Is King-Devick Testing, Compared With Other Sideline Screening Tests, Superior for the Assessment of Sports-related Concussion?

A Critically Appraised Topic

Karissa N. Arca, MD,* Amaal J. Starling, MD,‡ Marie D. Acierno, MD,§ Bart M. Demaerschalk, MD, MSc, FRCP(C),* Lisa Marks, MLS, AHIP,∥ and Cumara B. O’Carroll, MD, MPH*$

Background: Concussion affects almost 4 million individuals annually. There are many sideline screening tools available to assist in the detection of sports-related concussion. The King-Devick (K-D) test in association with Mayo Clinic utilizes rapid number naming to test saccadic eye movements in order to screen for concussion. An ideal screening tool for concussion would correctly identify all athletes with active concussion. The accuracy of K-D testing compared with other sideline screening tools is undetermined.

Objective: To critically assess current evidence regarding the utility of K-D testing as a sideline screening tool for acute concussion and compare K-D testing to other sideline concussion assessments.

Methods: The objective was addressed through the development of a critically appraised topic that included a clinical scenario, structured question, literature search strategy, critical appraisal, assessment of results, evidence summary, commentary, and bottom-line conclusions. Participants included consultant and resident neurologists, a medical librarian, clinical epidemiologists, and content experts in the field of concussion neurology and neuro-ophthalmology.

Results: A recent meta-analysis was selected for critical appraisal. Cohorts analyzing athletes with sports-related concussion were selected and utilized K-D testing as the main baseline and sideline assessment of concussion. K-D testing was found to have a high sensitivity and specificity for detecting concussion when there was worsening from baseline.

Conclusion: K-D testing has high sensitivity and specificity for detecting sideline concussion. Compared with other sideline screening tools that do not include vision testing, it has greater accuracy. Screening for concussion is optimized when multiple testing modalities are used in conjunction.

Key Words: concussion, mild traumatic brain injury, sideline, King-Devick test, evidence-based medicine, critically appraised topic

The Neurologist 2020;25:33–37

From the Departments of *Neurology; ‡Ophthalmology; and ¶Library Services, Mayo Clinic, Scottsdale, AZ.

The authors declare no conflict of interest.

Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved.

www.theneurologist.org | 33

CRITICALLY APPRAISED TOPICS

CLINICAL CASE

A 13-year-old competitive ice hockey player with 5 prior head injuries experienced a sixth head injury during a recent ice hockey game. He fell forward and was elbows in the right temporal aspect of his helmet, causing him to hit the boards below the Plexiglass and fall to the ice, striking his head against the ice. While his helmet was intact and there was no reported amnesia, loss of consciousness, or seizure activity at the time of the event, he later developed headache, fatigue, drowsiness, dizziness, vision difficulty, and orthostatic lightheadedness. He was evaluated in Concussion Clinic 8 days after the injury at which point he had a Post-Concussion Symptom Scale (PCSS), Standardized Assessment of Concussion (SAC), and King-Devick (K-D) testing performed. His total number of symptoms on the PCSS was 13 of 22 and severity was 35 of 132. He scored 27 of 30 on the SAC. His postinjury K-D test time was 44.0 seconds with no errors; no baseline was performed. A discussion arose as to whether or not sideline K-D testing at the time of initial injury would have been appropriate in the evaluation of this patient’s concussion.

BACKGROUND

Concussion

Concussion affects ~4 million individuals annually, but the incidence, including that of sports-related concussion (SRC), is often difficult to estimate given discrepancies in reporting and people that do not present for medical care.1–7 Concussion, or mild traumatic brain injury, is a disturbance in brain function that may be accompanied by loss of consciousness, focal neurological dysfunction, altered mentation at the time of the injury, or amnesia up to 24 hours. In order to be classified as “mild,” the loss of consciousness cannot exceed 30 minutes and the patient’s Glasgow Coma Scale score must be ≥13 after 30 minutes.8 Patients can experience physical, mental, emotional, and sleep symptoms following concussion and these symptoms can worsen with physical or mental exertion.9,10 At a cellular level, the physical brain injury experienced during concussion induces cation and anion exchange that ultimately results in mitochondrial dysfunction and the inability to produce cellular energy. After this occurs, tasks requiring exertion such as physical or mental exercise can cause worsening of symptoms.9,11,12

Sideline Assessment of Concussion

Many sideline screening tools for concussion are in use and are often administered by nonmedical personnel. These tests should only be used to screen for, rather than diagnose concussion.7,13 The gold standard for diagnosis of concussion is clinical evaluation by a trained medical professional.13 Sideline screening tools utilize different forms of testing ranging from...
balance to cognitive assessments to saccadic eye movements, and there are tools that integrate many of these components. In order to prevent a concussed athlete from returning to play prematurely and sustaining further cerebral damage, the ideal screening tool would capture all athletes with concussion at the time of injury. Some of the commonly used sideline screening tools in use include K-D test in association with Mayo Clinic, Sport Concussion Assessment Tool (SCAT; the most recent version of which is SCATS), SAC, and Balance Error Scoring System (BESS).

SCATS is used in athletes ages 13 and above (with Child-SCAT also available for ages 5 to 12) and includes 4 steps on the sideline followed by 5 “off-field” steps. After assessing for signs of emergent neurological deterioration on the field, the off-field steps include evaluation of orientation, immediate memory, concentration, delayed recall, a neurological screen that includes modified BESS, coordination, tandem gait, and a postconcussion symptom scale. SAC does not require any medical background to administer and assess orientation, immediate recall, concentration, and delayed recall. BESS assesses the athlete’s balance on multiple surfaces including the ground and a foam pad. Any worsening from baseline is a sign of a potential concussion. Military Acute Concussion Evaluation (MACE) comprises a variation of SAC as well as a brief neurological examination including pupillary response to light, extraocular movements, speech fluency, grip strength, pronator drift, and Romberg.

K-D Test in Association With Mayo Clinic
In the K-D test in association with Mayo Clinic, an athlete completes rapid number naming which measures the function of saccades, ocular convergence, accommodation, attention, and language. Together, the ability to perform these functions evaluates the integrity of the cortex, subcortex, basal ganglia, brainstem, and cerebellum. When an athlete’s time to complete the test or number of errors worsens from baseline by one second or more, this has been shown to correlate with concussion. Many small prospective cohort studies have evaluated the utility of the K-D test as a sideline screening assessment for concussion, with some also comparing it to other sideline evaluations, looking for evidence of superiority of one over the other. One of the first studies using the K-D test in sideline concussion screening was performed by Galetta and colleagues in 2011. Ten of 219 collegiate athletes in various sports experienced concussion over the course of one season. The K-D test was performed at baseline and immediately post-concussion. In the concussed group, K-D times were longer/worse by almost 10 seconds (P = 0.009). Another study by Galetta and colleagues in 2015 queried if K-D could complement the current measures used for sideline diagnosis of concussion and compared K-D testing with SAC and timed tandem gait. In 243 youth ice hockey and lacrosse athletes and 89 collegiate athletes, 12 sustained concussion and had worsening of K-D times by an average of 5.2 seconds. When compared with SAC and timed tandem gait, K-D had a greater capacity to distinguish a concussed athlete from a control (n = 14) who was tested under similar conditions.

Clinical Question
Is K-D testing, compared with other sideline screening tests, superior for the assessment of sports-related concussion?

SEARCH STRATEGY
The search was conducted using the Ovid MEDLINE database. The following MeSH terms were used and all were exploded: athletic injuries; brain concussion; brain injuries, traumatic and cranioencephalic trauma. The text/keyword words used included: athletic injury; concussion; brain trauma; SCAT or sport concussion assessment tool; King-Devick; SRC or sport related concussion; sport related head trauma; PCSS or post-concussion symptom scale and ImPACT or immediate post-concussion assessment cognitive testing. MeSH terms and text words were combined using the Boolean operators of “OR” and “AND” resulting in 39 citations. Of the 39 references, a total of 2 systematic reviews and 6 review articles were identified; no meta-analyses. While reviewing these studies, including some of the prospective cohort studies that directly identified the clinical question at hand, the references were hand-searched, and the meta-analysis Galetta and colleagues published in the journal Concussion, 2016 was identified. This meta-analysis was selected because it directly answered the clinical question and was the highest level of evidence available.

EVIDENCE, RESULTS, AND CRITICAL APPRAISAL
This 2016 meta-analysis by Galetta et al is the first to evaluate the efficacy of K-D testing and compare it to other sideline concussion screening assessments. The objective of this meta-analysis was to determine how accurately K-D testing distinguishes concussed athletes from controls and how it compares with other sideline screening tools. This study included 1419 athletes from 15 cohort studies with a mean age of 18.3 years across various levels of sports, from youth to professional. The pooled analysis average preseason K-D score was 44.5 seconds, whereas the weighted average preseason K-D score across all athletes was 43.8 seconds. No methodology was given for weighting technique. Of the 1419 athletes, 112 experienced concussion per standard definition of a witnessed or reported injury to the head or body with subsequent new neurological signs or symptoms. In all, 96 of these 112 players had postinjury K-D scores longer than their preseason values by an average of 4.8 seconds [95% confidence interval (CI): 3.7–5.8, P < 0.001]. In 202 nonconcussed control subjects, 181 did not have worsening of their K-D scores from their preseason baseline and on average were 1.9 seconds shorter (95% CI: −3.6 to −0.02, P = 0.05). On the basis of these results, the sensitivity of K-D testing for detecting sideline concussion was 86% and the specificity was 90% (Fig. 1). The relative risk of concussion if an athlete had a worsening score on K-D testing was 4.92 (95% CI: 3.07–7.89, P < 0.001). Calculations of sensitivity, specificity, and relative risk came from 8 of the 15
cohorts in the meta-analysis. Upon further discussion with one of the primary authors this was because these cohorts were the only cohort studies that included both concussed and control athletes, with differences between the 2 groups reported. Other cohorts may have had one or the other, but not simultaneously.

A total of 69 athletes, presumed part of the original sub-cohort of 112 concussed, had additional baseline and postinjury sideline assessments. The authors focused on SAC and timed tandem gait and compared these with K-D. Logistic regression models were used to develop receiver operating characteristic (ROC) curves for this data. While the raw data for the SAC and timed tandem gait scores were not available for this subcohort, it was found that worsening from baseline occurred in 85% of concussed athletes tested with K-D, 75% tested with timed tandem gait, and 48% tested with SAC. Their ROC curves were as follows: K-D 0.89 (95% CI: 0.82-0.96), timed tandem gait 0.81 (95% CI: 0.82-0.96), and SAC 0.66 (95% CI: 0.53-0.79), and the difference across ROC curves was statistically significant (P = 0.002). One of the most important findings from this data is that the worsening of at least one of the 3 screening tests was appreciated in all of the concussed athletes (100%). Furthermore, the authors also highlighted an age-effect on K-D test scores whereby the scores decreased with age. Specifically, athletes ages 5 to 18 showed the greatest improvements in K-D test time with increasing age (P < 0.001). In many cohorts studying K-D testing and concussion, athletes undergo a series of 2 baseline K-D tests. In this meta-analysis the authors concluded that there was no learning effect between the first and second trials despite improvement of times. A total of 1048 athletes had 2 baseline K-D trials with the first trial mean time of 49.5 seconds (95% CI: 45.8-47.6) and the second trial mean time of 46.7 seconds (95% CI: 45.8-47.6). While there was a statistically significant difference between these times (P < 0.0001), the intraclass correlation coefficient was 0.92 which suggests that the difference was due to the individual athlete rather than a learned effect from the repetition of the test by the same participant. In other studies, postseason K-D scores have been minimally lower than the best preseason scores. In contrast to the above conclusion from this cohort, the change from preseason to postseason is thought to likely reflect learning effects. It is important to note that an athlete’s K-D testing times are not altered by fatigue from recent play. To the contrary, K-D testing in persons without concussion but with fatigue demonstrates improvement in time. For all these reasons, a worsening K-D score after a witnessed or suspected head injury is even more reliable for underlying central nervous system dysfunction.15,19

A strength of our chosen manuscript is that Galetta et al20 is the only meta-analysis to compare K-D testing to other sideline assessment tools for concussion. In addition, there is low heterogeneity among the individual cohorts included in the final analysis. As noted throughout the above results, the I² values were all 0.0%. Statistically speaking, low heterogeneity is a strength, but this may not accurately represent some of the inherent differences among the athletes in the various studies. For instance, Duenas and colleagues studied athletes in Guam and suffered from participation bias due to the concerns from athletes and their parents that the trial data would be used against their future athletic careers and educational goals.21

There were also differences in sports, ages of athletes, and the types of sideline screening tools used aside from K-D. Women are poorly represented in this meta-analysis, but this may be in part due to the sports in which K-D testing was employed. Almost none of the cohort studies gave background information on the players which may also tend to the heterogeneity being 0.0%,15,19,21–33

Limitations of this study include the small sample size. Although 1419 athletes were enrolled in the study, only 112 who experienced concussion were studied, contributing to low sample size. Unfortunately, this is a problem inherent to concussion research since these are all cohort studies and are unable to predict the number of athletes that will sustain concussion over a period of time. In addition, a quality assessment by which to compare the pertinent features of each study included in the analysis was not available and no restrictions were used to determine if a cohort did not meet a particular quality standard. For future meta-analyses, we would propose quality measures such as a predetermined sample size, inclusion of standardized definition of concussion, inclusion of control subjects tested by a particular means, and inclusion of other sideline screening tools of the author’s choice. This method of quality analysis would help to account for and potentially eliminate some of the outliers that were noted, such as one athlete who was 63 years old.20

There were 7 studies left out of the relative risk analysis (figure 5 of Galetta et al20) without an explanation by the authors. Although it appears that the reason for excluding studies may have been due to lack of controls for which to derive a relative risk, this reasoning is difficult to follow when looking at the data from each of the individual studies. For instance, Marinides and colleagues is a retrospective study and by design does not include control athletes and yet it was included in the relative risk analysis.

**Clinical Bottom Lines**

1. K-D testing has high sensitivity (86%; 95% CI: 79%-92%) and specificity (90%; 95% CI: 85%-93%) for detecting active concussion on the sideline. Compared with other sideline tests that do not include vision testing, it is more accurate in the detection of concussion.

2. Worsening of K-D score by one second or longer from baseline suggests a 5-fold greater likelihood of concussion.

3. Identifying potentially concussed athletes on the sideline is optimized when multiple testing modalities are used in conjunction.

4. Yearly or seasonal K-D testing, especially among young athletes, is recommended due to improvements that occur with age.

5. K-D testing has high test-retest reliability and can be administered by laypersons.

**DISCUSSION**

The goal for any sideline concussion screening tool is to help keep athletes safe and minimize the potential for further brain injury until they are fully evaluated by a medical provider. There is a need for a comprehensive and dynamic test for sideline detection of concussion that utilizes a multimodal approach in order to detect all sideline concussions. The fact that 100% of sideline concussions of 69 players in this study were detected by at least one of the 3 selected tests (K-D, timed tandem gait, and SAC) speaks to the utility of multimodal testing. Many concussion screening evaluations are already being performed annually, but it is of particular importance to screen yearly for K-D testing due to the nature of the developing athlete’s brain, more specifically the frontal lobe which contributes to the frontal eye fields and saccadic eye movements.20 Furthermore, programs such as tele-concussion also provide a real-time assessment by trained physicians.
to determine the eligibility of an athlete to return to play based on these sideline assessments and the athlete’s overall symptomology.14

**Neuro-Ophthalmology and Neurology Expert Commentary**

Testing for concussion needs to be simple, timely, and reliable as there is a need in the acute setting of the brain injury to efficiently and effectively “look” for the earliest symptoms of a concussion. The K-D test is a good vision-based performance measure. It is a test in which the eyes help us to “see” how information from a rapid screening can quickly assess acute symptoms. In the acute setting, the K-D test adds data to other concussion tests that measure cognition and balance (i.e., SAC and BESS), increasing the accuracy of diagnosing concussion.19 The K-D test can be likened to a quick reading task capturing eye movements (saccades, convergence, and accommodation) as well as attention, and language. In our opinion, the KD test is reliable because it is purely objective. It is a rapid, straight-forward test that captures saccadic eye movements. Hence, the person administering the test need not assess the speed, rate, or accuracy of the saccades. The vast majority of concussions have subtle abnormalities in eye movements that can be captured by the K-D test. More than half of the brain pathways are dedicated to eye movements. Thus, the impaired eye movements demonstrated by K-D testing imply suboptimal brain function.15

The best combination for any medical test is one that can be both administered with ease and taken without much frustration. The K-D test is a portable test that can be quickly and simply administered on the sidelines via paper or tablet. The time to complete the test is usually less than 1 minute and no more than 2 minutes in younger athletes. The rapid naming of numbers makes the test fair easily for persons of all ages. The test has straightforward directions and therefore, lends itself to be administered by nonmedical personnel.15,19 The K-D test, a rapid screening test in the acute setting, adds the critical dimensions of speed, efficiency and accuracy to expedite concussion recognition.

The most worrisome consequence of a concussion is the increased immediate vulnerability of a subsequent concussion and the higher incidence of prolonged recovery in the setting of multiple concussions without adequate recovery in the interim.35,36 This highlights the importance of timely removal from play due to a suspected concussion. Rapid, reliable, accessible, easy to administer objective screening tools with high sensitivity can aid in the removal from play. Ideally, these tools would effectively screen for symptoms and signs seen in a concussion, but they would not diagnose a concussion. Concussion remains a clinical diagnosis. However, these tools are analogous to a thermometer. A thermometer does not provide a diagnosis, but is an accessible, easy to use tool that alerts the layperson that something may be wrong and prompts a healthcare office visit for additional work-up and diagnosis. We need a “thermometer” for concussion. This meta-analysis introduces a potential “thermometer” for concussion, the K-D test. In fact, the efficacy of the K-D administered by laypeople or “sports parents” has been preliminarily evaluated and appears to be feasible.30 However, this meta-analysis importantly highlights that a multimodal assessment can raise the sensitivity of testing to 100% which is the goal in removal from play in the setting of a suspected concussion. In addition, the critical appraisal of this meta-analysis raises several salient points. First, we must have a standard, well-accepted diagnostic criterion for concussion that is used in research. Otherwise, we cannot reliably perform quality meta-analyses and establish high levels of evidence. Second, we need to establish parameters for clinical studies in concussion including, but not limited to having a control/nonconcussed cohort for comparison. Concussion is a public health epidemic, especially in youth athletes, simply due to the sheer number of youth participating in sports. We need high levels of evidence to be able to apply accessible, well-studied tools on a population level to address this crisis.

**REFERENCES**


