*Journal of Athletic Training* 2024;59(6):600–607
doi: 10.4085/1062-6050-0317.23

www.natajournals.org



# False-Positive Rates, Risk Factors, and Interpretations of the Vestibular/Ocular Motor Screening in Collegiate Athletes

**Daniel Rosenblum, MEd, ATC; Catherine Donahue, PhD, ATC; Haven Higgins, BS; Madison Brna, BS; Jacob Resch, PhD, ATC**

Exercise and Sport Injury Laboratory (EaSIL), University of Virginia, Charlottesville. Dr Donahue is now at Sports Medicine Center, Children's Hospital of Colorado, Aurora.

***Context:*** Biological sex and history of motion sickness are known modifiers associated with a false-positive baseline Vestibular/Ocular Motor Screening (VOMS). However, other factors may be associated with a false-positive VOMS in collegiate athletes.

***Objective:*** To identify contributing factors to false-positive VOMS assessments using population-specific criteria. We also critically appraised previously reported interpretation criteria.

***Design:*** Descriptive laboratory study.

***Setting:*** Single-site collegiate athletic training clinic.

***Patients or Other Participants:*** National Collegiate Athletic Association Division I athletes (n = 462 [41% female]) aged 18.8 ± 1.4 years.

***Main Outcome Measure(s):*** Participants completed the Athlete Sleep Behavior Questionnaire, the 7-Item Generalized Anxiety Index, the Immediate Postconcussion Assessment and Cognitive Testing battery, the Patient Health Questionnaire-9, the Revised Head Injury Scale, the Sensory Organization Test, and the VOMS as part of a multidimensional baseline concussion assessment. Participants were classified into 2 groups based on whether they had a total symptom score of greater than or equal to 8 after VOMS administration, excluding the baseline checklist. We used $\chi^2$ and independent *t* tests to compare group demographics. A binary logistic regression with adjusted odds ratios (ORs) was used to evaluate the influence of sex, corrected vision, attention-deficit/hyperactivity disorder, Immediate Postconcussion Assessment and Cognitive Testing composite scores, concussion history, history of treatment for headache and/or migraine, Generalized Anxiety Index scores, Patient Health Questionnaire-9 scores, Athlete Sleep Behavior Questionnaire scores, and Sensory Organization Test equilibrium scores and somatosensory, visual, and vestibular sensory ratios on false-positive rates.

***Results:*** Approximately 9.1% (42 of 462 [30 females]) met criteria for a false-positive VOMS. A significantly greater proportion of females had false positives ($\chi^2_1 = 18.37$, $P < .001$). Female sex (OR = 2.79; 95% CI = 1.17, 6.65; $P = .02$) and history of treatment for headache (OR = 4.99; 95% CI = 1.21, 20.59; $P = .026$) were the only significant predictors of false-positive VOMS. Depending on cutoff interpretation, false-positive rates using our data ranged from 9.1% to 22.5%.

***Conclusions:*** Our results support the most recent interpretation guidelines for the VOMS in collegiate athletes due to a low false-positive rate and ease of interpretation. Biological sex and history of headaches should be considered when administering the VOMS in the absence of a baseline.

***Key Words:*** head injuries/concussion, athletic training, physical therapy/rehabilitation, clinical assessment/grading scales

**Key Points**

- False-positive rates in collegiate athletes have previously been found to be relatively low (approximately 10%) using a clinical cutoff described previously in an adolescent population.
- Using population-specific clinical cutoffs, the false-positive rate found in this study was similar to that found in previous work.
- Females were about 3 times more likely to exhibit a false-positive Vestibular/Ocular Motor Screening at baseline than males.

Vestibular dysfunction is one of the most reported symptoms after a sport-related concussion (SRC), with up to 84% of high school athletes reporting dizziness after injury.[1,2] Vestibular dysfunction may also manifest as vertigo, nausea, postural instability, and visual problems. Similarly, 81% (n = 200 of 247) of adolescents between 7 and 18 years of age who presented with at least 1 vestibular symptom after a diagnosed concussion took on median 53 days longer to return to school and 77 days to return to sport when compared with pediatric athletes without the same symptoms.[3] The sizable differences in time removed from school and sport emphasize the importance of evaluating for vestibular deficits after SRC.

The Vestibular/Ocular Motor Screening (VOMS) was developed to assess vestibular and ocular motor impairments via symptom provocation.[4] Previous clinical interpretation guidelines suggest that having any VOMS item symptom score greater than or equal to 2 (2 of 10) and a

near-point convergence (NPC) distance greater than or equal to 5 cm increases the probability of correctly identifying young athletes with concussion by 50% and 38%, respectively.[4] More recently, a cutoff criterion that consists of a post-VOMS symptom score greater than or equal to 8 was established in collegiate athletes.[5] A singular criterion of a symptom score provides clinicians with a simple and clear means for VOMS interpretation.[5] Independent evaluation of false-positive rates and contributing factors is warranted for this interpretation criterion in collegiate athletes.[6,7]

The reliability, discriminative validity, concurrent validity, and change scores of the VOMS have been previously described; however, as the clinical interpretation guidelines of the VOMS are variable, there is no general consensus on which guidelines to use.[4,5,7–14] The existence of multiple interpretation guidelines and a lack of consensus related to which criteria to use to interpret the VOMS is problematic. Using the interpretation criteria of a symptom score greater than or equal to 2 for any VOMS subtest and an NPC distance greater than or equal to 5 cm, the VOMS has been demonstrated to have a low false-positive rate of 11% in collegiate athletes during the preinjury (baseline) concussion assessment.[7] The false-positive rate is doubled (22%) in military personnel using these same criteria.[6] In adolescent athletes, false positives have ranged from approximately 9% to 29.3%.[9,11,15] Potential rationale for these false-positive rates include a history of motion sickness and migraine as well as an inability to accurately describe personal symptomology.[6,7,9,16,17] Though the evidence related to biological sex on recovery from concussion remains mixed, authors of previous studies have suggested that females are more likely to report suspected concussions and symptoms, which may contribute to a false-positive VOMS.[18–20] For example, female collegiate athletes are 3 times more likely to exhibit a false-positive VOMS than males.[7]

Given that motion sickness and biological sex are associated with a false-positive VOMS in healthy collegiate athletes, additional modifiers may confound the results of the assessment. The identification of additional modifiers is warranted to assist clinician understanding of potential sources of error that may influence the clinical interpretation of the VOMS before and after concussion. Moreover, the newly established clinical cutoff criteria in collegiate athletes have yet to be vetted in terms of false-positive rates in an independent sample.[5] Therefore, the purposes of this study were to (1) determine the false-positive rate at the baseline assessment, (2) identify additional risk factors that may contribute to a false-positive VOMS, and (3) compare the false-positive rate with other published clinical interpretation guidelines.

## METHODS

### Participants

Data were collected for this descriptive laboratory study between the 2019 and 2022 sport seasons as part of an ongoing prospective study evaluating SRC in collegiate athletes between 17 and 24 years of age.[21] Despite some of the data collection occurring during the COVID-19 pandemic, all participants completed the VOMS as administered before the implementation of preventative measures to mitigate the spread of the virus. Inclusion criteria consisted of being an active collegiate athlete and having complete data for all included assessments. This study was approved by the university's institutional review board and all participants provided informed consent before data collection.

### Measures

**Athlete Sleep Behavior Questionnaire.** The Athlete Sleep Behavior Questionnaire (ASBQ) is an 18-item tool that evaluates sleep practices in athletes during a 1-month period.[22] Participants respond to each item using a 5-point Likert scale: 1 = *never* to 5 = *always*, with higher composite scores indicative of worse sleep behaviors. The composite ASBQ score was used for analysis. The ASBQ has demonstrated moderate to large correlations with other validated sleep questionnaires ($r = 0.32–0.69$) and acceptable reliability (intraclass correlation coefficient [ICC] = 0.87).[22]

**Seven-Item Generalized Anxiety Index.** The 7-Item Generalized Anxiety Index (GAD-7) consists of 7 items measuring how often anxiety affects individuals during their daily lives over a 2-week period. Participants respond to each item using a 4-point Likert scale ranging from 0 = *not at all* to 3 = *nearly every day*. Higher GAD-7 composite scores are suggestive of higher anxiety. The GAD-7 composite score was used for analysis. The GAD-7 has been shown to be reliable, consistent, and valid for assessing generalized anxiety with an ICC = 0.83, Cronbach $\alpha = 0.92$, sensitivity of 89%, and specificity of 82%.[23]

**Immediate Postconcussion Assessment and Cognitive Testing.** Immediate Postconcussion Assessment and Cognitive Testing (ImPACT; ImPACT Applications) is a commonly used computerized neurocognitive test battery.[24] The ImPACT battery consists of self-reported medical history, a symptom scale, and 8 neurocognitive subtests that yield verbal memory, visual memory, visual motor speed, and reaction time outcome scores. The ImPACT baseline assessments were reviewed for validity based on the manufacturer's instructions and supplemental validity criteria.[25,26] The medical history domain was used to record the self-reported hours of sleep in the prior night, history of diagnosed attention-deficit disorder or attention-deficit/hyperactivity disorder, diagnosed concussion, and treatment for headache and/or migraine. The ImPACT has been demonstrated to have variable test-retest reliability (ICC = 0.19–0.88) and validity (ICC = 0.20–0.88) when compared with traditional neuropsychological tests.[27] The ImPACT has been demonstrated to be sensitive (79.2%–94.6%) and specific (89.4%–97.3%) to concussion during the acute phase of injury.[27,28]

**Patient Health Questionnaire-9.** The Patient Health Questionnaire-9 (PHQ-9) is a 9-item measure of depression.[29] Participants rated 9 items on the severity of their depression symptoms within the past 2 weeks using a Likert scale that ranged from 0 = *not at all* to 3 = *nearly every day*. A composite PHQ-9 score greater than or equal to 10 has been demonstrated to have equal sensitivity and specificity of 88% for diagnosing major depression.[29] The PHQ-9 composite score was used for analysis.

**Revised Head Injury Scale.** The Revised Head Injury Scale (HIS-r) is a 22-item symptom inventory specific to SRC.[30] For each endorsed item, the participant rated the duration and severity on a 6-point Likert scale based on a 24-hour period. Duration was rated from 1 = *brief* to 6 = *consistent* and severity from 0 = *not severe* to 6 = *severe*, with a potential maximum score of 132 for each qualifier. The HIS-r has been demonstrated to have a sensitivity and specificity of

77.5% and 100%, respectively, in collegiate athletes.[31] The HIS-r duration and severity composite scores were used for analysis.

**Sensory Organization Test.** The Sensory Organization Test (SOT) is a computerized balance measure that consists of somatosensory, vestibular, and visual challenges. The SOT was administered on the Natus Smart Balance Master System (NeuroCom International) or the Bertec Computerized Dynamic Posturography Unit (Bertec Corp). The Natus Smart Balance Master was discontinued during the study period and therefore the Bertec unit was then used to administer the SOT. During the SOT, participants completed 3 trials of 6 conditions for a total of 18 trials. Each trial lasted 20 seconds. The 6 SOT conditions consist of 3 different visual conditions (eyes open, eyes closed, sway-referenced visual surround) and 2 different surface conditions (fixed, sway referenced). The term *sway reference* refers to when the visual surround, the force platform, or both tilt in response to a participant's center-of-gravity sway. The SOT yields an equilibrium score and vestibular, visual, and somatosensory sensory ratios, which were all included in our analyses. The Natus SOT has been previously demonstrated to have a sensitivity of 12.8% to 72.5% and a specificity of 85.0% to 94.9% in collegiate athletes.[31,32]

**The VOMS.** The VOMS consists of 7 subtests: (1) smooth pursuits, (2) horizontal saccades, (3) vertical saccades, (4) NPC, (5) horizontal vestibular-ocular reflex, (6) vertical vestibular-ocular reflex, and (7) visual motion sensitivity.[4,5] Before and after each VOMS subtest, the patients were asked to rate their symptoms of headache, dizziness, nausea, and fogginess on a Likert scale that ranged from 0 = *symptom not present* to 10 being the most severe. The NPC was assessed by averaging the distances (centimeters) across 3 trials, with the "normal" cutoff of NPC distance reported to be less than 5 cm.[33] For this study, the total VOMS symptom score was calculated by adding the total number of symptoms reported after each subtest (excluding the baseline symptoms), with a maximum of 40 points for each subtest and 280 points for the entirety of the assessment.[5]

For this study we used population-specific interpretation criteria that consisted of a total symptom score of greater than or equal to 8.[5] This criterion was chosen as it reflects a simpler way to score and interpret the VOMS. Total symptom score has been reported to have the highest accuracy in identifying SRC compared with controls (area under the curve = 0.91).[5] Additionally, it was determined that NPC did not distinguish concussed athletes from healthy controls.

### Procedures

Upon arrival at a sports medicine clinic, all participants reviewed and provided consent before data collection. Participants then completed the ASBQ, GAD-7, HIS-r, and PHQ-9. Next, athletes were randomly divided into pairs in which they completed the SOT and VOMS individually at 2 separate stations or completed the ImPACT in groups of 2. All assessments were administered by a trained research team member on a desktop computer equipped with Microsoft Windows 10 and the latest version of Google Chrome. For ImPACT, participants were provided written, supplemental, and standardized verbal directions throughout the assessment. The total testing duration was 75 minutes.

To complete the SOT, participants were provided a standardized set of instructions before they stepped onto the force platform and were positioned in accordance with the manufacturer's guidelines.[34,35] Participants were then administered practice trials that lasted no more than 5 seconds per condition. The SOT trials were administered in a randomized order. As this was a baseline assessment, if a participant received a score for a trial that was below the manufacturer's normative data (ie, *red* trial), that trial was repeated up to 3 times until their score was within normal limits (ie, *green* trial). If after the third attempt, the trial remained below normative values, the trial was kept as a red trial.

To complete the VOMS, participants received standardized instructions prior to reporting their symptoms before and after each subtest. For the NPC, a tongue depressor with a printed, 14-point size "X" was used as the point of focus.[4] The NPC was measured (centimeters) using a tape measure as the average distance from the tip of the participant's nose to the tongue depressor across 3 trials.

### Data Analysis

Participants were divided into 2 groups based on if they did or did not have a VOMS total symptom score of greater than or equal to 8.[4,5] Group comparisons for sex and age were performed using $\chi^2$ and independent *t*-test analyses, respectively. A binary logistic regression (LR) model with adjusted odds ratios (ORs) using an enter method estimated the association between the independent variables—each outcome score from the baseline assessment—and the dependent variable—a VOMS false-positive baseline test (false-positive or true-negative). Despite entering 21 predictor variables into our model, we maintained a number of events per variable of at least 10, which has been previously described.[36] The final independent variable list included biological sex; corrected vision; history of attention-deficit disorder or attention-deficit/hyperactivity disorder; number of prior concussions; history of treatment for headache; history of treatment for migraine; treatment for anxiety and/or depression; ImPACT composite scores (verbal memory, visual memory, visual motor speed, reaction time); composite scores for the GAD-7, PHQ-9, and ASBQ; hours of sleep (via ImPACT); symptom duration (HIS-r); symptom severity (HIS-r); and the SOT equilibrium score and somatosensory, visual, and vestibular sensory ratios. Explained pseudovariance in the LR equation was determined using Nagelkerke $R^2$. All analyses were performed with SPSS (version 28.0.1; IBM Corp) with $\alpha = .05$.

## RESULTS

### Demographics

Of the original 511 National Collegiate Athletic Association Division I collegiate athletes, 42 were excluded due to missing data on one or more of the primary outcome measures, whereas 7 declined to participate. Therefore, a total of 462 (187 female [40%], 275 male [60%]) collegiate athletes were included in our analyses. Groups were similar in age (true-negative = 18.8 ± 1.4 years vs false-positive = 18.5 ± 1.1 years, $P > 0.05$) and other demographic variables (Table 1). Medians and interquartile ranges for each continuous risk factor included in our model across groups are presented in Table 2.

**Table 1. Medical History Demographics per Group**

| Group | No. (% of Group) | |
|---|---|---|
| | True-Negative (n = 420) | False-Positive (n = 42) |
| ADD/ADHD | 30 (7) | 14 (3) |
| Headache | 7 (16) | 8 (19) |
| Migraine | 18 (4) | 4 (9) |
| Anxiety/depression | 27 (6) | 9 (21) |
| Concussion history | 101 (24) | 17 (40) |
| 1 concussion | 65 (15) | 13 (31) |
| 2 concussions | 27 (6) | 3 (7) |
| 3 concussions | 8 (2) | 1 (2) |
| 4 concussions | 1 (0.2) | 0 (0) |

Abbreviations: ADD, attention-deficit disorder; ADHD, attention-deficit/hyperactivity disorder.

**Table 2. Medians and Interquartile Ranges (IQRs) for Continuous Risk Factors per Group**

| Risk Factor | Median (IQR) | |
|---|---|---|
| | False-Positive | True-Negative |
| VOMS | | |
| Total symptom score | 15.00 (10.75) | 0.00 (0.00) |
| NPC, cm | 1.50 (2.79) | 1.83 (3.83) |
| ImPACT composite scores | | |
| Verbal memory | 91.50 (12.00) | 93.00 (11.00) |
| Visual memory | 82.00 (15) | 83.00 (15.00) |
| Visual motor speed | 41.08 (7.83) | 42.96 (8.56) |
| Reaction time | 0.57 (0.06) | 0.54 (0.08) |
| Patient-reported outcomes | | |
| Hours of sleep | 7.00 (1.38) | 7.00 (1.00) |
| GAD-7 total | 2.50 (5.75) | 0.00 (1.00) |
| ASBQ total | 38.00 (9.75) | 35.00 (8.25) |
| HIS-r duration | 4.00 (10.25) | 0.00 (2.00) |
| HIS-r severity | 2.50 (7.00) | 0.00 (1.00) |
| PHQ-9 total | 2.00 (3.00) | 0.00 (1.00) |
| Sensory Organization Test | | |
| Equilibrium score | 80.00 (7.46) | 82.00 (5.21) |
| Somatosensory ratio | 97.44 (3.45) | 97.19 (3.18) |
| Visual ratio | 88.58 (9.87) | 92.03 (6.50) |
| Vestibular ratio | 77.45 (11.71) | 78.34 (10.61) |

Abbreviations: ASBQ, Athlete Sleep Behavior Questionnaire; GAD-7, Generalized Anxiety Index-7; HIS-r, Revised Head Injury Scale; ImPACT, Immediate Postconcussion Assessment and Cognitive Testing; NPC, near-point convergence; PHQ-9, Patient Health Questionnaire-9; VOMS, Vestibular/Ocular Motor Screening.

### False-Positive Rates

Approximately 9.1% of participants (12 males, 30 females) demonstrated a false-positive VOMS assessment. Among participants with false positives, there was a significantly greater proportion of females (n = 462, $\chi^2 = 18.37$, $P < .001$, 4% males vs 16% females). Table 3 summarizes the false-positive rates in the current study as well as previously reported false-positive rates in relationship to different cutoff criteria and samples. Table 4 demonstrates how false-positive rates can vary because of different cutoff interpretations, using the data from this study.

### Multivariable LR

The overall binary LR was statistically significant ($P < .001$) and accounted for about 36% of the total variance ($R^2 = 0.355$). Sex (OR = 2.79; 95% CI = 1.17, 6.65; $P = .020$) and history of treatment for headache (OR = 4.99; 95% CI = 1.21, 20.59; $P = .026$) were the only significant predictors of a false-positive VOMS assessment. Results of the LR model are presented in Table 5.

## DISCUSSION

The primary aim of our study was to cross-validate population-specific VOMS clinical interpretation guidelines. A secondary aim of this study was to determine contributing factors to a false-positive VOMS based on these and other population-specific interpretation criteria (Tables 3 and 4).[6,7,11,13,16,37] The overall false-positive rate using newly established cutoff criteria in collegiate athletes was 9.1%, which is consistent with prior research.[7] We observed that being a female and receiving prior treatment for headache were associated with an increased risk for a false-positive VOMS. Lastly, we determined that false-positive rates are largely dependent on interpretation guidelines. Our findings should be considered when administering the VOMS, especially in the absence of a baseline assessment.

Reported false-positive rates of the VOMS range from 9% to 29.3% in adolescents to 21.9% in US military personnel and 11% in collegiate athletes based on criteria established in young athletes.[6,7,9,11] For this study, we applied an interpretation criterion (≥8 total symptoms) specific to collegiate athletes.[5] Using this criterion, our sample of healthy collegiate athletes had a low false-positive rate of 9.1%, which is consistent with previous research.[6,7,9,11] Given the low false-positive rate observed in this study and the previously published high area under the curve (0.91), clinicians that work with collegiate athletes should consider using the total symptom score (≥8) to interpret the VOMS.[5] This is the first study to cross-validate the cutoff score reported by Kontos et al in concussed collegiate athletes and is in partial alignment with the recently published Sport Concussion Office Assessment Tool-6, as it does not include the NPC.[5,38] In sum, using a total symptom score (≥8) is a relatively easy way to interpret the VOMS with an acceptable false-positive rate and in conjunction with the high discriminative ability found previously in collegiate athletes.[5]

Biological sex and a history of motion sickness are known risk factors for false-positive VOMS performance in collegiate athletes. Kontos and colleagues reported that 72% of athletes who exceeded clinical cutoffs reported a history of motion sickness and that female athletes were 3 times more likely to have one or more VOMS items above the clinical cutoff, leading to a false-positive outcome.[7] Similarly, Henry et al observed that males reported fewer vestibular-ocular symptoms when using the VOMS within 2 weeks of a diagnosed concussion.[39] Given the consistency of biological sex differences with this study and related research associated with vestibular-ocular function after concussion, we believe this warrants further discussion.

As mentioned in our introduction, the influence of female sex on recovery from concussion remains equivocal.[18] That said, some evidence suggests that high school female athletes are more likely to develop persisting symptoms (ie, lasting greater than 28 days) after a concussion.[40–42] In this study, biological sex was identified as a key factor associated with a false-positive assessment. This finding may be due to females being more likely to report symptoms.[43] Brown et al

**Table 3. Previously Established Vestibular/Ocular Motor Screening False-Positive Rates**

| Population | Cutoff Criteria | False-Positive Rate, % | Source |
|---|---|---|---|
| Adolescent | ≥2 symptoms on any subtest *or* an NPC ≥5 cm (original cutoff)[a] | N/A | Mucha et al,[4] 2014 |
| Adolescent | ≥2 symptoms on any subtest *and/or* NPC ≥5 cm | 9–16 | Moran et al,[9] 2018; Moran et al,[15] 2019 |
| Adolescent | ≥2 symptoms on any subtest *or* NPC ≥5 cm | 29.3 | Iverson et al,[11] 2021 |
| Military | ≥2 symptoms on any subtest *and* NPC ≥5 cm | 21.9 | Kontos et al,[6] 2022 |
| Collegiate | ≥2 symptoms on any subtest *or?* NPC ≥5 cm (uses *and* but is in an *or* context) | 11 | Kontos et al,[7] 2016 |
| Collegiate | ≥8 total symptom score *without* NPC | 9.1 | Current study |

Abbreviation: NPC, near-point convergence.
[a] According to the conclusion section of the study.

concluded that healthy female athletes were more likely to report vision or hearing problems, headache/migraine, and difficulty concentrating during baseline assessments.[43] These symptoms are consistent with those of the VOMS. That female athletes are more willing to disclose symptoms may be a result of psychology but also of physiology in the absence of a concussion. For example, the menstrual cycle may be associated with vestibular-ocular symptom reporting in the absence of a concussion. Symptoms such as headache and difficulty concentrating have been associated with premenstrual syndrome and are similar to the VOMS symptoms headache and fogginess.[44] Nausea, a symptom recorded as part of the VOMS, and motion sickness are predictors of a false-positive VOMS. Each symptom has been demonstrated to be significantly higher in collegiate females who do not use oral contraceptives during symptom provocation tasks similar to the VOMS.[45] Accounting for the menstrual cycle at the baseline and postinjury assessments is important and may assist clinicians in determining if symptoms provoked during the administration of VOMS are due to vestibular-ocular dysfunction or another cause. Future researchers should consider the development of sex-specific cutoffs to ensure that clinicians are providing individualized care to their patients in accordance with the biopsychosocial model of health care.

Interestingly, a history of treatment for headaches was also a significant predictor in this study; however, the confidence interval was relatively large and thus this result should be interpreted with caution. Despite the inclusion of multiple predictors in our regression model, no other modifiers were identified. Surprisingly, the SOT outcome scores were not associated with false-positive VOMS assessments. The SOT was included as a predictor in our analysis as the vestibulo-ocular and vestibulospinal systems are challenged during the assessment. Though the SOT outcome scores were not associated with a false-positive VOMS, this does not imply that the VOMS and SOT are always measuring different constructs. However, authors of 2 separate studies concluded that the VOMS measured different aspects of vestibular function when compared with the Balance Error Scoring System and King-Devick assessments, which also purport to measure vestibulo-ocular and vestibulospinal function.[4,10] Future researchers should investigate the relationship between the SOT and VOMS outcome scores.

**Table 4. False-Positive Rates Associated With Different Cutoff Interpretations in Current Study**

| Clinical Cutoff | Clinical Classification | No. | Percentage[a] |
|---|---|---|---|
| ≥8 total symptoms | True-negative | 420 | 90.9 |
| | False-positive | 42 | **9.1** |
| ≥8 total symptoms *AND* NPC ≥5 cm | True-negative | 397 | 85.9 |
| | False-positive | 65 | **14.1** |
| ≥8 total symptoms *OR* NPC ≥5 cm | True-negative | 358 | 77.5 |
| | False-positive | 104 | **22.5** |
| NPC ≥5 cm only | True-negative | 392 | 84.8 |
| | False-positive | 70 | **15.2** |
| NPC >5 cm only | True-negative | 396 | 85.7 |
| | False-positive | 66 | **14.3** |

Abbreviation: NPC, near-point convergence.
[a] False-positive rates are bolded.

Lastly, we compared different VOMS cutoff interpretation guidelines in terms of false-positive rates. To date, the most frequently used clinical cutoff to evaluate VOMS performance has been a symptom score greater than or equal to 2 for any domain or an NPC of greater than or equal to 5 cm.[4] However, discrepancies exist in terms of the wording of these clinical interpretation guidelines. The underlying issues associated with variable clinical interpretation criteria are based on the use of *and*, *or*, and *and/or*. For example, when discussing the clinical cutoffs in their abstract, Mucha et al stated: "An NPC distance ≥ 2 5 cm and any VOMS item symptoms score ≥ 2 resulted in an increase in the probability of correctly identifying concussed patients of 38% and 50%, respectively."[4] In their conclusion, Mucha and colleagues stated that "cutoff scores of ≥ 2 total symptoms after any VOMS item or an NPC distance of ≥ 5 cm resulted in high rates of identifying concussions."[4] Although those sentences technically mean the same thing, it may lead to confusion in the clinical interpretation of the VOMS. Similarly, articles have reported using the symptom score *and/or* as well as *or* NPC (Table 3).[6,7,9,11,16] For example, authors of 1 study explicitly mentioned using greater than or equal to 2 symptoms *and/or* a NPC of greater than or equal to 5 cm; however, in the same article the authors reported that "clinical cutoff scores are defined as ≥2 symptoms on VOMS subscales or ≥5 cm on NPC distance."[9] These criteria can be interpreted in 2 distinct ways:

1. An individual would need to be above *both* cutoffs to be considered positive.
2. An individual would need to be above *only one* cutoff to be considered positive.

These alternate interpretations based on *and/or* versus just *or* may change interpretation and decision-making, though the VOMS should not be used in isolation for clinical decision-making after a concussion.

**Table 5. Results of Binary Logistic Regression**

| Predictor | Coefficient | *P* Value | Odds Ratio | 95% CI | |
|---|---|---|---|---|---|
| | | | | Lower | Upper |
| Demographics | | | | | |
| Sex (female) | 1.02 | .020[a] | 2.79 | 1.17 | 6.65 |
| Corrected vision | 0.56 | .163 | 1.75 | 0.79 | 3.85 |
| ADD/ADHD | 0.65 | .260 | 1.93 | 0.61 | 6.05 |
| Treatment for headache | 1.60 | .026[a] | 4.99 | 1.21 | 20.59 |
| Treatment for migraine | −0.49 | .591 | 0.60 | 0.09 | 3.74 |
| Depression/anxiety | −0.59 | .361 | 0.55 | 0.15 | 1.97 |
| Concussion history | 0.67 | .115 | 1.95 | 0.84 | 4.50 |
| ImPACT composites | | | | | |
| Verbal memory | 0.01 | .602 | 1.01 | 0.96 | 1.05 |
| Visual memory | −0.01 | .589 | 0.99 | 0.95 | 1.02 |
| Visual motor speed | −0.05 | .194 | 0.94 | 0.87 | 1.02 |
| Reaction time | 0.21 | .953 | 1.24 | 0.00 | 1608.87 |
| Patient-reported outcomes | | | | | |
| GAD-7 total | 0.10 | .235 | 1.10 | 0.93 | 1.30 |
| ASBQ total | 0.04 | .154 | 1.04 | 0.98 | 1.11 |
| Hours of sleep (ASBQ) | 0.14 | .366 | 1.15 | 0.84 | 1.59 |
| HIS-r duration | 0.16 | .239 | 1.18 | 0.89 | 1.55 |
| HIS-r severity | −0.09 | .559 | 0.90 | 0.64 | 1.26 |
| PHQ-9 total | 0.16 | .121 | 1.17 | 0.95 | 1.44 |
| Sensory Organization Test | | | | | |
| Equilibrium score | −0.12 | .142 | 0.88 | 0.74 | 1.04 |
| Somatosensory ratio | 0.03 | .705 | 1.03 | 0.87 | 1.22 |
| Visual ratio | −0.001 | .982 | 0.99 | 0.89 | 1.11 |
| Vestibular ratio | 0.01 | .362 | 1.01 | 0.97 | 1.06 |

Abbreviations: ADD, attention-deficit disorder; ADHD, attention-deficit/hyperactivity disorder; ASBQ, Athlete Sleep Behavior Questionnaire; GAD-7, Generalized Anxiety Index-7; HIS-r, Revised Head Injury Scale; PHQ-9, Patient Health Questionnaire-9.
[a] Indicates statistical significance at $P < .05$.

A separate study reported a false-positive rate of 11% based on symptom scores less than 2 on all VOMS items; however, the authors seemed to have included the NPC distance in their original operational definition of what constituted a false-positive.[7] A discrepancy exists between the results found in Table 3 within that manuscript, as the table shows 11% based on NPC distance only, not symptoms.[7] Another discrepancy in the literature related to the VOMS includes interpreting the cutoffs as a function of a symptom provocation change score from baseline, despite the Mucha et al article's not accounting for a change from the baseline assessment.[4,46,47] Additionally, in terms of the NPC distance cutoff, researchers have reported using both greater than and greater than or equal to 5 cm, sometimes even within the same study.[6,16] Table 5 provides examples of how these different verbiages influence the false-positive rates when applied to the data of this study. Overall, these observations reflect small, but important, changes and inconsistencies in verbiage that can make it difficult to interpret the VOMS and to compare findings across multiple studies. Our findings support a simple and cross-validated way to interpret the VOMS.[5]

This study is not without limitations. As with any study involving patient-reported outcomes, recall bias may have influenced our demographic results. It is unclear whether the order of tests, which was relatively random based on clinician availability, may have influenced our results. For example, it has been shown that strenuous exercise can influence VOMS results.[48] However, our participants were tested in small groups of no more than 4 at a time (tested individually during the VOMS assessment) by 1 of 3 clinicians who were all trained in the administration of the VOMS by the same individual. The SOT was administered on 1 of 2 balance units due to the discontinuation of one (Natus) and an upgraded unit (Bertec) being installed during the study. It is unclear whether the use of 2 different units administering the SOT had any significant effect on the SOT outcome scores, as the Bertec Computerized Dynamic Posturography unit's measurement properties have yet to be established in collegiate athletes. That said, the SOT guidelines and algorithms are the same between the 2 units. Additionally, we did not query athletes about a history of motion sickness, a known risk factor that may also have contributed to misclassification.[6,7] As this study is part of an ongoing prospective study, motion sickness was not queried because the ongoing study was not specific to false-positive VOMS assessments a priori; however, the addition of an item asking about a history of motion sickness is warranted if the screen is to be used. Said again, as we did not evaluate the role of motion sickness as a risk for a false-positive VOMS, we cannot support or refute the findings of related literature. Regarding our LR model, there were relatively few individuals in the false-positive group given the number of predictors entered. Importantly, balancing our data is not necessarily the only valid approach in this instance. If one of the classes (ie, true negatives) is actually much more common in the population (which it is according to a variety of other VOMS studies examining false-positive rates) and not solely in our sample, a naïve model (ie, classifying everything as belonging to the most common category [ie, true-positive]) can actually be a good estimate. In other words, although there are few false-positive cases, we feel that the model provides a good idea of what "normal" is, given the large number of true negatives. Additionally, history of headache as a predictor has a wide

confidence interval and so that finding should be interpreted with caution. Given that the only other predictor was sex, and that sex differences have been found in previous work, we feel that, although the model might not be perfect, it gives a good picture of what might classify someone as false-positive. Lastly, we did not assess interrater or intrarater reliability of VOMS assessors; however, all assessors were certified athletic trainers who were instructed by the same individual in the administration of the VOMS.

## CONCLUSIONS

When using VOMS interpretation guidelines specific to collegiate athletes, the false-positive rate was similar to that found in previous research. Female athletes and athletes with a history of treatment for headaches had a higher false-positive rate than those who did not. Biological sex and prior treatment for headaches should be considered when interpreting the VOMS before and after a concussion. Understanding what an athlete is "trying to get back to" after a concussion when using the VOMS coupled with the ability to detect potential vestibular dysfunction in the absence of concussion emphasizes the importance of the VOMS baseline assessment if it is to be used clinically. Our findings support using population-specific interpretation criteria for the VOMS, as false-positive rates vary when alternate interpretation guidelines are used. Lastly, clinicians and researchers should consider their verbiage when incorporating the VOMS into concussion protocols for the purposes of simplified and consistent interpretation. The use or misuse of *and*, *or*, or *and/or* may change the interpretation of the VOMS and influence clinical decision-making after a concussion.

## REFERENCES


1. Lau BC, Kontos AP, Collins MW, Mucha A, Lovell MR. Which on-field signs/symptoms predict protracted recovery from sport-related concussion among high school football players? *Am J Sports Med*. 2011;39(11):2311–2318. doi:10.1177/0363546511410655
2. Merritt VC, Rabinowitz AR, Arnett PA. Injury-related predictors of symptom severity following sports-related concussion. *J Clin Exp Neuropsychol*. 2015;37(3):265–275. doi:10.1080/13803395.2015.1004303
3. Corwin DJ, Wiebe DJ, Zonfrillo MR, et al. Vestibular deficits following youth concussion. *J Pediatr*. 2015;166(5):1221–1225. doi:10.1016/j.jpeds.2015.01.039
4. 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. doi:10.1177/0363546514543775
5. Kontos AP, Eagle SR, Marchetti G, et al; CARE Consortium Site Investigators. Discriminative validity of vestibular ocular motor screening in identifying concussion among collegiate athletes: a National Collegiate Athletic Association–Department of Defense Concussion Assessment, Research, and Education Consortium study. *Am J Sports Med*. 2021;49(8):2211–2217. doi:10.1177/03635465211012359
6. Kontos AP, Monti MK, Eagle SR, et al. False-positive rates and associated risk factors on the Vestibular-Ocular Motor Screening and Modified Balance Error Scoring System in US military personnel. *J Athl Train*. 2022;57(5):458–463. doi:10.4085/1062-6050-0094.21
7. Kontos AP, Sufrinko A, Elbin RJ, Puskar A, Collins MW. Reliability and associated risk factors for performance on the Vestibular/Ocular Motor Screening (VOMS) tool in healthy collegiate athletes. *Am J Sports Med*. 2016;44(6):1400–1406. doi:10.1177/0363546516632754
8. Kontos AP, Monti K, Eagle SR, et al. Test-retest reliability of the Vestibular Ocular Motor Screening (VOMS) tool and modified Balance Error Scoring System (mBESS) in US military personnel. *J Sci Med Sport*. 2021;24(3):264–268. doi:10.1016/j.jsams.2020.08.012
9. Moran RN, Covassin T, Elbin RJ, Gould D, Nogle S. Reliability and normative reference values for the Vestibular/Ocular Motor Screening (VOMS) tool in youth athletes. *Am J Sports Med*. 2018;46(6):1475–1480. doi:10.1177/0363546518756979
10. 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. doi:10.1177/1941738116678411
11. Iverson GL, Cook NE, Howell DR, et al. Preseason vestibular ocular motor screening in children and adolescents. *Clin J Sport Med*. 2021;31(4):e188–e192. doi:10.1097/JSM.0000000000000767
12. Eagle SR, Feder A, Manderino LM, et al. Concurrent validity of the Vestibular/Ocular Motor Screening (VOMS) tool with the Dizziness Handicap Inventory (DHI) among adolescents with vestibular symptoms/impairment following concussion. *Phys Ther Sport*. 2022;53:34–39. doi:10.1016/j.ptsp.2021.11.003
13. Eagle SR, Ferris LM, Mucha A, et al. Minimum detectable change and false positive rates of the Vestibular/Ocular Motor Screening (VOMS) tool: an NCAA-DoD Care Consortium analysis. *Brain Inj*. 2021;35(12–13):1563–1568. doi:10.1080/02699052.2021.1973561
14. Elbin RJ, Eagle SR, Marchetti GF, et al. Using change scores on the Vestibular Ocular Motor Screening (VOMS) tool to identify concussion in adolescents. *Appl Neuropsychol Child*. 2022;11(4):591–597. doi:10.1080/21622965.2021.1911806
15. Moran RN, Covassin T, Elbin RJ. Sex differences on vestibular and ocular motor assessment in youth athletes. *J Athl Train*. 2019;54(4):445–448. doi:10.4085/1062-6050-220-18
16. Moran RN, Covassin T, Wallace J. Premorbid migraine history as a risk factor for vestibular and oculomotor baseline concussion assessment in pediatric athletes. *J Neurosurg Pediatr*. 2019;23(4):465–470. doi:10.3171/2018.10.PEDS18425
17. Ellis MJ, Leddy JJ, Willer B. Physiological, vestibulo-ocular and cervicogenic post-concussion disorders: an evidence-based classification system with directions for treatment. *Brain Inj*. 2015;29(2):238–248. doi:10.3109/02699052.2014.965207
18. Iverson GL, Gardner AJ, Terry DP, et al. Predictors of clinical recovery from concussion: a systematic review. *Br J Sports Med*. 2017;51(12):941–948. doi:10.1136/bjsports-2017-097729
19. Broshek DK, Kaushik T, Freeman JR, Erlanger D, Webbe F, Barth JT. Sex differences in outcome following sports-related concussion. *J Neurosurg*. 2005;102(5):856–863. doi:10.3171/jns.2005.102.5.0856
20. Wallace J, Covassin T, Beidler E. Sex differences in high school athletes' knowledge of sport-related concussion symptoms and reporting behaviors. *J Athl Train*. 2017;52(7):682–688. doi:10.4085/1062-6050-52.3.06
21. Thompson XD, Newman TM, Donahue CC, Erdman NK, Statuta SM, Resch JE. Kinesiophobia is related to acute musculoskeletal injury incidence following concussion. *J Sport Rehabil*. 2022;32(2):145–150. doi:10.1123/jsr.2022-0134
22. Driller MW, Mah CD, Halson SL. Development of the Athlete Sleep Behavior Questionnaire: a tool for identifying maladaptive sleep practices in elite athletes. *Sleep Sci*. 2018;11(1):37–44. doi:10.5935/1984-0063.20180009
23. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092–1097. doi:10.1001/archinte.166.10.1092
24. Meehan WP III, d'Hemecourt P, Collins CL, Taylor AM, Comstock RD. Computerized neurocognitive testing for the management of sport-related concussions. *Pediatrics*. 2012;129(1):38–44. doi:10.1542/peds.2011-1972
25. Walton SR, Broshek DK, Freeman JR, Cullum CM, Resch JE. Valid but invalid: suboptimal ImPACT baseline performance in university athletes. *Med Sci Sports Exerc*. 2018;50(7):1377–1384. doi:10.1249/MSS.0000000000001592

26. ImPact, version 4. Administration and interpretation manual. ImPACT Applications. Published 2022. Accessed February 19, 2024. https://impacttest.com/wp-content/uploads/LBL-01_v14_ImPACT-Version-4_Administration_Manual.pdf
27. Resch JE, McCrea MA, Cullum CM. Computerized neurocognitive testing in the management of sport-related concussion: an update. *Neuropsychol Rev*. 2013;23(4):335–349. doi:10.1007/s11065-013-9242-5
28. Czerniak LL, Liebel SW, Garcia GP, et al; CARE Consortium Investigators. Sensitivity and specificity of computer-based neurocognitive tests in sport-related concussion: findings from the NCAA-DoD CARE Consortium. *Sports Med*. 2021;51(2):351–365. doi:10.1007/s40279-020-01393-7
29. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606–613. doi:10.1046/j.1525-1497.2001.016009606.x
30. Piland SG, Motl RW, Ferrara MS, Peterson CL. Evidence for the factorial and construct validity of a self-report concussion symptoms scale. *J Athl Train*. 2003;38(2):104–112.
31. Resch JE, Brown CN, Schmidt J, et al. The sensitivity and specificity of clinical measures of sport concussion: three tests are better than one. *BMJ Open Sport Exerc Med*. 2016;2(1):e000012. doi:10.1136/bmjsem-2015-000012
32. Register-Mihalik JK, Guskiewicz KM, Mihalik JP, Schmidt JD, Kerr ZY, McCrea MA. Reliable change, sensitivity, and specificity of a multidimensional concussion assessment battery: implications for caution in clinical practice. *J Head Trauma Rehabil*. 2013;28(4):274–283. doi:10.1097/HTR.0b013e3182585d37
33. Scheiman M, Gallaway M, Frantz KA, et al. Nearpoint of convergence: test procedure, target selection, and normative data. *Optom Vis Sci*. 2003;80(3):214–225. doi:10.1097/00006324-200303000-00011
34. Roller M, Boismier T, Krzak J. Balance Manager Clinical Interpretation Guide; NeuroCom International; 2009.
35. Bertec Balance Advantage Dynamic System user manual. Interacoustics. Published 2014. Accessed March 7, 2023. https://www.interacoustics.com/images/files/80P-0002_BER_DynamicUserManual_2021-06.pdf
36. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49(12):1373–1379. doi:10.1016/s0895-4356(96)00236-3
37. Moran RN, Wallace J, Murray NG, Covassin T. Effects of attention deficit hyperactivity disorder and learning disability on vestibular and ocular baseline concussion assessment in pediatric athletes. *Appl Neuropsychol Child*. 2021;10(3):276–282. doi:10.1080/21622965.2019.1683453
38. Patricios JS, Davis GA, Ahmed OH, et al. Introducing the Sport Concussion Office Assessment Tool 6 (SCOAT6). *Br J Sports Med*. 2023;57(11):648–650. doi:10.1136/bjsports-2023-106860
39. Henry LC, Elbin RJ, Collins MW, Marchetti G, Kontos AP. Examining recovery trajectories after sport-related concussion with a multimodal clinical assessment approach. *Neurosurgery*. 2016;78(2):232–241. doi:10.1227/NEU.0000000000001041
40. Miller JH, Gill C, Kuhn EN, et al. Predictors of delayed recovery following pediatric sports-related concussion: a case-control study. *J Neurosurg Pediatr*. 2016;17(4):491–496. doi:10.3171/2015.8.PEDS14332
41. Morgan CD, Zuckerman SL, Lee YM, et al. Predictors of postconcussion syndrome after sports-related concussion in young athletes: a matched case-control study. *J Neurosurg Pediatr*. 2015;15(6):589–598. doi:10.3171/2014.10.PEDS14356
42. Zemek R, Barrowman N, Freedman SB, et al; Pediatric Emergency Research Canada (PERC) Concussion Team. Clinical risk score for persistent postconcussion symptoms among children with acute concussion in the ED. *JAMA*. 2016;315(10):1014–1025. doi:10.1001/jama.2016.1203
43. Brown DA, Elsass JA, Miller AJ, Reed LE, Reneker JC. Differences in symptom reporting between males and females at baseline and after a sports-related concussion: a systematic review and meta-analysis. *Sports Med*. 2015;45(7):1027–1040. doi:10.1007/s40279-015-0335-6
44. Schmelzer K, Ditzen B, Weise C, Andersson G, Hiller W, Kleinstäuber M. Clinical profiles of premenstrual experiences among women having premenstrual syndrome (PMS): affective changes predominate and relate to social and occupational functioning. *Health Care Women Int*. 2015;36(10):1104–1123. doi:10.1080/07399332.2014.954701
45. Matchock RL, Levine ME, Gianaros PJ, Stern RM. Susceptibility to nausea and motion sickness as a function of the menstrual cycle. *Womens Health Issues*. 2008;18(4):328–335. doi:10.1016/j.whi.2008.01.006
46. Büttner F, Howell DR, Doherty C, Blake C, Ryan J, Delahunt E. Clinical detection and recovery of vestibular and oculomotor impairments among amateur athletes following sport-related concussion: a prospective, matched-cohort study. *J Head Trauma Rehabil*. 2021;36(2):87–95. doi:10.1097/HTR.0000000000000608
47. Ferris LM, Kontos AP, Eagle SR, et al. Predictive accuracy of the Sport Concussion Assessment Tool 3 and Vestibular/Ocular-Motor Screening, individually and in combination: a National Collegiate Athletic Association-Department of Defense Concussion Assessment, Research and Education Consortium analysis. *Am J Sports Med*. 2021;49(4):1040–1048. doi:10.1177/0363546520988098. Published correction appears in *Am J Sports Med*. 2021;49(13):NP66–NP67.
48. Ratka J, Cheever K, Mansell JL, Tierney RT. The effect of an interval fatigue protocol on Vestibular/Ocular Motor Screening (VOMS) performance. *Brain Inj*. 2020;34(1):110–114. doi:10.1080/02699052.2019.1682194

*Address correspondence to Daniel Rosenblum, MEd, ATC, Exercise and Sport Injury Laboratory (EaSIL), University of Virginia, 550 Brandon Avenue, Charlottesville, VA 22903. Address email to dr6gz@virginia.edu.*