Adding Vision to Concussion Testing: A Prospective Study of Sideline Testing in Youth and Collegiate Athletes

Kristin M. Galetta, MS, Jennifer Morganroth, BA, Nicholas Moehringer, Bridget Mueller, PhD, Lisena Hasanaj, BA, Nikki Webb, MS, ATC, Courtney Civitano, MS, ATC, Dennis A. Cardone, DO, Arlene Silverio, MD, Steven L. Galetta, MD, Laura J. Balcer, MD, MSCE

Background: Sports-related concussion commonly affects the visual pathways. Current sideline protocols test cognition and balance but do not include assessments of visual performance. We investigated how adding a vision-based test of rapid number naming could increase our ability to identify concussed athletes on the sideline at youth and collegiate levels.

Methods: Participants in this prospective study included members of a youth ice hockey and lacrosse league and collegiate athletes from New York University and Long Island University. Athletes underwent preseason baseline assessments using: 1) the King-Devick (K-D) test, a <2-minute visual performance measure of rapid number naming, 2) the Standardized Assessment of Concussion (SAC), a test of cognition, and 3) a timed tandem gait test of balance. The SAC and timed tandem gait are components of the currently used Sport Concussion Assessment Tool, 3rd Edition (SCAT3 and Child-SCAT3). In the event of a concussion during the athletic season, injured athletes were re-tested on the sideline/rink-side. Nonconcussed athletes were also assessed as control participants under the same testing conditions.

Results: Among 243 youth (mean age 11 ± 3 years, range 5–17) and 89 collegiate athletes (age 20 ± 1 years, range 18–23), baseline time scores for the K-D test were lower (better) with increasing participant age (P < 0.001, linear regression models). Among 12 athletes who sustained concussions during their athletic season, K-D scores worsened from baseline by an average of 5.2 seconds; improvement by 6.4 seconds was noted for the nonconcussed controls (n = 14). The vision-based K-D test showed the greatest capacity to distinguish concussed vs control athletes based on changes from preseason baseline to postinjury (receiver operating characteristic [ROC] curve areas from logistic regression models, accounting for age = 0.92 for K-D, 0.87 for timed tandem gait, and 0.68 for SAC; P = 0.0004 for comparison of ROC curve areas).

Conclusions: Adding a vision-based performance measure to cognitive and balance testing enhances the detection capabilities of current sideline concussion assessment. This observation in patients with mild traumatic brain injury reflects the common involvement and widespread distribution of brain pathways dedicated to vision.
the Balance Error Scoring System (BESS) are commonly used at the collegiate level (10, 11). The Sport Concussion Assessment Tool, 3rd Edition (SCAT3) combines both cognitive and balance testing, including the SAC and a modified BESS or timed tandem gait test (11, 12). Child-SCAT3, a battery similar to SCAT3, is recommended for use of athletes younger than 13 years and is under investigation. The combination of SAC and BESS (cognitive and balance testing) has been used to help diagnose concussion on the sidelines in collegiate cohorts (12). However, this composite of tests lacks a vision-based performance measure.

Because approximately 50% of the brain’s circuits are dedicated to vision, the ability to test these pathways may increase our ability to detect concussion after acute injury (2, 13). A study of collegiate athletes suggests that adding a vision-based test of rapid number naming to the SAC and BESS allows for identification of most or all concussions (14). The King–Devick (K-D) test requires intact saccades and other eye movements to perform quickly; literature to date on the K-D test has shown worsening of time scores from a preseason/competition baseline in 60+ concussed athletes (15, 16). Given its simplicity (timed rapid number naming), rapid administration (<1 minute in collegiate athletes), and high test–retest reliability (ICC = 0.96–0.97) (16, 17), the K-D test can be used by nonmedical personnel, including parents—frequently the only adults present on the sidelines for youth sports (18). Having a preseason baseline score for each athlete available in the event present on the sidelines/rink-side as soon as medically feasible. Control athletes playing a similar position who were also consented/assented for the study underwent testing under the same conditions as the concussed athlete. Controls were in place to examine the potential role for fatigue or other factors related to play/practice on test scores in the absence of concussion. Previous studies of the K-D test have demonstrated that vigorous exercise/scrimgame alone is associated with improvements in scores, consistent with learning effects inherent in performance measures in the absence of injury (15, 17).

**King–Devick Test**

K-D is a test of rapid number naming that takes <1 minute to administer in collegiate athletes; times are slightly longer (<2 minutes) in younger athletes. The test consists of 3 cards with variably spaced single digit numbers (Fig. 1). Participants are asked to read each card as quickly as possible; the time to read each card is recorded. Times for all 3 cards are summed to give a total time score (15, 16,19–22). Higher testing times compared with baseline indicate worsening of performance. Worsening scores are not observed after competition/fatigue alone (15); healthy athletes typically show improvement with exercise (15, 17). Because shorter times are expected in non-concussed athletes, any worsening of K-D scores from baseline is consistent with concussion. In previous studies investigating the utility of the K-D test, baseline examinations were performed in a noisy hot locker room, and we found no difference in inter- or intra-rater reliability (15). Although these examinations were not performed during game time, they were performed under uncomfortable circumstances. In this study, baseline examinations of the K-D test were also administered for 2 trials to obtain the best possible baseline score.

**Standardized Assessment of Concussion**

The SAC is a brief cognitive test. A maximum total score of 30 is generated by adding the 4 subscores: Orientation (maximum score = 5), Immediate Memory (maximum score = 15), Concentration (maximum score = 5), and Delayed Recall (maximum = 5). Recent evidence-based guidelines suggested that a worsening of 2–4 points from baseline is a sensitive threshold for clinically meaningful change for SAC (12). We used this threshold as a dichotomous criterion for worsening of the SAC.

**Timed Tandem Gait Test**

The timed tandem gait is a balance component of the SCAT3 and Child-SCAT3. To perform this test, the participant is instructed to walk along a 38-mm wide,
3-m long sports tape. Athletes place 1 foot in front of the other along the line as quickly and accurately as possible. The best time of 4 trials back and forth along the tape is recorded as the official score. Studies show that dynamic balance and coordination tests, such as timed tandem gait, are less impacted by exercise fatigue than are static balance tasks, such as the BESS (23). Resiliency of dynamic balance testing was reinforced by a study showing that high-intensity exercise decreased performance of both static and dynamic balance testing; timed tandem gait, however, was unaffected by moderate exercise (20).

**Testing Procedures**

Baseline preseason assessments for the sideline tests (K-D, SAC, and timed tandem gait) were performed preseason before practices started. Tests for these sessions were administered by trained study volunteers or by athletic trainers in the case of collegiate athletes. NYU study personnel were present and performed training of testers for baseline assessments. Components of SCAT3 are performed routinely as part of athletic training for collegiate athletes; in the youth athletes, these tests were performed for research purposes as part of this study.

Concussion was defined using the standard definition of witnessed or reported impulse blow to the head or body followed by any neurological symptom(s). Judgments about whether a concussion had occurred were made by athletic trainers for collegiate cohorts. For youth athletes, the judgments of volunteer parents specifically assigned to assess injured athletes at each game/practice were used.
For athletes without witnessed trauma who presented later with symptoms consistent with concussion, testing was performed as soon as possible after the athlete self-reported. The diagnosis of concussion was confirmed in all cases by an expert physician.

Statistical Analyses
Statistical analyses were performed using Stata 13.0 (StataCorp, College Station, TX). The Wilcoxon signed-rank test was performed to determine changes in scores from baseline to postinjury. Linear regression models, accounting for age, were used to examine regression models, accounting for age, with calculation of areas under the receiver operating characteristic (ROC) curves. ROC curve areas represent the probability that a test or combination of tests can distinguish concussed from controls and range from 0.5 (probability no better than chance) to 1.0 (perfect ability to distinguish). Comparisons of the logistic regression-derived ROC curve areas were made for combinations of the tests (K-D, SAC, total score, and balance testing using linear combination methods.

RESULTS
Preseason baseline and postinjury test scores are reported in Table 1. There were 243 youth athletes aged 11 ± 3 (range 5–17 years, 16% female) and 89 collegiate, aged 20 ± 1 (range 18–23 years, 26% female), for a total of 332 participants. Baseline scores for all tests (Table 1) improved with increasing age in this combined cohort (P < 0.001, linear regression models); this age effect was particularly evident for K-D card 3, where vertical crowding of the test numbers is the greatest (P = 0.001 for card 3 vs card 1, linear regression; Fig. 2). Twelve athletes sustained a concussion during their athletic season.

A composite of tests including SAC, timed tandem gait, and...
K-D together (ROC curve area = 0.97) was a greater discriminator of concussed vs control athlete groups than was the combination of SAC and timed tandem gait as used in the SCAT3/Child-SCAT3 (ROC curve area, 0.88, Fig. 3B). The combination of timed tandem gait and K-D (ROC curve area = 0.98) was nearly identical to the composite of 3 tests (ROC area = 0.97). This means that athletic trainers had a 92% probability of correctly distinguishing a concussed vs nonconcussed athlete based on the result of the K-D test alone.

When the test scores were analyzed using published cutoffs for worsening from baseline in the setting of concussion, the SAC showed a 2 point or more worsening (12) in 2/10 concussed players (20%) and 3/14 controls (21%). The timed tandem gait showed worsening in 10/12 concussed players (83%) and 5/14 controls (36%). K-D times demonstrating worsening in 9/12 (75%) concussed players and 1/14 controls (7%).

**DISCUSSION**

Results of this investigation demonstrate that adding a rapid simple vision-based performance measure to cognitive and balance tests enhances the detection capabilities of current sideline assessments for concussion. Because rapid number naming captures visual function, K-D is a useful tool to aid in the diagnosis of concussed athletes at all levels of sport (14–16,19–22). Use of a measure that requires saccadic eye movements is particularly effective for several reasons. Studies have shown that patients with impaired saccades post-concussion have both cortical and subcortical deficits. These deficits correlate with worse scores for quality of life assessments (24). Saccadic eye movements require relay of information throughout the brain, including frontal eye fields,
supplementary eye fields, dorsolateral prefrontal cortex, intraparietal sulcus, and deeper structures of the brainstem (25–27). Eye movement testing enables the analysis of a number of circuits throughout the brain including visual–spatial integration, motor planning, attention, motivation, and spatial organization (26). The wide distribution of neuronal networks required for saccades thus makes a vision-based sideline screening test particularly effective.

The K-D test has been successful in identifying concussion in boxers and Mixed Martial Arts fighters. In those studies, worse K-D scores were associated with lower scores for the Military Acute Concussion Evaluation, a brief cognitive test, both postfight (\( r_t = -0.79, P = 0.0001 \)) and regarding changes from prefight baseline (\( r_t = 0.90, P < 0.0001 \)) (19). Worsening of K-D times was associated with worsening SAC immediate memory scores (\( P < 0.001, R^2 = 0.62 \)) (16). Studies have shown that K-D scores correlate with Immediate Post-Concussion Assessment and Cognitive Testing (IMPACT) subscores that are visual in nature (14, 17, 28). As an added benefit, the K-D test can accurately and easily be performed by non-medically trained observers, including parents of youth athletes (18).

Another factor that adds to the simplicity and relevance of the K-D test in youth athletes is the use of preseason baseline scores. Baseline scores obviate the need for parents or others on the sidelines to determine normative values in the acute setting of an injury. Furthermore, as shown in this study, K-D time scores decrease (improve) with advancing age of youth athletes. Although these factors make determination of new baseline scores essential at the start of each athletic season, the use of baseline scores eases interpretation when time is of the essence. Using modern definitions, a concussion should be suspected when an athlete has 1) an impulse blow to the head or body and 2) any new neurological symptom. Tests such as K-D, therefore, are used to remove some of the guesswork from this process and should not substitute for clinical or parental judgment that a concussion has occurred.

In our youth athlete cohort, worse scores were noted among younger players for all sideline tests. K-D time scores in particular were significantly slower for younger players (\( P < 0.001 \)). This association with age and improved overall K-D scores (faster times) could be explained by developmental changes in saccadic eye movements and cognition. Diffusion tensor imaging MRI studies have shown that both white matter and gray matter changes continue in the frontal lobes throughout childhood (29). Eye movement tasks, which require frontal lobe circuits, begin to reach stabilization around adolescence, in concert with other developmental changes in the brain (29).

Saccades have been described by their components: peak velocity, latency, and accuracy. Although changes in velocity of saccadic eye movements with age have been inconsistently described, saccadic latency decreases throughout childhood. Changes in accuracy also stabilize with age (29). Age-related changes in saccade latency and accuracy may extend beyond the oculomotor system and may reflect changes in cognitive processing (29). The K-D test, which requires saccadic eye movements with a superimposed cognitive task, may also be impacted by the normal developmental changes of the brain with age and that may explain the improved times we observed with increasing age.

Performance on test card 3 with the greatest degree of vertical visual crowding had the most variability in terms of testing times. Scores on this card improved with older age within the cohort of athletes younger than 18 years (Fig. 2, \( P < 0.001, \) linear regression). It is suspected that the effect on this specific card may also be secondary to visual crowding, an age-dependent ability to visualize objects among clutter (30).

Future studies of the K-D test will explore the possibility of age-related norms. In this study, baseline scores for collegiate athletes averaged 37.4 seconds, very similar to collegiate athlete scores in previous studies (average 37.0 seconds, range 36.0–40.2 seconds) (15). Reference ranges for baseline scores in children are under development; these are likely to correlate with age as suggested by our results and by literature suggesting an impact of the changing brain on measures of saccadic eye movement performance (30). Studies are also ongoing to examine the eye movement dynamics and correlates of K-D test performance using formal eye movement recordings. These investigations will determine how prolonged K-D test times may relate to transient slowing of saccades, saccadic inaccuracy, increased latency, or a combination of these factors. The potential role for antisaccades in eye movement-related tasks after concussion will also be examined.

To our knowledge, this is the first investigation to examine the use of timed tandem gait in children. Our data show that the timed tandem gait is a potentially useful tool in the assessment of concussions in youth athletes. In terms of sensitivity in our cohort, the timed tandem gait fell only slightly behind the K-D test as a diagnostic tool. Further investigation will determine test–retest reliability. In this cohort, we did not find SAC testing to be helpful in distinguishing the concussed athlete. A previous study of the SAC in a pediatric cohort presenting to the emergency department for concussion also did not find a significant difference in scores vs nonconcussed controls (31). Baseline SAC scores in youth populations can be very low, as observed in our study, making it difficult to find a decrement in SAC scores after concussion in certain athletes. Similar to the timed tandem gait, SAC testing requires further validation in youth population.

REFERENCES