCI to represent an improvement in scores.

distance was good to high (0.91 [95% CI, 0.85-0.95]) across the 3 times, fair to high (0.87 [95% CI, 0.77-0.93]) from baseline to prepractice, poor to good (0.76 [95% CI, 0.56-0.87]) from baseline to removal, and high (0.94 [95% CI, 0.90-0.97]) from prepractice to removal.

Reliable Change Indices

RCIs were calculated for K-D test total times and NPC distance. RCIs for K-D test total times using 95% CIs including mean improvements (practice effects) between time points were 9.51 seconds from baseline to prepractice, 3.61 seconds from baseline to removal, and 5.06 seconds from prepractice to removal (Table 4). RCIs for NPC distance using 95% CIs including practice effects between time points were 4.44 cm from baseline to prepractice, 5.79 cm from baseline to removal, and 2.14 cm from prepractice to removal.

Multivariate Base Rates at Removal

Multivariate base rates presented in Tables 5 and 6 are to assist with clinical interpretation of abnormal test performance within a healthy, adolescent student-athlete cohort during administration at removal. Sixteen percent of the cohort performed the K-D test slower at removal compared with their first 2 time points. Eighteen percent committed errors at removal, and 36% either committed an error at removal or performed slower than either baseline or prepractice. One participant performed worse (slower time) at removal and committed an error. Using the 0.89 predictive probability model from Mucha et al[19] ($\geq$2 symptom provocation and NPC distance $\geq$5 cm), 1 false-positive result occurred at removal. Examining symptom provocation alone at removal, 16% reported $\geq$2 symptom provocation on 1 VOMS item, 9% reported $\geq$2 symptom provocation on 2 items, and 4% reported $\geq$2 symptom provocation on 3 VOMS items (Table 6). An NPC distance $\geq$5 cm occurred in 24% of the cohort.

TABLE 5
Prevalence of Abnormal Performance on the King-Devick Test During Removal

	% of Cohort
Increase in time from baseline	8.9
Increase in time from prepractice	15.6
Any increase in time at removal	20.0
Errors during removal	17.8
Any increase in time at removal or any errors	35.6
Errors during removal and increased time from baseline	0.0
Errors during removal and increased time from prepractice	2.2

Removal Conditions

Thirty participants completed the VOMS and then the K-D test, and 15 participants completed the K-D test and then the VOMS. Twenty-four participants reported a noncontact

TABLE 6
Prevalence of Flagged VOMS Items During Removal[a]

	% of Cohort	Cumulative %
≥2 symptom provocation alone for VOR or VMS		
3 positive VOMS items	4.4	4.4
2 positive VOMS items	4.4	8.9
1 positive VOMS item	6.7	15.6
0 positive VOMS items	84.4	100.0
Any ≥2 symptom provocation and NPC distance ≥5 cm	2.2	
Any ≥3 symptom provocation and NPC distance ≥5 cm	2.2	
Any ≥4 symptom provocation and NPC distance ≥5 cm	0.0	

[a]NPC, near point of convergence; VMS, visual motion sensitivity; VOMS, Vestibular/Ocular Motor Screening; VOR, vestibulo-ocular reflex.

practice, and 21 reported a contact practice. The mean time between baseline and prepractice/removal was 41.7 ± 18.5 days (range, 12-69 days). The mean time between prepractice and removal was 113 ± 54 minutes (range, 45-256 minutes). The mean RPE was 11.53 ± 2.1, ranging from "very, very light" to "hard." The mean duration of time practicing before removal was 59 ± 30 minutes (range, 29-145 minutes). When determining the effect of test order (VOMS first or K-D test first), there was not a disproportionate number of false-positive or normal results (P = .38). When determining the effect of practice type (contact or noncontact), there was not a disproportionate number of false-positive or normal results (P = .36). When comparing the normal and false-positive groups, there were no significant differences between the days between testing (30.8 ± 14.4 days vs 31.81 ± 15.3 days, respectively; P = .82), time between prepractice and removal (115.8 ± 59.9 minutes vs 108.0 ± 42.0 minutes, respectively; P = .65), RPE (11.41 ± 2.1 vs 11.75 ± 2.1, respectively; P = .61), or time practicing (59.7 ± 25.6 minutes vs 58.0 ± 27.2 minutes, respectively; P = .83).

DISCUSSION

We examined the use of the VOMS and K-D test in the context of sport participation in a representative high school athlete cohort. Results show that the K-D test has a considerably higher rate of false-positive results than the VOMS (36% vs 2%, respectively) in a sample of healthy high school student-athletes participating in a sport and tested within 5 minutes of removal from practice. Previous research on the VOMS has only been performed in a clinical setting; therefore, findings could not be generalizable to the acute sideline evaluation.[19,25] The stability of the K-D test has previously been explored while participants exercised; however, these cohorts were predominantly male and older than 18 years of age, and sample sizes were less than 40.[8,9,11,13] The VOMS provoked minimal symptoms across 3 administration time points while the NPC distance slightly increased with exercise. In terms of absolute agreement of symptom provocation, the VOMS ranged from 48.9% (vertical VOR) to 88.9% (smooth pursuit) across 3 times. Two false-positive results (4%) occurred at baseline, and 1 false-positive result (2%) occurred at prepractice and removal. The K-D test-retest reliability ranged from 0.85 to 0.91, demonstrating good to high reliability and similar mean total times (44.57 ± 7.57 seconds vs 44.57 ± 8.49 seconds, respectively) for the 2 resting test sessions. Improvement in total times was seen at removal in ≥84% of the cohort compared with the 2 resting conditions. Improvements are consistent with previous studies examining exercise and K-D test performance.[8,9,11,13] One-third of the cohort (36%) performed slower at removal than baseline or prepractice or committed errors at removal. When accounting for reliable change, only 1 participant would have fallen outside the 95% CI with practice effects for total time. It is difficult to determine if the effect of physical exertion could fully account for the increased reading times and presence of errors committed on the K-D test during removal, but athletes tested in different environments (quiet vs simulated crowd noise) did not have significantly different reading times for the baseline K-D test.[24] The false-positive group and normal performance group did not differ in RPE reporting, so perceived exercise intensity may not significantly affect test performance.

Clinical Relevance

The need for objective and standardized sideline assessments of concussions, including brief screening of both neurocognitive and neurological functioning in addition to symptoms, led to the original development of the Standardized Assessment of Concussion (SAC) in 1997.[15-17] Since 1997, sideline concussion assessments have become increasingly difficult, with a number of neurocognitive, neurological, and symptom report screening tools available for the clinician. Of the commonly used sideline concussion tools available (SAC, SCAT, BESS, K-D test), only the K-D test and finger-to-nose task from the SCAT2 have demonstrated improved test performance after exercise.[6,7,9,11-13,26,27,30] Based on the current study, the VOMS demonstrates greater stability under the constraints of exertion in a healthy sample when compared with the K-D test. Sideline concussion assessments will likely continue to include some form of neurocognitive, neurological, and symptom report. Going forward, the VOMS may be more stable than other neurological and symptom report screening measures and less vulnerable to false-positive results than the K-D test.

Strengths and Limitations

To the authors' knowledge, this was the first study to investigate the use of the VOMS in a sport competition setting. It was also the first time that K-D test performance in the context of sport participation was examined with a sizeable female sample (47% of the cohort) and the largest

cohort using the K-D test with exercise to our knowledge. Additionally, this study was only the second to report test-retest reliability of the VOMS; however, a previous study included only 2 administration times.[31] The participants' reporting of medical history, RPE, contact practice type, and symptoms on the VOMS items were subject to self-report bias, but it was assumed that the responses were provided accurately and with integrity. Given the limited sample size (n = 45), and the strict inclusion criteria, findings may only be generalizable to all adolescent athletes competing in contact and noncontact sports without a recently diagnosed concussion. The order of administration and practice type (contact or noncontact) did not appear to affect the presence of a positive result on the K-D test or VOMS, which supports the use of multiple tests during an evaluation. When the false-positive group was compared with the normal performance group, there were no significant differences in days between testing, time between prepractice and removal, RPE, and duration of practice, suggesting little or no influence on test performance. It is difficult to directly compare the K-D test and VOMS performance, as the nature of the variables did not allow us to perform identical statistical tests. However, because both tools used multiple variables to determine a positive result (errors or time vs provocation and NPC distance), multivariate base rates may provide the most appropriate comparison.

Future Directions

Future research comparing the VOMS and K-D test in a large, acutely concussed sample will provide support for or against the use of the VOMS in a sideline setting. Additionally, the VOMS will need to be explored in a youth, collegiate, and professional athlete sample as the operating definition for specific symptoms may vary across age groups.

CONCLUSION

An immediate sideline concussion evaluation in sports is a complex process, as decisions often must be made quickly and accurately with limited or unvalidated diagnostic tools. Ongoing effort, both empirically and clinically, continues to progress toward the development of valid and reliable sideline concussion assessment measures intended to provide accurate determinations regarding whether an athlete requires immediate removal from competition or is able to safely return to activity. This article provides initial support for the VOMS as a potential option as a neurological and symptom report screening measure in a sideline environment.

ACKNOWLEDGMENT

The authors thank Emanuel Boutzoukas, Carrie Jones, Sadie Grace, and Kurt Zimmerman for their assistance with data collection and coordination with the participating high schools.

REFERENCES

1. Borg GA. Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med*. 1970;2:92-98.
2. Chelune GJ, Naugle RI, Lüders H, et al. Individual change after epilepsy surgery: practice effects and base-rate information. *Neuropsychology*. 1993;7(1):41-52.
3. Corwin DJ, Wiebe DJ, Zonfrillo MR, et al. Vestibular deficits following youth concussion. *J Pediatr*. 2015;166(5):1221-1225.
4. Covassin T, Weiss L, Powell J, Womack C. Effects of a maximal exercise test on neurocognitive function. *Br J Sports Med*. 2007;41(6):370-374.
5. Docherty CL, Valovich McLeod TC, Shultz SJ. Postural control deficits in participants with functional ankle instability as measured by the Balance Error Scoring System. *Clin J Sport Med*. 2006;16(3):203-208.
6. Erkmen N, Takin H, Kaplan T, Saniolu A. The effect of fatiguing exercise on balance performance as measured by the Balance Error Scoring System. *Isokinet Exerc Sci*. 2009;17(2):121-127.
7. Fox ZG, Mihalik JP, Blackburn JT, Battaglini CL, Guskiewicz KM. Return of postural control to baseline after anaerobic and aerobic exercise protocols. *J Athl Train*. 2008;43(5):456-463.
8. Galetta KM, Barrett J, Allen M, et al. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology*. 2011;76(17):1456-1462.
9. Galetta KM, Brandes LE, Maki K, et al. The King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort. *J Neurol Sci*. 2011;309(1-2):34-39.
10. Iverson GL, Schatz P. Advanced topics in neuropsychological assessment following sport-related concussion. *Brain Inj*. 2015;29(2):263-275.
11. King D, Brughelli M, Hume P, Gissane C. Concussions in amateur rugby union identified with the use of a rapid visual screening tool. *J Neurol Sci*. 2013;326(1-2):59-63.
12. Koscs M, Kaminski TW, Swanik CB, Edwards DG. Effects of exertional exercise on the Standardized Assessment of Concussion (SAC) score. *Athl Train Sports Health Care*. 2009;1(1):24-30.
13. Leong DF, Balcer LJ, Galetta SL, Evans G, Gimre M, Watt D. The King-Devick test for sideline concussion screening in collegiate football. *J Optom*. 2015;8(2):131-139.
14. Lovell MR, Iverson GL, Collins MW, et al. Measurement of symptoms following sports-related concussion?: reliability and normative data for the post-concussion scale. *Appl Neuropsychol*. 2006;13(3):166-174.
15. McCrea M, Kelly JP, Kluge J, Ackley B, Randolph C. Standardized Assessment of Concussion in football players. *Neurology*. 1997;48(3):586-588.
16. McCrea M, Kelly JP, Randolph C, et al. Immediate neurocognitive effects of concussion. *Neurosurgery*. 2002;50(5):1032-1042.
17. McCrea M, Kelly JP, Randolph C, et al. Standardized Assessment of Concussion (SAC): on-site mental status evaluation of the athlete. *J Head Trauma Rehabil*. 1998;13(2):27-35.
18. McCrory P, Meeuwisse WH, Aubry M, et al. Consensus statement on concussion in sport: the 4th International Conference on Concussion in Sport, Zurich, November 2012. *J Athl Train*. 2013;48(4):554-575.
19. Mucha A, Collins MW, Elbin RJ, et al. A brief Vestibular/Ocular Motor Screening (VOMS) assessment to evaluate concussions: preliminary findings. *Am J Sports Med*. 2014;42(10):2479-2486.
20. Olson BL, Oberlander TJ, Weidauer L. Test-retest reliability of the King-Devick test in an adolescent population. *J Athl Train*. 2017;52(2):439-445.
21. Rahn C, Munkasy BA, Barry Joyner A, Buckley TA. Sideline performance of the Balance Error Scoring System during a live sporting event. *Clin J Sport Med*. 2015;25(3):248-253.
22. Schatz P. Long-term test-retest reliability of baseline cognitive assessments using ImPACT. *Am J Sports Med*. 2010;38(1):47-53.
23. Sell K, Verity T, Worrell T, Pease B, Wigglesworth J. Two measurement techniques for assessing subtalar joint position: a reliability study. *J Orthop Sport Phys Ther*. 1994;19(3):162-167.

24. Spradley B, Wiriyanpinit S, Magner A. Baseline concussion testing in different environments: a pilot study. *Sport J*. 2014;19:1-8.
25. Sufrinko AM, Marchetti GF, Cohen PE, Elbin RJ, Re V, Kontos AP. Using acute performance on a comprehensive neurocognitive, vestibular, and ocular motor assessment battery to predict recovery duration after sport-related concussions. *Am J Sports Med*. 2017;45(5):1187-1194.
26. Sullivan SJ, Schneiders AG, Handcock P, Gray A, McCrory PR. Changes in the timed finger-to-nose task performance following exercise of different intensities. *Br J Sports Med*. 2011;45(1):46-48.
27. Susco TW, Valovich McLeod TC, Gansneder BM, Shultz SJ. Balance recovers within 20 minutes after exertion as measured by the Balance Error Scoring System. *J Athl Train*. 2004;39(3):241-246.
28. Ventura RE, Balcer LJ, Galetta SL. The neuro-ophthalmology of head trauma. *Lancet Neurol*. 2014;13(10):1006-1016.
29. Weber AF, Mihalik JP, Register-Mihalik JK, Mays S, Prentice WE, Guskiewicz KM. Dehydration and performance on clinical concussion measures in collegiate wrestlers. *J Athl Train*. 2013;48(2):153-160.
30. Wilkins JC, McLeod TCV, Perrin DH, Gansneder BM. Performance on the Balance Error Scoring System decreases after fatigue. *J Athl Train*. 2004;39(2):156-161.
31. Yorke AM, Smith L, Babcock M, Alsalaheen B. Validity and reliability of the Vestibular/Ocular Motor Screening and associations with common concussion screening tools. *Sports Health*. 2017;9(2):174-180.