ABSTRACT

Background: A battery of tests is commonly used to measure disability with and recovery from concussion. A number of different concussion-oriented assessment tests exist and each is considered useful. To the authors' knowledge, no study has compared the scores of these tests during recovery in the middle school and high school aged population to see how each change over time.

Purpose: The purposes of this study were to analyze clinical data of concussed middle school and high school aged athletes to determine the concurrent and predictive validity for post-concussion syndrome (PCS) of the Post-Concussion Symptom Scale (PCSS), Balance Error Scoring System (BESS), and the five subscales of the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT).

Methods: The study was a retrospective chart review performed on middle school and high school aged individuals with a diagnosis of concussion from the years 2008-2010 within the Akron Children's Hospital Sports Medicine system. To be eligible for inclusion in the dataset, each subject required a baseline measurement for each of the three tests (and all five subscales of the ImPACT) and a post-test measure. The mean age of the population was 15.38 years (SD = 1.7) and ranged from 11 to 19 years. Pearson product correlation tests (correlation matrix) were used to analyze the concurrent validity of the test items during recovery following a concussion. Receiver operating characteristics (ROC) curves were used to determine the predictive validity of initial scores for developing PCS.

Results: The correlation matrix captured five statistically significant findings; however, these suggested only weak to mild correlations. Five test items yielded an area under the curve (AUC) greater than 0.50 but only one was statistically significant. After qualitative evaluation, only one of the three tests (including the five subscales of the ImPACT) was useful in predicting post-concussion syndrome.

Conclusion: This study suggests that there is poor concurrent validity among three commonly used concussion tests and there is no baseline score that predicts whether post-concussion syndrome will occur.

Level of Evidence: 2b

Key Words: Concussion, Diagnostic accuracy, Post-concussion syndrome, Validity
INTRODUCTION
There are an estimated 1.7 million traumatic brain injuries (TBIs) each year in the United States, of these 1.36 million are treated and released from an emergency room setting. Approximately 75% of TBI cases are classified as a mild traumatic brain injury or concussion. TBIs cost the country approximately $17 billion dollars annually. Of all the concussions that occur each year, 300,000 are a result of playing a sport. An estimated 55.2% of all high school aged students in the country participate in some form of an organized sport, thus the potential for concussions for this population during sporting events yearly is exceptionally high.

A concussion has been defined as a complex pathophysiologic process that occurs from a blow to the head, face, or neck, in which the force is transferred to the head. The concussion affects the brain and can lead to a rapid onset of short-lived impairments that are not always seen via imaging. Recognition of symptoms associated with concussion is imperative for the health care provider. Post-concussion syndrome (PCS) is a poorly understood and typically non-life threatening condition, which occurs when concussive symptoms are prolonged for weeks or potentially months after the injury that caused the concussion. Post-concussion syndrome is generally treated by providing information, education, and reassurance but may also include pharmacologic therapies designed to reduce prolonged symptoms such as sleep disturbances or anxiety. Second impact syndrome is a disorder where the brain swells rapidly after a person suffers a second concussion before symptoms from an earlier concussion have subsided, is life threatening and differs significantly from post-concussion syndrome. The occurrence of second impact syndrome has been documented almost exclusively in immature brains, which suggests that young athletes are at the greatest risk.

To decrease this risk, clinical management of concussion has included restriction of participation in physically and mentally stressful activities until the individual is asymptomatic.

A battery of tests is commonly used to determine post-concussion recovery, which includes neurocognitive testing, postural stability assessment, and self-reported concussion symptoms. A battery is recommend for a comprehensive assessment because the tools lack sensitivity and often fail to accurately identify positive findings in acute concussion cases diagnosed by a physician. Several studies have looked at the relationships between the individual tests and there is a general consensus that the constructs of each are important and unique in the assessment of concussion. To the authors knowledge, no study has compared the scores of these tests during recovery in the middle school and high school aged population to see if each changes over time.

It has been suggested that the initial resolution of self-reported concussive symptoms may not be indicative of full recovery. Evidence exists that neurocognitive deficits remain for as many as 14 days even if the adolescent is not reporting any symptoms. In addition, it has been reported that anywhere from 15 to 50% of individuals who sustain a concussion experience prolonged symptoms and are diagnosed with (PCS). Although no current diagnostic criterion is universally agreed upon, most PCS diagnoses include a history of head injury, a dedicated time parameter of continued symptoms, and resultant behavioral criteria such as headache, apathy, irritability, dizziness and fatigue. Even though these criteria exist for guidance, the final PCS diagnosis has historically been a clinical diagnosis which runs the risk of variability among physicians. At present, few studies have examined baseline predictors of PCS and outside of gender, litigation, and neurocognitive testing, there is little evidence to identify individuals who may be at risk for developing PCS using clinical testing. Neurocognitive testing is recommended to help aid in the diagnosis of concussion and return to activity decisions, and is becoming increasingly popular to ensure that full recovery has been achieved. Clinical testing, using methods commonly employed by practicing clinicians often involves measures of balance, agility, dexterity, and computer-related tasks.

The purpose of this study was to retrospectively analyze clinical data of concussed middle school and high school aged individuals from a regional children’s hospital to determine if there was concurrent validity among three commonly used tests for assessing concussion, and how these test scores changed over time following the injury. In addition, a goal was to determine if there is predictive validity of any of these test scores at baseline to be able to predict a subsequent diagnosis of post-concussion syndrome.
that the patient rates on a scale between 0 to 6 with 0 indicating the absence of that symptom and 6 indicating a severe presence of the symptom. Lovell et al have shown the PCSS to be a reliable test with an internal consistency reliability of 0.93.

**Balance Error Scoring System (BESS)**
The BESS is an assessment tool use for an inexpensive assessment of postural stability. The test includes single, double, and tandem stance assessment on firm and foam (unstable) surfaces, each held for 20 seconds, with the athletes hands on hips and eyes closed. The BESS is the most widely used balance assessment tool for concussion evaluation and management and is based on the premise that concussion injuries lead to a decline in postural stability and an increase in measurable postural sway during fatigue or intentional demands. The BESS has been shown to be a reliable and valid clinical tool to help determine balance deficits after head trauma with an intrarater reliability ranging from 0.74 to 0.87 and an interrater reliability of 0.57 which improved to 0.98 with serial administration.

**Immediate Post-Concussion and Cognitive Testing (ImPACT)**
The ImPACT is the most widely used computerized neurocognitive test. The test has several subscales that evaluate intentional processes, verbal recognition memory, visual working memory, visual processing speed, reaction time, numeric sequencing ability and learning. Research has shown that the ImPACT test is both reliable and stable with a test-retest reliability ranging from 0.65 to 0.86 in comparison to other neurocognitive tests. The sensitivity of the ImPACT test was determined to be 81.9% and the specificity was 89.4% to rule a concussion in or out.

**Statistical Procedures**
All statistical analyses were performed using Predictive Analytics Software (PASW) Statistics version 18.0.1. The descriptive statistics were reported after running frequency distributions, mean and standard deviations when applicable. Pearson product correlation tests (in the form of a correlation matrix) were used to analyze the concurrent validity of the BESS, PCSS and individual items on the ImPACT during recovery following a concussion. Change scores were used to represent the BESS, PCSS, and ImPACT score.
subscales. Correlational strength was defined as: 
<0.25 = little or no relationship, 0.25–0.50 = fair relationship, 0.50–0.75 moderate to good relationship and >0.75 = good to excellent relationship.37

Receiver Operating Curve (ROC) statistics were used to identify cut-points within each initial concussion measure to an end point measure of post-concussive syndrome. ROC curve values, including area under the curve (AUC) were calculated (in lieu of regression analyses), because each baseline score represented a continuous data point and the authors were interested in determining whether a cut point or dedicated baseline value was associated with a long-term diagnosis of PCS. AUC measures range from .50 to 1.0, with values closer to 1.0 reflecting stronger relationships. Values below .50 reflect an inverse or negative relationship. For all analyses, an alpha value of ≤0.05 was considered statistically significant.

RESULTS
Following the review of all available charts with a diagnosis of concussion between the years of 2008 to 2010 the medical records of 106 patients satisfied the inclusion criteria. The mean age of the population was 15.38 years (SD=1.7) and ranged from 11 to 19 years. The study population consisted of 65% males and 35% females. Of the 69 males, 38 (55%) were subsequently diagnosed with PCS whereas 17 of the 37 (46%) females were diagnosed with PCS; accounting for 51.9% of the total sample. Of the 106 individuals included in the study, 40.6% of patients previously suffered a diagnosed concussion and 12.3% had suffered multiple concussions. A majority of the concussions occurred during an athletic event with football as the sport with the greatest number of concussions at 33.0%, followed by basketball and soccer at 11.3% and 10.4% respectively. Most (59%) of concussions were first time events, followed by 28.3% which was second concussions, 8.5% which were third concussions, 2.8% which were the fourth reported concussion, and less that 1% indicating a fifth concussion. The duration of care ranged from 2 to 97 days with a mean of 15.5 days (SD=14.1). There was no single standard of care for all patients and treatment processes were variable. Table 1 outlines the descriptive statistics of the sample.

The results from the Pearson product correlation matrix are found in Table 2. The correlation matrix captured five statistically significant findings. Mild correlations existed between the change scores of the BESS and ImPACT Impulse Control (r = −0.31; p = .002) and also between the change scores of the BESS and ImPACT Verbal(r = 0.37; p = .000). The ImPACT Impulse Control score also demonstrated a weak correlation with the ImPACT Visual (r = −0.24; p = .015) and Verbal change scores(r = −0.22; p = .026). Lastly, the ImPACT Verbal and Visual change scores demonstrated a mild correlation (r = 0.31; p=.001).

When assessing the ROC baseline cut points, 5 of the findings found in Table 3 yielded an area under the curve (AUC) greater than 0.50, but only one was statistically significant. The PCSS yielded an AUC of 0.34, suggesting an inverse relationship with PCS (lower baseline scores are more commonly associated with PCS). The results of the ROC curves are displayed in Table 3.

DISCUSSION
The purpose of this study was to determine the concurrent validity of the change scores of the ImPACT,
BESS and PCSS, as well as determine the predictive validity of this battery of tests along with the duration of treatment and a history of previous concussions in determining PCS. Neurocognitive testing, which involves examination methods that are not routinely available to a practicing sports medicine clinician, along with a balance assessment and a self-reported symptom scale (two methods available and frequently used), are commonly used to evaluate patients after a concussion and have been widely recommended for determining when an individual can return to normal activity.8-10,12,14-16 To the best of the authors knowledge, no study to date has examined the concurrent validity and predictive validity of clinical testing of the BESS, PCSS, or ImPACT scores in the middle school and high school aged population.

This study found low concurrent validity between the change scores of the battery of tests despite the fact that there were five statistically significant results in the correlation matrix. The test items in the correlation matrix that were most closely related was the

<table>
<thead>
<tr>
<th>Test</th>
<th>Area Under the Curve (AUC)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCSS Initial Score</td>
<td>.377*</td>
<td>.03*</td>
</tr>
<tr>
<td>BESS Initial Score</td>
<td>.481</td>
<td>.733</td>
</tr>
<tr>
<td>ImPACT Impulse Control Initial Score</td>
<td>.447</td>
<td>.351</td>
</tr>
<tr>
<td>ImPACT Verbal Initial Score</td>
<td>.535</td>
<td>.538</td>
</tr>
<tr>
<td>ImPACT Visual Initial Score</td>
<td>.556</td>
<td>.319</td>
</tr>
<tr>
<td>ImPACT Reaction Time Initial Score</td>
<td>.544</td>
<td>.433</td>
</tr>
<tr>
<td>ImPACT Visual Motor Initial Score</td>
<td>.555</td>
<td>.325</td>
</tr>
<tr>
<td>Duration of Treatment</td>
<td>.571</td>
<td>.643</td>
</tr>
<tr>
<td>History of Concussion</td>
<td>.494</td>
<td>.919</td>
</tr>
</tbody>
</table>

*Significant at the p ≤ .05 level
ImPACT verbal change score and the BESS change score with $r = 0.37$; however, this only represents a fair correlation. The other relationships found in the correlation matrix ranged from weak to mild. At present, it is suggested that the use of a battery of tests to assess recovery and provide guidance for the return to activity following a concussion is the most meaningful and discriminative method for appropriate clinical practice.\textsuperscript{8,10,12,20,31} The results of this study lend tangential support to this suggestion, as improvement in one test did not correlate strongly with improvement in any of the other tests included in this study. The current findings are similar to those from a previous study performed with college athletes that found that not all concussion tests scores for an individual with a concussion improve at a similar rate.\textsuperscript{38} A myriad of symptoms are present following a concussion and the variable tests may be necessary to capture all of the potential representations.

The second major finding in this study is that there is no evidence that a baseline score of the BESS, ImPACT, or PCSS had the ability to accurately identify individuals who will be diagnosed with PCS at a later date. The PCSS was the only measure with a statistically significant ROC curve, however, the area under the curve was 0.38, which suggests those with a higher PCSS score (indicating greater severity of injury), have a decreased likelihood of developing PCS (and vice versa). These findings are counterintuitive, and it is unlikely that a patient with a lower initial PCSS score is more likely to develop PCS when compared to a patient with a higher initial score. Ironically, this finding is consistent with findings reported in a systematic review by Carroll et al.\textsuperscript{25} who concluded the long term prognosis for persistent symptoms following a concussion may be attributable to factors other than the severity of initial impairments.

Another possible explanation for this curious result is that there is uncertainty in the literature of the definition of PCS.\textsuperscript{38} PCS is a clinical diagnosis, representing signs and symptoms and a subsequent decision made by a physician. The current study utilized medical charts from multiple physicians and because PCS is a clinical diagnosis and is considered a syndrome, the diagnosis may have varied from one physician to another. Another possible explanation is that subjects with higher initial scores on the PCSS may have been treated more conservatively. The typical protocol following a concussion is physical and mental rest.\textsuperscript{8,12,40-42} The individuals with few impairments, as measured by the battery of tests, may have begun activity earlier than those with multiple impairments, or they may not have followed the protocol as closely because they were less symptomatic.

**Limitations**

This study was retrospective in nature, which reduces the amount of control over extraneous variables such as post-concussive treatment. This is evident by the large variance in time from initial visit to final visit. Also, the inclusion criteria required subjects to have at least two visits during a single episode of care in which all three of the tests were performed. There is the risk that certain subjects performed these tests multiple times because of poor performance during the initial test. This may have skewed the data to include patients with a higher level of impairment. Also, the patients who were experiencing less impairment after a single visit may not have returned for a second visit.

**CONCLUSION**

Following a concussion injury (in this sample of middle and high school aged athletes), a single clinical measurement tool does not appear to be adequate in order to accurately determine whether the resolution of symptoms and impairments is sufficient to permit a full return to activity. It is important to monitor progress using a battery of tests to ensure confidence of clinical decision making. Most concussion oriented clinical tests do not appear to have any predictive value for determining the risk of an individual for experiencing prolonged post-concussive symptoms. Future research is needed to identify variables that are associated with PCS and until this is more clearly identified individuals with concussion injuries should be treated with extreme caution and careful assessment should be conducted prior to return to activity.

**REFERENCES**

1. Faul M, Xu L, Wald MM, Coronado VG. Traumatic brain injury in the United States: emergency department visits, hospitalizations, and deaths. Atlanta (GA): Centers for Disease Control and
Prevention, National Center for Injury Prevention and Control; 2010.


42. McGrath N. Supporting the student athletes return to the classroom after a sport-related concussion. *J Athl Train* 2010;45(5):492-8.