Maximizing the Accuracy of Adolescent Concussion Diagnosis Using Individual Elements of Common Standardized Clinical Assessment Tools

Daniel J. Corwin, MD, MSCE*†; Francesca Mandel, PhD‡; Catherine C. McDonald, PhD, RN, FAAN*§; Fairuz N. Mohammed, MPH*; Susan Margulies, PhDII; Ian Barnett, PhD‡; Kristy B. Arbogast, PhD*†; Christina L. Master, MD*¶

*Center for Injury Research and Prevention, Children's Hospital of Philadelphia, PA; †Division of Emergency Medicine, Children's Hospital of Philadelphia, PA; ‡Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia; §School of Nursing, University of Pennsylvania, Philadelphia; IIDepartment of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta; ¶Sports Medicine and Performance Center, Children's Hospital of Philadelphia, PA

Context: Multiple clinical evaluation tools exist for adolescent concussion with various degrees of correlation, presenting challenges for clinicians in identifying which elements of these tools provide the greatest diagnostic utility.

Objective: To determine the combination of elements from 4 commonly used clinical concussion batteries that maximize discrimination of adolescents with concussion from those without concussion.

Design: Cross-sectional study.

Setting: Suburban school and concussion program of a tertiary care academic center.

Patients or Other Participants: A total of 231 participants with concussion (from a suburban school and a concussion program) and 166 participants without concussion (from a suburban school) between the ages of 13 and 19 years.

Main Outcome Measure(s): Individual elements of the visio-vestibular examination (VVE), Sport Concussion Assessment Tool, fifth edition (SCAT5; including the modified Balance Error Scoring System), King-Devick test (K-D), and Postconcussion Symptom Inventory (PCSI) were evaluated. The 24 subcomponents of these tests were grouped into interpretable factors using sparse principal component analysis. The 13 resultant factors were combined with demographic and clinical covariates into a logistic regression model and ranked by frequency of inclusion into the ideal model, and the predictive

performance of the ideal model was compared with each of the clinical batteries using the area under the receiver operating characteristic curve (AUC).

Results: A cluster of 4 factors (factor 1 [VVE saccades and vestibulo-ocular reflex], factor 2 [modified Balance Error Scoring System double-legged stance], factor 3 [SCAT5/PCSI symptom scores], and factor 4 [K-D completion time]) emerged. A model fit with the top factors performed as well as each battery in predicting concussion status (AUC = 0.816 [95% CI = 0.731, 0.889]) compared with the SCAT5 (AUC = 0.784 [95% CI = 0.692, 0.866]), PCSI (AUC = 0.776 [95% CI = 0.674, 0.863]), VVE (AUC = 0.711 [95% CI = 0.602, 0.814]), and K-D (AUC = 0.708 [95% CI = 0.590, 0.819]).

Conclusions: A multifaceted assessment for adolescents with concussion, comprising symptoms, attention, balance, and the visio-vestibular system, is critical. Current diagnostic batteries likely measure overlapping domains, and the sparse principal component analysis demonstrated strategies for streamlining comprehensive concussion assessment across a variety of settings.

Key Words: visio-vestibular examination, Sport Concussion Assessment Tool, King-Devick test, Postconcussion Symptom Inventory

Key Points

- A multifaceted concussion assessment for injured adolescents is key, including assessment of symptoms, attention, balance, and the visio-vestibular system.
- A model including key components of 4 common concussion batteries performed as well as any individual battery in distinguishing adolescents with concussion from those without concussion.
- Current concussion diagnostic batteries likely measure overlapping concepts. The analytical approach identified strategies for streamlining concussion assessment across a variety of settings.

oncussion is a common injury in adolescents that has heterogeneous presentations and signs and symptoms, spanning the somatic, visio-vestibular, cognitive, emotional, and sleep domains,¹ ultimately manifesting as multiple phenotypes.² Although advances have been made in our understanding of the physiological changes that occur after injury,^{3,4} the mainstay of diagnosis according to both expert statements and consensus guidelines remains symptom based,^{1,5,6} with symptoms often assessed via standardized scales such as the Postconcussion Symptom Inventory (PCSI).⁷ To address the multiple phenotypes of concussion and augment symptom evaluation, a multifaceted approach to concussion diagnosis using multiple physiological tools beyond symptom assessment is recommended.^{1,5,6} These tools include the Sport Concussion Assessment Tool, fifth edition (SCAT5),8 with measures of attention, memory, and balance; tests of visual and cognitive function, such as the King-Devick (K-D) test⁹; and assessment of visio-vestibular functioning, including the visio-vestibular examination (VVE)¹⁰ or the Vestibular/ Oculomotor Motor Screening (VOMS) examination.¹¹ Individually, these clinical batteries have demonstrated variable diagnostic performance for adolescent concussion.9,11 Evaluations of these batteries in adolescents without concussion have also shown varying proportions with deficits when not injured, indicating that multidimensional testing should be used to identify true injury.^{12,13}

Researchers have demonstrated that these batteries and their subcomponents have different levels of correlation and overlap,^{12,14} and performing all 4 of the above batteries (PCSI, SCAT5, K-D, and VVE) for a single patient could take a provider up to 30 minutes to complete. Given the time constraints to concussion assessment, knowing which elements of each clinical battery provide the greatest diagnostic utility would be useful for clinicians attempting to diagnose a concussion, allowing them to streamline their assessment by reducing the number of elements performed. Therefore, the ideal combination of clinical testing elements for concussion diagnosis needs to be identified. Previous investigators have compared the performance of the individual elements of these batteries in college-aged athletes, leading to an algorithm for using the individual components of the SCAT to diagnose concussion,¹⁵ with vestibular testing augmenting symptom scores in distinguishing young adults with concussion from those without concussion.¹⁶ However, the optimal testing battery for adolescents, who have the highest concussion risk, is unknown.¹⁷

The purpose of our study was to use sparse principal component analysis with regression modeling to determine the combination of elements from 4 commonly used clinical concussion batteries (VVE, SCAT5, K-D, and PCSI) that maximize discrimination of adolescents with and those without concussion, thereby reducing the overall number of components required to evaluate adolescents with concussion. A secondary purpose was comparing the discriminatory ability of each of the 4 batteries with the ideal combined model.

METHODS

Participants

We recruited participants aged 13 to 19 years between August 2017 and October 2020 as part of a large, prospective observational cohort study assessing deviceand nondevice-based diagnostic measures of concussion.¹³ Our age cutoffs were chosen to limit our sample to the adolescent age period given age-related differences in some of the batteries when used to assess younger children.^{18,19} Participants without concussion who completed testing as part of either pre- or postseason testing for a scholastic sport season (including basketball, field hockey, lacrosse, and soccer) were recruited from a local suburban school.^{20,21} We included all assessments for a given participant in the sample, as an individual without concussion could engage in these assessments across multiple sports, multiple years, or both. We also recruited participants with concussion from the same suburban school and from the sports medicine concussion program of a tertiary care academic medical center. The suburban school and the concussion program included adolescents from the same geographic area with similar sociodemographic characteristics as demonstrated by the demographic similarities of our cohorts with and those without concussion (Table 1). All participants with concussion received their diagnosis in accordance with the 5th International Consensus Statement on Concussion in Sport¹ from a sports medicine physician and sustained their injuries via sport- or recreation-related mechanisms. Participants recruited for the cohort without concussion who subsequently sustained a concussion were only studied as part of the cohort with concussion. The inclusion criterion for participants with concussion was completing the first set of clinical assessments within 28 days of injury. Subsequent evaluations for participants with concussion occurred at the clinical discretion of the treating team per standard clinical care, and we analyzed all follow-up assessments. Exclusion criteria for participants with and those without concussion were active recovery from a previous concussion (\leq 30 days of physician clearance from the previous injury) and any lower extremity trauma that would affect gait assessment, balance assessment, or both. Trained research staff conducted all assessments in either the sports medicine clinical setting or the athletic training facility of the suburban high school. Before enrollment, participants and guardians provided written informed assent and consent as appropriate. The study was approved by the Institutional Review Board of Children's Hospital of Philadelphia.

Clinical Assessments and Batteries

Demographic and Clinical Covariates. We obtained age, sex, race and ethnicity, and concussion history from electronic health records for the cohort with concussion and from self-report for the cohort without concussion. We abstracted the time from injury to assessment for participants with concussion from electronic health records.

The VVE. The VVE consists of 9 examination elements that evaluate vision and vestibular function.^{13,22} Adapted from the VOMS assessment,¹¹ key differences are a larger number of repetitions for saccadic eye movement and vestibular ocular reflex (VOR) testing, which enhances sensitivity,²⁰ the inclusion of abnormal signs in addition to symptom provocation for smooth pursuit testing,¹³ and the addition of monocular accommodation and complex tandem gait.²¹ The examination has been shown to be reliable across multiple clinical settings in which adolescents with

Table 1. Participant Characteristics

| Characteristic | Concussion | |
|--|------------------------------|------------------|
| | Yes (n = 231) | No (n = 166) |
| | No. (%) ^a | |
| Sex | | |
| Female | 123 (53.2) | 95 (57.2) |
| Male | 108 (46.8) | 71 (42.8) |
| Race and ethnicity | | |
| Hispanic | 6 (2.6) | 6 (3.6) |
| Non-Hispanic Black | 20 (8.7) | 18 (10.8) |
| Non-Hispanic White | 183 (79.2) | 128 (77.1) |
| Other ^b | 22 (9.5) | 14 (8.4) |
| Previous concussion | 105 (45.5) | 48 (28.9) |
| | Median (Interquartile Range) | |
| Age, y | 15.6 (14.6-16.7) | 15.6 (14.6-16.8) |
| No. of previous concussions | 0 (0-1.5) | 0 (0-1) |
| Time from injury to first visit, d | 10.0 (5.0-18.0) | NA |
| Time from injury to last visit, d | 33.0 (21.5-69.5) | NA |
| Postconcussion Symptom Inventory at initial assessment | 28.0 (12.5-53.0) | 5.0 (1.0-9.0) |
| Time spent in exercise, h/wk | 3 (2-3) | 2 (2-3) |

Abbreviation: NA, not applicable.

^a Percentages were rounded, so sums may not be 100%.

^b Other includes American Indian or Alaska Native, Asian, mixed, other, unknown, and not reported.

potential concussion receive care.²² Elements of the VVE are (1) smooth pursuit, testing the ability of participants to track in a single (horizontal) plane for 5 repetitions, with an abnormality defined as either symptom provocation (headache, nausea, dizziness, eye fatigue, and eye pain) or abnormal signs (jerky eye movements, jumpy eye movements, and >1 beat of nystagmus)¹³; (2) horizontal and (3) vertical saccades, measuring symptom provocation (headache, nausea, dizziness, eye fatigue, and eye pain), with 20 repetitions of the eyes moving rapidly between fixed objects²⁰; (4) horizontal and (5) vertical VOR, or gaze stability, assessing symptom provocation with 20 repetitions during which the participants' eyes are fixed and their heads move in either the horizontal or vertical plane²⁰; (6) near point of convergence (NPC), evaluating break (double vision) using a standard Astron accommodation rule (Gulden Ophthalmics) with a single-column 20/30 card and abnormal defined as a break occurring at >6 cm²³; (7) right and (8) left monocular accommodation, testing clear-to-blur distance with 1 eye open using an Astron accommodative rule, with abnormal distance defined by age based on the formula of Hofstetter^{24,25}; and (9) complex tandem gait, observing participants for 5 steps forward and backward with their eyes open and then closed and recording sway or steps off a straight line, with abnormal defined as a composite of at least 5 (steps off the line or sway) out of 24 (a scale of 0 to 6 for each of the 4 conditions).²¹

The SCAT5. The SCAT5 is a concussion assessment battery developed by the Concussion in Sport Group that assesses symptom burden, attention, memory, and concentration.⁸ The SCAT5 consists of the following variables: (1) symptom and (2) symptom severity score, assessing 22 concussion symptoms on a 7-point Likert scale (0 = none and 6 = severe) for a possible symptom score of 0 to 22 and a possible severity score of 0 to 132; (3) orientation (naming the date, day of the week, month, year, and time); (4) immediate word memory (3 trials of a list of 5 words);

(5) delayed word memory (repeated list of 5 words after 5 minutes have elapsed); (6) concentration (list of digits backward and months of year backward); and the modified Balance Error Scoring System (mBESS) in (7) double-legged stance, (8) single-legged stance, and (9) tandem stance, with all conditions performed on a firm surface.

The K-D Test. The K-D test evaluates a combination of eye tracking, attention, and language by having the participant read a series of numbers from left to right across 3 test cards with increasingly difficult orientations.⁹ We measured total K-D completion time.

The PCSI. We administered the PCSI adolescent selfreport, which contains 21 concussion symptoms rated on a 7-point Likert scale (0 = none and 6 = severe), to all participants.⁷ The test generates 4 categories of symptoms (physical, fatigue, emotional, and cognitive). We further categorized the physical symptom category into somatic symptoms (headache, nausea, light sensitivity, and noise sensitivity) and vestibular symptoms (visual problems, balance problems, dizziness, and clumsiness).¹³

Modeling and Statistical Analyses

We summarized participant characteristics using standard descriptive statistics. The individual subcomponents of each battery used for modeling (model variables) are listed in Table 2. We first determined the degree of incomplete data (Figure 1). Observations with values at only 1 clinical assessment and observations missing all subcomponents of the VVE and SCAT5 were subsequently excluded from the analysis. We considered whether the incomplete data satisfied a missing-at-random assumption and handled the missing values using multivariate imputation by chained equations.^{26,27} We imputed each variable with missing values via a separate regression model that used all other variables in the data set (the individual battery subcomponents and the demographic and clinical covariates). Continuous and integer-

| Clinical Assessment | Subcomponent | |
|---------------------------------------|---|--|
| Visio-vestibular examination | Smooth pursuit Horizontal saccades Vertical saccades Horizontal gaze stability Vertical gaze stability Near point of convergence Right monocular accommodation Left monocular accommodation Complex tandem gait | |
| Postconcussion Symptom Inventory | Somatic score Vestibular score Fatigue score Emotional score Cognitive score | |
| Sport Concussion Assessment Tool-5 | Total symptom score Symptom severity score Orientation Immediate memory Delayed recall Concentration mBESS double-legged stance mBESS single-legged stance mBESS tandem stance | |
| King-Devick test | Total completion time | |

Abbreviation: mBESS, modified Balance Error Scoring System.

valued variables were imputed using linear regression models; we applied predictive mean matching rather than direct prediction to accommodate skewed and other nonnormal distributions.²⁸ Binary variables were imputed using direct

prediction from logistic regression models. We then created 10 imputed data sets and conducted the factor analysis procedure (described in the next paragraph) separately for each set. We combined the results across the multiple imputed data sets; further discussion of imputation is provided in the Appendix. In addition, we performed a sensitivity analysis to establish the influence of the imputation procedure. All analyses were carried out via R Statistical Software (version 4.2.1; The R Foundation for Statistical Computing).

To account for correlations between the 24 subcomponents, we grouped them into interpretable factors using sparse principal component analysis.^{29,30} The number of factors and level of sparsity were chosen to minimize overlap and maximize interpretability. The resulting 13 factors were linear combinations of the original 24 variables in which the coefficients of the linear combinations were termed loadings. The contributing variables and loadings of each factor are shown in Figure 2. We then adapted a logistic regression model that used the 13 factors plus 4 demographic and clinical covariates (age, sex, presence or absence of a concussion history, and number of previous concussions) to predict concussion status, accounting for the variability in time from presentation to first visit among our participants with concussion. To this end, we fixed a point in time, t, and weighted the log likelihood such that cases with observations closest to t were upweighted relative to other cases. Given that t was not applicable to participants without concussion, all control individuals were weighted equally in the likelihood. The estimated coefficients are then more representative of a comparison of



Figure 1. Overview of the analysis. In step 1 (data cleaning), participants with observed values at only 1 assessment and participants missing all subcomponents of the VVE and Sport Concussion Assessment Tool, fifth edition (SCAT-5), were excluded. In step 2 (imputation), the remaining missing values were imputed using multivariate imputation by chained equations. In step 3 (creation of factors), sparse principal component analysis was used to create interpretable factors from the 24 subcomponents of the 4 clinical assessments. In step 4 (regression), a forward-selection procedure was implemented to determine the factors that maximized discrimination of participants with from those without concussion. Abbreviations: K-D, King-Devick test; PCSI, Postconcussion Symptom Inventory; VVE, visio-vestibular examination.



Figure 2. Factors created by the sparse principal component analysis. Thirteen factors were created from the 24 subcomponents of the clinical assessments listed in Table 2. The loading for each variable is reported in parentheses. Abbreviations: K-D, King-Devick test; mBESS, modified Balance Error Scoring System; NPC, near point of convergence; PCSI, Postconcussion Symptom Inventory; SCAT-5, Sport Concussion Assessment Tool, fifth edition; VOR, vestibulo-ocular reflex; VVE, visio-vestibular examination.

participants with concussion observed close to time t and those without concussion. We repeated this estimation for a range of values of t, specifically for every observed value of time since injury in the data set, yielding a separate set of estimated coefficients for each value of t, which was used to calculate a sequence of predicted probabilities of concussion across the range of t values. Treating the sequence of predicted probabilities as a function, we integrated over the range of values of t and used the integrated values to construct a receiver operating characteristic (ROC) curve and calculate the area under the ROC curve (AUC). The Appendix provides a further discussion of our modeling procedure.

To determine which factors were most highly predictive of concussion status, we implemented a forward-selection procedure whereby the regression models were estimated on a training data set, and AUCs calculated on a validation data set were used as the criteria. To increase the robustness of the feature selection procedure, we performed random subsampling cross-validation 100 times and ranked the 13 factors by the number of times they were included in the model and their average rank when selected into a model. To identify the optimal number of features, we added the factors 1 at a time to the weighted logistic regression model in order of highest selection frequency, and the Akaike information criterion and Bayesian information criterion were calculated at each step. The model with the lowest Akaike information criterion and Bayesian information criterion was considered the optimal model,³¹ and the factors included in that model were classified as top factors. Given that the evaluation of significant factors was based on selection frequencies and ranks, not regression coefficients and standard errors, repeated measures among participants were treated as independent observations.³² Although this can introduce bias into the estimates of regression coefficient variances, it did not hinder our analysis because we were not relying on standard errors for inference. Therefore, we included all observations for each study participant in our analysis. This procedure was repeated on each of the 10 imputed data sets, and selection frequency and rank were averaged across all models. The number of times each factor was selected as a top feature out of the 10 runs was used to detect the overall top factors.

The predictive performance of these top factors was then compared with the individual clinical batteries. For each of the 4 clinical batteries (VVE, SCAT5, K-D, and PCSI), a separate weighted logistic regression model was fit and contained all subcomponents of the assessment, and a fifth model was fit with the top factors as determined by the forward-selection procedure across the 10 imputed data sets. All 5 models included the 4 demographic and clinical covariates of age, sex, presence or absence of a concussion history, and number of previous concussions and were estimated on a training data set. The AUC for each model was calculated on a testing set in the same



Figure 3. Results of the forward-selection procedure. A, Scatterplot of the average number of times a factor was selected into the model out of 100 (*y* axis) versus the average rank in selection order when present in a model (*x* axis). Values were averaged across results from 10 imputed data sets. The top 4 factors are circled in red. B, Bar plot of the number of times each factor was selected as a top feature out of 10 imputations. Factors not listed were never selected as top features. The top 4 factors are outlined in red. Abbreviations: K-D, King-Devick test; mBESS, modified Balance Error Scoring System; NPC, near point of convergence; PCSI, Postconcussion Symptom Inventory; SCAT-5, Sport Concussion Assessment Tool, fifth edition; VOR, vestibulo-ocular reflex; VVE, visio-vestibular examination.

manner as described earlier, and 95% CIs for the AUC were calculated based on 1000 hierarchical bootstrap samples of the data.

RESULTS

The study cohort consisted of 231 participants with concussion, who provided 619 observations, and 166 participants without concussion, who provided 406 observations. The median number of observations per participant was 2 (interguartile range = 2-3, range = 1-10) for the group with concussion and 2 (interquartile range = 1-3, range =1-10) for the group without concussion. A summary of the participant characteristics appears in Table 1. The distribution of characteristics was similar between individuals with and those without concussion. Of 1236 original observations, we excluded 136 due to incomplete data and 75 for not satisfying the inclusion criteria. Of those 211 excluded observations, 126 were from participants whose complete observations at other visits were included. Of the remaining 1025 observations across 32 features (32 800 total potential individual values) used in the analysis, 10.9% (n = 3587 individual values) of assessments were missing and were handled by multivariate imputation via chained equations. Results of the sensitivity analysis on the influence of the imputation procedure are reported in the Appendix.

The results of the forward selection procedure averaged over the 10 imputed data sets are presented in Figure 3. The spread in Figure 3A shows a cluster of 4 factors with a high average rank and a large number of times selected; these are factor 1 (saccades + VOR), factor 2 (mBESS double-legged stance), factor 3 (SCAT5/PCSI symptoms), and factor 4 (K-D completion time). We also identified a second cluster of the remaining factors with a low average rank and a small number of times selected. The separation between these groups is reiterated in Figure 3B, which indicates the number of times out of the 10 imputed data sets that each factor was selected as a top feature. Factor 3 (SCAT5/PCSI symptom score) was selected in 75.29% (n = 752) of the models with an average rank of 1.4 and was selected as a top feature in all 10 imputed data sets. Factor 4 (K-D completion time) was selected in 65.5% (n = 655) of the models with an average rank of 2.4 and was selected as a top feature in all 10 imputed data sets. Factor 2 (mBESS double-legged stance) was selected in 76.9% (n =769) of the models with an average rank of 3.2 and was selected as a top feature in all 10 imputed data sets. Finally, factor 1 (VVE saccades + VOR) was selected in 74.2% (n =742) of the models with an average rank of 3.0 and was selected as a top feature in all 10 imputed data sets.

Among the 9 remaining factors, 6 were never selected as a top feature: factors 8 (VVE complex tandem gait), 9 (VVE NPC + accommodation), 10 (mBESS single-legged



Figure 4. Area under the receiver operating characteristic curve (AUC) of model fit with the top 4 factors from the sparse principal component analysis and logistic regression procedure and models including the subcomponents of each clinical assessment. Each of the 5 models includes the demographic and clinical covariates listed in Table 1. *Factors* refers to the combination of factor 1 (saccades and vestibulo-ocular reflex from the visio-vestibular examination [VVE]), factor 2 (modified Balance Error Scoring System double-legged stance), factor 3 (Sport Concussion Assessment Tool, fifth edition [SCAT-5], and Postconcussion Symptom Inventory [PCSI] symptoms), and factor 4 (King-Devick [K-D] test completion time). Error bars represent 95% Cls for the AUC calculated from 1000 hierarchical bootstrap samples.

stance), 11 (SCAT5 delayed recall), 12 (SCAT5 concentration), and 13 (SCAT5 orientation). The remaining 3 factors were selected in either 1 or 2 of the imputed data sets.

A comparison of the model fit with the top 4 factors (factors 1 to 4) and models fit with all subcomponents from each battery is supplied in Figure 4. The model with the top 4 factors performed as well as any individual battery in predicting concussion status (AUC = 0.816 [95% CI = 0.731, 0.889]) compared with the individual subcomponents of the SCAT5 (AUC = 0.784 [95% CI = 0.692, 0.866]), PCSI (AUC = 0.776 [95% CI = 0.674, 0.863]), VVE (AUC = 0.711 [95% CI = 0.602, 0.814]), and K-D (AUC = 0.708 [95% CI = 0.590, 0.819]). Individual ROC curves for each of the models are provided in Appendix Figure 3.

DISCUSSION

We identified the combination of subcomponents from 4 commonly used clinical concussion assessments that maximized the discrimination of adolescents with concussion (evaluated approximately 10 days from injury in a concussion specialty clinic) from adolescents without concussion. Overall, a combination of 4 factors contributed the most information to identifying concussion status: saccades and VOR from the VVE, symptom scores from the SCAT5/ PCSI, mBESS double-legged stance, and K-D test completion time. Conversely, the factors that contributed the least to identifying concussion status were the VVE complex tandem gait; NPC; monocular accommodation; mBESS single-legged stance; and SCAT5 delayed recall, concentration, and orientation. These results offer a streamlined framework for those using these multiple testing batteries to potentially reduce the number of elements needed to discriminate adolescents with concussion from those without concussion.

Traditionally, symptoms have been the mainstay of concussion diagnosis recommendations.^{5,6} However, given the nonspecific and subjective nature of symptoms, as well as physiological disturbances in cognition, vision, and vestibular function identified after injury, recommendations have been made to augment the symptom evaluation with clinical testing batteries.¹ In 2 large studies of collegiate athletes from the Concussion Assessment, Research and Education Consortium, researchers evaluated augmenting symptombased assessment with testing from other standardized clinical batteries. Among this population of collegiate athletes, Broglio et al¹⁵ used a classification and regression tree analysis and found that sequentially adding mBESS testing to symptom scores provided the highest diagnostic accuracy. These results, augmenting symptom scales with balance testing from the mBESS, partially mirror ours, as 2 of our 4 highest-performing factors were the mBESS double-legged stance (factor 2) and a combination of the SCAT5 and PCSI symptoms (factor 3). Although both the SCAT5 and PCSI symptom scales were included in our model for analytic purposes, because of their redundancy, 1 should suffice for practical purposes in translating these findings into practice. Based on the recommendation to limit the use of the SCAT5 to a more acute timeframe,³³ the PCSI would appear to be the optimal choice for a symptom scale in the subacute timeframe. In our study, the addition of subcomponents from other batteries testing eye tracking, attention, and the visio-vestibular system further contributed unique elements to the concussion diagnosis, as demonstrated by the strong performance of VVE saccadic and VOR testing (factor 1) and K-D total completion time (factor 4). In a separate investigation of the Concussion Assessment, Research and Education Consortium collegiate athlete population, Ferris et al¹⁶ found increased diagnostic sensitivity by adding the complete VOMS to the complete SCAT, third edition, although the authors evaluated the batteries as complete tests rather than the subcomponent analysis we used. Of particular interest is the inclusion of saccadic eve movement and VOR testing from the VVE in the highperforming factors in our analysis. One of the key differences between the VVE and the VOMS is the use of 20 versus 10 repetitions to assess symptom provocation on saccadic eye movement and VOR testing; this difference has been shown to enhance diagnostic sensitivity (without sacrificing specificity).²⁰ The results of our factor analysis further emphasize the value of this evaluation.

The rankings in Figure 3B indicate the relative importance of the factors in the joint model. The number of times a factor was selected as a top feature does not necessarily translate to its individual influence (in isolation from the other factors). A factor, such as the factor 12, which included SCAT5 concentration, that is ranked lower than others in Figure 3B may still individually distinguish participants with from those without concussion effectively but when considered in a joint model with the other assessment measures, it may not contribute much unique information when combined with the other top factors. Given that factors with redundant information are unlikely to be selected together in a single model by the forward-selection procedure, factors that are individually influential in determining concussion status can have low rankings in the joint model. Ultimately, the results in Figure 3 suggest that the top 4 factors, when combined, may provide the most discriminatory information related to a concussion diagnosis, whereas the remaining factors may contain information already captured by the top factors. The redundancy in these measures is particularly salient given the time constraints of treating clinicians. Various degrees of redundancy between these batteries have been demonstrated in other populations; for example, in a population of collegiate athletes, Clugston et al¹⁴ noted a correlation between results on the SCAT, third edition, concentration score and the K-D test. In our assessment, the factor that included the SCAT5 concentration score (factor 12) was used less often than factor 4, which included the K-D composite time (factor 4), also suggesting overlap between the tests, as the K-D tests concentration in addition to attention and eye-tracking ability.⁹

As shown in Figure 4, a diagnostic test including only the subcomponents in the top 4 factors in this analysis performed as well as any of the 4 complete batteries alone (with a higher individual AUC and the tightest CIs, but given overlapping CIs, we cannot rule out similar performance across the other batteries). Because of the possibility of redundant information in the assessment measures. limiting patient assessment to only the features contained in the top 4 factors could reduce both time and cost for clinicians. However, in doing so, clinicians must keep in mind the heterogeneous nature of concussion as well as the utility of these tests for information beyond concussion diagnosis. For example, although deficits elicited by complex tandem gait in this population did not uniquely identify patients with concussion substantially beyond our 4 highestperforming factors, a subset of adolescents with concussion may present primarily with gait disturbances that our highest-performing factors would not capture. Although the mBESS is included in the top 4 factors in our analysis, Corwin et al²¹ found enhanced diagnostic sensitivity of complex tandem gait when directly compared with the mBESS, likely due to the complexity involved in completing a tandem walk backward with one's eyes closed. A phenotype of concussion may exist in which the mBESS and complex tandem gait provide unique information.

Beyond heterogeneous diagnostic considerations, these batteries can be used for both prognosis and tailoring anticipatory guidance. For example, each element of the VVE has displayed a strong correlation with prolonged concussion symptoms, particularly monocular accommodation, which would be eliminated in an assessment paradigm guided by our factor analysis.^{34,35} Eliminating some of these elements might not sacrifice substantial diagnostic power; however, decreasing their use may produce less effective prognostication, a characteristic of the assessment that we did not evaluate. Finally, given the eye-tracking demands in the school setting, multiple elements of visiovestibular testing can also serve as a functional guide for school reentry, enabling clinicians to tailor accommodations to individuals' needs that would be lost with elimination of our 9 lower-performing factors.³⁶ Future researchers should evaluate the ability of our factors to provide prognostic, in addition to diagnostic, information.

Our study had several limitations. Given that the adolescents with concussion were primarily enrolled from a referral sports medicine concussion program, the median number of days from injury to first visit for our participants was 10; therefore, our findings may not be generalizable to the hyperacutely injured adolescent seen in an emergency or urgent care setting or shortly after injury on the sideline, and we caution providers in immediately using these findings in this scenario. Although heterogeneous times existed from injury to initial presentation, our modeling accounted for the time-varying nature of presentation. Next, this population comprised participants with sport- and recreation-related concussions; previous investigators have observed unique recovery trajectories for nonsport-related concussion (eg, due to motor vehicle collisions or assault),³⁷ suggesting perhaps that the features most salient to diagnosis may differ due to the mechanism of concussion. Our participants with concussion largely consisted of patients who were referred, so they may indicate a subset of more severely injured adolescents for whom the diagnostic factors might differ versus generalizing to the overall population of allcomers with adolescent concussion. Future authors should validate our factors across a broader population, including those with different times from injury, as well as various injury mechanisms. In addition, whereas the psychometric properties of the 4 individual batteries have been reported extensively in previous work,7-9,22 we did not assess the psychometric properties of our ideal model in this study, and the interrater and testretest reliability for a streamlined, combined diagnostic battery should be determined. Finally, as demonstrated in Figure 1, although missingness was present in several of our measures, we showed that our imputation measures were robust and avoided substantial bias with a sensitivity analysis that assessed the influence of the imputation procedure, yielding results that were highly consistent with the original analyses (Appendix).

CONCLUSIONS

Our study highlights the importance of a multifaceted concussion assessment for the diagnosis of an injured adolescent, testing various aspects including symptoms, attention, balance, and the visio-vestibular system. Current concussion diagnosis batteries likely measure overlapping domains, and the sparse principal component analysis conducted herein identified strategies for streamlining concussion assessment for clinicians across a variety of settings by reducing the number of elements needed for diagnosis. Although the streamlined battery we presented, which is an initial attempt at such a reduction, may not be ready for immediate clinical implementation, we characterized the redundancy among our current diagnostic tests in the battery. Conducting future studies to further assess both the utility and psychometric properties of a streamlined battery is an important next step.

ACKNOWLEDGMENTS

Funding for this study was provided by the Pennsylvania Department of Health (D.J.C., C.C.M., K.B.A., and C.L.M.). This study was also supported by research grant No. R01NS097549 from the National Institute of Neurological Disorders and Stroke of the National Institutes of Health (I.B., K.B.A., and C.L.M.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank Ronni S. Kessler, Anne Mozel, Taylor Valerio, Olivia Podolak, Ari Fish, Shelly Sharma, and Julia Vanni of the Children's Hospital of Philadelphia Minds Matter Concussion Program for their contribution to data collection for this study. We thank Melissa Pfeiffer for her contributions to data curation. In addition, we thank the students and parents from the Shipley School for their participation. We appreciate the support from the Shipley School administration, faculty, and athletic department, including Steve Piltch, Mark Duncan, Katelyn Taylor, Dakota Carroll, Kimberly Shaud, and Kayleigh Jenkins.

REFERENCES

- McCrory P, Meeuwisse W, Dvořák J, et al. Consensus statement on concussion in sport—the 5th International Conference on Concussion in Sport held in Berlin, October 2016. *Br J Sports Med.* 2017;51(11):838– 847. doi:10.1136/bjsports-2017-097699
- Maruta J, Lumba-Brown A, Ghajar J. Concussion subtype identification with the Rivermead Post-Concussion Symptoms Questionnaire. *Front Neurol.* 2018;9:1034. doi:10.3389/fneur.2018.01034
- Giza CC, Hovda DA. The new neurometabolic cascade of concussion. *Neurosurgery*. 2014;75(suppl 4):S24–S33. doi:10.1227/NEU. 000000000000505
- Barkhoudarian G, Hovda DA, Giza CC. The molecular pathophysiology of concussive brain injury—an update. *Phys Med Rehabil Clin N Am.* 2016;27(2):373–393. doi:10.1016/j.pmr.2016.01.003
- Lumba-Brown A, Yeates KO, Sarmiento K, et al. Centers for Disease Control and Prevention guideline on the diagnosis and management of mild traumatic brain injury among children. *JAMA Pediatr.* 2018;172(11): e182853. doi:10.1001/jamapediatrics.2018.2853
- Halstead ME, Walter KD, Moffatt K; Council on Sports Medicine and Fitness. Sport-related concussion in children and adolescents. *Pediatrics*. 2018;142(6):e20183074. doi:10.1542/peds.2018-3074
- Sady MD, Vaughan CG, Gioia GA. Psychometric characteristics of the Postconcussion Symptom Inventory in children and adolescents. *Arch Clin Neuropsychol.* 2014;29(4):348–363. doi:10.1093/arclin/ acu014
- Shehata N, Wiley JP, Richea S, Benson BW, Duits L, Meeuwisse WH. Sport Concussion Assessment Tool: baseline values for varsity collision sport athletes. *Br J Sports Med.* 2009;43(10):730–734. doi:10. 1136/bjsm.2009.059832
- Galetta KM, Morganroth J, Moehringer N, et al. Adding vision to concussion testing: a prospective study of sideline testing in youth and collegiate athletes. *J Neuroophthalmol.* 2015;35(3):235–241. doi:10. 1097/WNO.0000000000226
- Corwin DJ, Propert KJ, Zorc JJ, Zonfrillo MR, Wiebe DJ. Use of the vestibular and oculomotor examination for concussion in a pediatric emergency department. *Am J Emerg Med.* 2019;37(7):1219–1223. doi:10.1016/j.ajem.2018.09.008
- 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
- 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
- Corwin DJ, McDonald CC, Arobgast KB, Mohammed FN, Grady MF, Master CL. Visio-vestibular deficits in healthy child and adolescent athletes. *Clin J Sport Med.* 2022;32(4):376–384. doi:10.1097/ JSM.00000000000095
- Clugston JR, Houck ZM, Asken BM, et al. Relationship between the King-Devick test and commonly used concussion tests at baseline. J Athl Train. 2019;54(12):1247–1253. doi:10.4085/1062-6050-455-18

- Broglio SP, Harezlak J, Katz B, Zhao S, McAllister T, McCrea M; CARE Consortium Investigators. Acute sport concussion assessment optimization: a prospective assessment from the CARE Consortium. *Sports Med.* 2019;49(12):1977–1987. doi:10.1007/s40279-019-01155-0
- 16. 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
- Zhang AL, Sing DC, Rugg CM, Feeley BT, Senter C. The rise of concussions in the adolescent population. *Orthop J Sports Med.* 2016; 4(8):2325967116662458. doi:10.1177/2325967116662458
- Gubanich PJ, Gupta R, Slattery E, Logan K. Performance times for the King-Devick test in children and adolescents. *Clin J Sport Med.* 2019;29(5):374–378. doi:10.1097/JSM.00000000000670
- Corwin DJ, Zonfrillo MR, Wiebe DJ, Master CL, Grady MF, Arbogast KB. Vestibular and oculomotor findings in neurologicallynormal, non-concussed children. *Brain Inj.* 2018;32(6):794–799. doi:10.1080/02699052.2018.1458150
- Storey EP, Corwin DJ, McDonald CC, et al. Assessment of saccades and gaze stability in the diagnosis of pediatric concussion. *Clin J Sport Med.* 2022;32(2):108–113. doi:10.1097/JSM.000000000000897
- Corwin DJ, McDonald CC, Arbogast KB, et al. Clinical and devicebased metrics of gait and balance in diagnosing youth concussion. *Med Sci Sports Exerc.* 2020;52(3):542–548.
- 22. Corwin DJ, Arbogast KB, Swann C, Haber R, Grady MF, Master CL. Reliability of the visio-vestibular examination for concussion among providers in a pediatric emergency department. *Am J Emerg Med.* 2020;38(9):1847–1853. doi:10.1016/j.ajem.2020.06.020
- Scheiman M, Cotter S, Kulp MT, et al; Convergence Insufficiency Treatment Trial Study Group. Treatment of accommodative dysfunction in children: results from a randomized clinical trial. *Optom Vis Sci.* 2011;88(11):1343–1352. doi:10.1097/OPX.0b013e31822f4d7c
- Nunes AF, Monteiro PML, Ferreira FBP, Nunes AS. Convergence insufficiency and accommodative insufficiency in children. *BMC Ophthalmol.* 2019;19(1):58. doi:10.1186/s12886-019-1061-x
- Hofstetter H. Useful age-amplitude formula. Optom World. 1950; 38:42–45.
- White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med.* 2011;30(4):377– 399. doi:10.1002/sim.4067
- Rubin DB. Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons; 1987.
- Rubin D. Statistical matching using file concatenation with adjusted weights and multiple imputations. *J Bus Econ Stat.* 1986;4(1):87–94. doi:10.1080/07350015.1986.10509497
- Zou H, Hastie T, Tibshirani R. Sparse principal component analysis. J Comput Graph Stat. 2006;15(2):265–286. doi:10.1198/ 106186006X113430
- Zhang Z, Castelló A. Principal components analysis in clinical studies. Ann Transl Med. 2017;5(17):351. doi:10.21037/atm.2017.07.12
- Schwartz G. Estimating the dimension of a model. Ann Stat. 1978;6(2): 461–464. doi:10.1214/aos/1176344136
- Cannon MJ, Warner L, Taddei JA, Kleinbaum DG. What can go wrong when you assume that correlated data are independent: an illustration from the evaluation of a childhood health intervention in Brazil. *Stat Med.* 2001;20(9–10):1461–1467. doi:10.1002/sim. 682
- Echemendia RJ, Meeuwisse W, McCrory P, et al. The Sport Concussion Assessment Tool 5th Edition (SCAT5): background and rationale. Br J Sports Med. 2017;51(11):848–850. doi:10.1136/bjsports-2017-097506

- Master CL, Master SR, Wiebe DJ, et al. Vision and vestibular system dysfunction predicts prolonged concussion recovery in children. *Clin J Sport Med.* 2018;28(2):139–145. doi:10.1097/JSM. 000000000000507
- 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
- Corwin DJ, Grady MF, Master CL, Joffe MD, Zonfrillo MR. Evaluation and management of pediatric concussion in the acute setting. *Pediatr Emerg Care*. 2021;37(7):371–379. doi:10.1097/PEC.00000000002498
- Lumba-Brown A, Tang K, Yeates KO, Zemek R; Pediatric Emergency Research Canada (PERC) 5P Concussion Team. Post-concussion symptom burden in children following motor vehicle collisions. J Am Coll Emerg Physicians Open. 2020;1(5):938–946. doi:10.1002/emp2.12056

Address correspondence to Daniel J. Corwin, MD, MSCE, Children's Hospital of Philadelphia, Roberts Center for Pediatric Research, 2716 South Street, Philadelphia, PA 19146. Address email to corwind@chop.edu. D.J.C. and F.M. are cofirst authors.

Appendix. Data Imputation and Modeling for the Multivariate Model

Data Imputation

Missing values in the predictors were handled using multivariate imputation by chained equations, which assumes that the incomplete data are missing at random. This assumption was evaluated by testing for correlations between the propensity of each variable to be missing and other variables in the data set, especially the demographic and clinical covariates and time since injury. In particular, age and time since injury were highly correlated with missingness. In addition, multivariate imputation by chained equations was well suited for our data because it can handle various types of variables, such as continuous, binary, and integer-valued variables. Finally, the use of multiple imputations rather than a single imputation accounted for uncertainty in the imputed values.

In the imputation procedure, we fit a set of regression models, one for each variable with missing values. Logistic regression models were used for binary variables, and linear regression models were used for continuous and integer-valued variables. In each model, the dependent variable was regressed on all other variables in the data (individual battery subcomponents and demographic and clinical covariates).

Predictive mean matching was used for integer-valued variables and continuous variables with skewed and other nonnormal distributions. Each missing entry was assigned a value that was randomly selected from the observations and had regression-predicted values closest to the regression-predicted value for the missing entry. Given that the imputed values were taken from other observed values of the variable in the data, predictive mean matching ensured that the imputed values were plausible. This approach is more appropriate than direct imputation with the predicted values when the variable with missing entries does not satisfy the normality assumption. The imputation procedure was conducted 10 times to generate 10 unique imputed data sets.

We also conducted a sensitivity analysis to assess whether the results were robust to changes in the imputation procedure. Although time since injury was associated with missingness in the data, it was not included in the imputation models because it was not observed for participants without concussion. To ensure that excluding time since injury from the imputation models did not greatly affect the results, we performed a secondary imputation procedure that included time since injury and repeated the analyses presented in the main paper. Given that controls were not injured and, therefore, did not have an observed value for time since injury, we used an inverted form of time since injury and assigned control participants a value of zero. Except for the inclusion of time since injury, the imputation procedure followed the same steps as described above.

We analyzed the new set of imputed data using the same steps presented in the Methods, and the results were consistent with the primary analyses. The results of the forward-selection procedure are shown in Appendix Figure 1. Appendix Figure 1A shows the same cluster of 4 factors (Sport Concussion Assessment Tool, fifth edition [SCAT5]/Postconcussion Symptom Inventory [PCSI] symptoms, saccades + vestibulo-ocular reflex, modified Balance Error Scoring System double-legged stance, and King-Devick [K-D] test completion time) with a high average rank and a large number of times selected, as shown in Figure 3. The remaining factors form a second cluster with a lower average rank and smaller number of times selected. Appendix Figure 1B reinforces the separation between the 2 groups. The SCAT5/PCSI symptoms factor was selected in 79.3% (n = 793) of the models with an average rank of 1.4 and was selected as a top feature in all 10 imputed data sets. The saccades + vestibuloocular reflex factor was selected in 75.4% (n = 754) of the models with an average rank of 3.1 and was selected as a top feature in all 10 imputed data sets. The modified Balance Error Scoring System double-legged stance factor was selected in 73.1% (n = 731) of the models with an average rank of 3.4 and was selected as a top feature in all 10 imputed data sets. The K-D test completion time factor was selected in 65.1% (n = 651) of the models with an average rank of 2.7 and was selected as a top feature in 9 of the 10 imputed data sets. The remaining factors were selected in less than half of the imputed data sets.

Appendix Figure 2 presents the comparison between a model fit with the top 4 factors and models for each individual battery. All 5 models included the 4 demographic and clinical covariates of age, sex, presence or absence of a concussion history, and number of previous concussions. The model with the top 4 factors performed best in predicting concussion status (area under the receiver operating characteristic curve [AUC] = 0.771 [95% CI = 0.667, 0.860]), followed by the PCSI (AUC = 0.748 [95% CI = 0.647, 0.838]), SCAT5 (AUC = 0.747 [95% CI = 0.639, 0.840]), visio-vestibular examination (AUC = 0.723 [95% CI = 0.616, 0.818]), and K-D test (AUC = 0.701 [95% CI = 0.576, 0.807]). The hierarchical bootstrap CIs showed considerable overlap.

Modeling

To predict concussion status, we set up the following logistic regression model:

$$\operatorname{logit}[\mu_i(t)] = \sum_{j=1}^p \beta_j(t) X_{ij} + \sum_{k=1}^4 \alpha_k(t) Z_{ik,j}$$

where $\mu_i(t) = P(y_i = 1)$, y_i is the indicator that the *i*th (i = 1, ..., n) observation belongs to a participant with concussion (as opposed to a participant without concussion), X_{ij} (j = 1, ..., p) are the factors, Z_{ik} (k = 1, ..., 4) are the demographic and clinical covariates, and $\beta_j(t)$, $\alpha_k(t)$ are the respective coefficients. To account for the variability in time from presentation to first visit among the participants with concussion, we fixed a point in time, *t*, and fit the above model where the log likelihood was weighted such that cases with observations



Appendix Figure 1. Results of the forward-selection procedure on data imputed using time since injury. A, Scatterplot of the average number of times a factor was selected into the model out of 100 (y axis) versus the average rank in selection order when present in a model (x axis). Values were averaged across results from 10 imputed data sets. The top 4 factors are circled in red. B, Bar plot of the number of times each factor was selected as a top feature out of 10 imputations. Factors that are not listed were never selected as top features. The top 4 factors are outlined in red. Abbreviations: K-D, King-Devick test; mBESS, modified Balance Error Scoring System; NPC, near point of convergence; PCSI, Postconcussion Symptom Inventory; SCAT-5, Sport Concussion Assessment Tool, fifth edition; VOR, vestibulo-ocular reflex.



closer to *t* were upweighted relative to other cases. This resulted in a set of coefficients, $\beta(t)$, $\alpha(t)$, which were more representative of a comparison of participants with concussion observed close to time *t* and those without concussion. We repeated this estimation of $\beta(t)$, $\alpha(t)$ for a range of values of *t*, specifically all values of time since injury observed in the data.

The likelihood is specified as follows:

$$\beta[(t), \ \alpha(t)] = \sum_{i=1}^{n} w_i(t) \{ y_i \ \log[\mu_i(t)] + (1 - y_i) \ \log[1 - \mu_i(t)] \},$$

where the weights, $w_i(t)$, are

$$w_i(t) = \begin{cases} c(t)\phi\left(\frac{t_i - t}{\sigma}\right), & y_i = 1\\ 1, & y_i = 0 \end{cases}$$

The constant c(t) was picked for each value of t such that $\sum_{i=1}^{n} I\{y_i = 1\} \ w_i(t) = \sum_{i=1}^{n} I\{y_i = 0\}.$ This ensured that the contribution of participants with concussion to the likelihood relative to participants without concussion remained the same regardless of weighting. The number of days since injury for observation i is t_i . The constant σ determined the range of days around t that was given more weight in the estimation of the regression coefficients and was tuned on a validation data set.

Appendix Figure 2. Area under the receiver operating characteristic curve (AUC) of the model fit with the top 4 factors and each battery on data imputed using time since injury. Each of the 5 models includes the demographic and clinical covariates listed in Table 2. *Factors* refers to the combination of factor 1 (saccades and vestibulo-ocular reflex from the visio-vestibular examination [VVE]), factor 2 (modified Balance Error Scoring System double-legged stance), factor 3 (Sport Concussion Assessment Tool, fifth edition [SCAT-5] and Postconcussion Symptom Inventory [PCSI] symptoms), and factor 4 (King-Devick [K-D] test completion time). Error bars represent 95% Cls for the AUC calculated from 1000 hierarchical bootstrap samples.

We derived AUC values with 95% CIs based on 1000 hierarchical bootstrap samples for models fit with all subcomponents of each clinical assessment and a model fit with the top 4 factors. In Appendix Figure 3, an example of the underlying receiver operating characteristic curves for each of the 5 models is presented for 1 hierarchical bootstrap sample.



Appendix Figure 3. Receiver operating characteristic (ROC) curves for each battery and for a model fit with the top 4 factors (Factors). The ROC curves are calculated from 1 hierarchical bootstrap sample used to calculate the area under the ROC curves and 95% Cls presented in Figure 4. Abbreviations: K-D, King-Devick test; PCSI, Postconcussion Symptom Inventory; SCAT-5, Sport Concussion Assessment Tool, fifth edition; VVE, visio-vestibular examination.