Sensitivity and specificity of the ImPACT Test Battery for concussion in athletes

Philip Schatz a,*, Jamie E. Pardini b, Mark R. Lovell b, Michael W. Collins b, Kenneth Podell c

a Saint Joseph’s University, 5600 City Avenue, Post Hall 222, Philadelphia, PA 19131, USA
b UPMC Center for Sports Medicine, Pittsburgh, PA, USA
c Henry Ford Health System, Detroit, MI, USA

Accepted 3 August 2005

Abstract

This study explored the diagnostic utility of the composite scores of Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) and Post Concussion Symptom Scale scores (PCSS). Recently concussed high school athletes (N = 72) were tested within 72 h of sustaining a concussion, and data were compared to non-concussed high school athletes with no history of concussion (N = 66). Between-groups MANOVA revealed a significant multivariate effect of concussion on test performance (p < .001); univariate ANOVAs revealed all six measures contributed to the between-groups differences. A discriminant function analyses was conducted to measure the ability of the five ImPACT composite scores, as well as the PCSS to classify concussion status. One discriminant function was identified that consisted of the Visual Memory, Processing Speed, and Impulse Control composite scores PCSS, which correctly classified 85.5% of the cases. Approximately 82% of participants in the concussion group and 89% of participants in the control group were correctly classified. Using these data, the sensitivity of ImPACT was 81.9%, and the specificity was 89.4%. As part of a formal concussion management program, ImPACT is a useful tool for the assessment of the neurocognitive and neurobehavioral sequelae of concussion, and can also provide post-injury cognitive and symptom data that can assist a practitioner in making safer return to play decisions.

Keywords: Processing Speed; Impulse Control; ImPACT; Concussion

The diagnosis, treatment, and management of sports related concussion has gained widespread attention, in recent years, in the fields of neuropsychology and sports medicine. This increase in interest is not spurious, given that approximately 300,000 sports-related mild traumatic brain injuries (MTBIs) occur each year; at the high school level alone, approximately 62,816 sports-related concussions occur yearly, with high school football players acquiring 69% of recorded concussions (Powell & Barber-Foss, 1999). Despite the vast number of concussions suffered on a yearly basis, data to aid in the diagnosis and management of sports concussion have only begun to emerge. Such clinical data may assist in the development of empirically-based concussion management guidelines, and can surely contribute to athletes’ return-to-play decisions. Athletes with a history of concussion have been shown to have cumulative

* Corresponding author. Tel.: +1 610 660 1804; fax: +1 610 660 1819.
E-mail address: pschatz@sju.edu (P. Schatz).

© 2005 National Academy of Neuropsychology. Published by Elsevier Ltd. All rights reserved.

doi:10.1016/j.acn.2005.08.001
cognitive effects, as well as decreased cognitive performance relative to non-concussed and fully-recovered peers, as well as athletes with a history of only one previous concussion (Collins, Grindel, et al., 1999; Collins et al., 2002; Moser & Schatz, 2002). Further, as sustaining a cerebral concussion has been shown to increase the likelihood of sustaining another concussion (Guskiewicz et al., 2003), prevention of premature return-to-play following concussion may decrease the likelihood of sustaining subsequent concussions during athletic competition and participation.

In the early phases of concussion management, grading scales were utilized to classify the severity of injury and make return to play decisions. At least 14 (Collins, Lovell, & McKeag, 1999) different return-to-play scales and 25 (Johnston, McCrory, Mohtadi, & Meeuwisse, 2001) different injury-grading systems exist, each offering often differing recommendations for the diagnosis and management of concussion, including guidelines for making return-to-play decisions. These scales, though perhaps beneficial in classifying concusive injuries on a grand scale, were not empirically based, and the management and return-to-play strategies recommended were based on subjective clinical experience rather than empirical evidence or research. In addition, these scales were unable to account for inter-individual variations in injury features and recovery course (Collins, Lovell, et al., 1999). Overall, the traditional grading-scale approach provided a predetermined recovery period following concussion based upon the concussion grade (though this period often varied from one grading or return-to-play system to the next).

Also, in this model, an athlete was not allowed to return to play unless they reported being asymptomatic. Because of the dependence on self-report of symptoms and the lack of individualized cognitive assessment tools and return-to-play guidelines, it was impossible to ensure that an individual injured athlete had regained cognitive functioning consistent with pre-injury levels (the defining factor in the recovery from a MTBI), regardless of symptom status (Collins, Stump, & Lovell, 2004).

Barth et al. (1989) first developed the initial prospective study of sport-related concussion, which established the use of an athlete’s pre-season baseline levels of performance for comparison to post-concussion levels. This approach addressed the problem of subjectivity inherent in the many concussion grading and guideline scales. With the advent of various neuropsychological testing batteries, and the increasing need for a more efficient way of assessing and managing concussive injuries, the advent of newer computerized neuropsychological testing batteries, in that they were able to utilize efficiently variable test paradigms, which lessened the likelihood of a practice effect, and can be administered in larger groups and supervised by team physicians and athletic training staffs. However, the Vienna committee was concerned with the reliability and sensitivity of such testing batteries, and called for research into these aspects of the new test batteries. This endorsement of neuropsychological testing as a key component in concussion management was reaffirmed by a second international conference in Prague in 2004 (McCleney et al., 2005).

In an attempt to increase the availability of neuropsychological testing within the athletic environment, Lovell et al. developed the ImPACT (Lovell, Collins, Podell, Powell, & Maroon, 2000; Maroon et al., 2000) Test Battery. Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) consists of three main parts: demographic data, neuropsychological tests, and the Post-Concussion Symptom Scale (PCSS), and these three sections combine to provide data to assist in accurately assessing and managing concussive injuries. The demographic data section supplies all relevant sport, medical, and concussion history information. The test modules and symptom checklist sections are described in detail below.

ImPACT (version 2.0) consists of six neuropsychological tests, each designed to target different aspects of cognitive functioning including attention, memory, Processing Speed, and reaction time. From these six tests, four separate...
Table 1

<table>
<thead>
<tr>
<th>Test name</th>
<th>Neurocognitive domain measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Memory</td>
<td>Verbal recognition memory (learning and retention)</td>
</tr>
<tr>
<td>Design Memory</td>
<td>Spatial recognition memory (learning and retention)</td>
</tr>
<tr>
<td>X’s and O’s</td>
<td>Visual working memory and cognitive speed</td>
</tr>
<tr>
<td>Symbol Match</td>
<td>Memory and visual-motor speed</td>
</tr>
<tr>
<td>Color Match</td>
<td>Impulse inhibition and visual-motor speed</td>
</tr>
<tr>
<td>Three Letter Memory</td>
<td>Verbal working memory and cognitive speed</td>
</tr>
<tr>
<td>Symptom Scale</td>
<td>Rating of individual self-reported symptoms</td>
</tr>
</tbody>
</table>

Composite scores are generated: Verbal Memory, Visual Memory, visuomotor speed, and reaction time. Table 1 lists the individual neuropsychological tests that make up the ImPACT Neuropsychological Test Battery, as well as the neuropsychological domains measured by each subtest. For a more thorough description, see Lovell et al. (2003), Iverson, Gaetz, Lovell, and Collins (2004), or Podell (2004). The Post-Concussion Symptom Scale (PCSS; Lovell & Collins, 1998) is also utilized in the ImPACT Test battery. The scale is used by many sports organizations to document and track concussion symptoms (Lovell, 1999; Lovell & Burke, 2002). The 21-symptom checklist asks the injured athlete to rate each symptom on a seven-point scale, with zero indicating no experience of a symptom and six indicating a severe symptom. This particular scale is quite useful because it presents “common” terms to describe symptoms and avoided jargon and less familiar medical terminology (e.g., sensitivity to light was used instead of photophobia).

ImPACT has been shown to be an effective tool for concussion management, and is not subject to the large practice effects sometimes seen on pencil and paper tests (Lovell et al., 2003). In the Lovell et al. (2003) study, concussed high school athletes followed for 1-week post-injury, and were compared to age-matched control participants. Control group scores on the neuropsychological indicator examined (memory composite) did not increase with multiple testing opportunities, thus indicating that the ImPACT memory composite was not hindered by a practice effect. Concussed athletes also performed much lower on the Verbal Memory Test at 36h, 4 and 7 days post-concussion compared to their individual baselines, and also took longer to rebound to baseline than did the concussed athletes in the less severe group. A follow-up study of concussed high school athletes supported the above findings that Verbal Memory and symptom indicators on the ImPACT evaluation are indicative of the concussion injury and its severity (Lovell, Collins, Iverson, Johnston, & Bradley, 2004).

Iverson, Lovell & Collins (2002) examined several validity measurements of ImPACT using 120 high school and college athletes. Concurrent validity was established by examining the composite scores and their sensitivity to the acute effects of concussion. Concussed athletes reported significantly more symptoms, and performed worse on Memory and Reaction Time Indices. Decreased performance on the Symbol Digit Modalities Test was significantly correlated with ImPACT Processing Speed and Reaction Time Indices (Iverson, Gaetz, Lovell, & Collins, 2005) and post-concussive symptoms were significantly related to decreased performance on ImPACT Reaction Time, Verbal Memory, and Processing Speed Indices (Iverson, Gaetz, Lovell, & Collins, 2004), suggesting that ImPACT is sensitive to the acute effects of concussion. Divergent validity was examined through an inter-correlation matrix of composite
scores at preseason and post concussion. The non-significant correlations found between different test components (at preseason baseline testing) indicate they do not have much shared variance, and therefore appear to be measuring different constructs.

Research to date has empirically demonstrated that the neuropsychological test indices and Post-Concussion Symptom Scale on ImPACT reflect changes occurring as a result of concussion, and that these deficits resolve, or return to baseline, upon concussion recovery. The aim of this study was to further explore the diagnostic utility of ImPACT. We compared concussed athletes with non-concussed controls using MANOVA to establish group differences on the dependent measures, the five composite scores and symptom scale score. We then conducted a discriminant function analysis to identify those measures that contributed to identifying group membership.

1. Methods

1.1. Participants

The University of Pittsburgh and Saint Joseph’s University Institutional Review Boards conducted appropriate reviews of our research with human participants and approved our study. The study extracted available data of 138 participants (72 concussed athletes, 66 non-concussed athletes with no history of concussion) from a much larger data set of approximately 1500 individuals. Criteria for inclusion in the concussion group required that participants were concussed high school athletes who were tested within 72 h of sustaining a concussion. Athletes in the control group were high school athletes with no history of concussion who completed baseline assessments. All athletes in the study participated in the Sports Medicine Concussion Program at the University of Pittsburgh Medical Center (in accordance with Field, Collins, Lovell, & Maroon, 2003). Athletes within the sample were included from high schools within the states of Pennsylvania, Michigan, Illinois, Oregon, and Maine, as part of an ongoing clinical program implementing baseline and post-injury neuropsychological testing to assist team sports medicine personnel in making return-to-play decisions after the occurrence of sports-related concussions. Concussed athletes participated primarily in football (73%) while controls participated primarily in non-contact sports (79%) such as track and tennis (Table 2). Participants in the concussion group were significantly more likely to be male, and younger in age (Table 3). Of note, age did not emerge as a significant contributor when included as a covariate in between-groups analyses, as discussed in the results section below.

Data were collected regarding athletes’ concussion status (independent variable) and ImPACT was the method by which baseline and post-concussion presentation of symptoms and constellation of cognitive abilities (dependent variables) were documented. As part of the Sports Medicine Concussion Program at the University of Pittsburgh Medical Center, all athletes in this study underwent a baseline or pre-injury evaluation, and were administered ImPACT before the 2000, 2001, or 2002 athletic seasons. Eleven individuals were excluded from the study, having achieved a score of 20 or higher on the Impulse Control, which raised questions regarding the general testing approach taken by those participants (Lovell, 2004).

<table>
<thead>
<tr>
<th>Sport</th>
<th>Concussed (N=72)</th>
<th>Control (N=66)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Football</td>
<td>52 (72.7%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Soccer</td>
<td>5 (6.5%)</td>
<td>12 (18.2%)</td>
</tr>
<tr>
<td>Ice Hockey</td>
<td>4 (5.6%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Field Hockey</td>
<td>3 (4.2%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Basketball</td>
<td>3 (3.9%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Other contacta</td>
<td>5 (6.5%)</td>
<td>2 (3.0%)</td>
</tr>
<tr>
<td>Non-contactb</td>
<td>0 (0%)</td>
<td>52 (78.8%)</td>
</tr>
</tbody>
</table>

* Other contact sports: softball, volleyball, and gymnastics.

b Non-contact sports: track and tennis.
### Table 3
Demographic data for concussed and non-concussed groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Concussion group</th>
<th>F/χ²</th>
<th>Significance</th>
<th>ES²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Concussed</td>
<td>Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>16.5 (2.3)</td>
<td>17.3 (1.7)</td>
<td>5.35</td>
<td>.022</td>
</tr>
<tr>
<td>Education</td>
<td>10.4 (2.0)</td>
<td>10.8 (1.8)</td>
<td>2.1</td>
<td>.149</td>
</tr>
<tr>
<td>Gender</td>
<td>79.2%</td>
<td>43.9%</td>
<td>0.23</td>
<td>.60</td>
</tr>
<tr>
<td>Female</td>
<td>20.8%</td>
<td>56.1%</td>
<td>0.28</td>
<td>.60</td>
</tr>
<tr>
<td>Special education</td>
<td>3%</td>
<td>2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning disability</td>
<td>3%</td>
<td>1%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* ES = Effect Size; φ for χ² analyses, d for ANOVA
* χ² = 18.2, significance = .001, ES: φ = .38

### 1.2. Materials and procedures
All baseline data were collected during the off-season (i.e., before preseason contact drills). ImPACT is inclusive of a standardized demographic questionnaire that requires the athlete to document relevant educational, sports participation, and personal medical history. Test administrators were trained to define concussion as a “traumatically induced alteration in mental status that may or may not be accompanied by a loss of consciousness”, based on the standard American Academy of Neurology nomenclature (AAN, 1997). High school athletes in our study population who experienced a cerebral concussion were referred for and received post-injury ImPACT evaluation within 72 h of injury. In-season concussions were diagnosed on the basis of the following criteria (in accordance with Field et al., 2003): (1) any observable alteration in mental status or consciousness on following a blow to the head or body during sport participation, and/or (2) the presence of LOC and/or anterograde or retrograde amnesia identified in an on-field examination, and/or (3) any self-reported symptoms such as cognitive “fogginess”, headache, nausea and/or vomiting, dizziness, balance problems, and visual changes after a collision involving the head or body. Certified athletic trainers or team physicians who were present on the sideline at the time of injury made the initial diagnosis of concussion.

### 1.3. Analyses
One-way analyses of variance were conducted to identify between-group differences on age and education, and χ²-square analyses were conducted to identify between-group differences on gender, handedness, diagnosis of learning disability, and history of special education. MANOVA was conducted to establish between-group differences on the dependent measures. Stepwise discriminant analysis was performed to identify variables that discriminated between concussion groups, with total score on the Symptom Scale, and performance on the five ImPACT Composite scores as the independent variables. All analyses were conducted with an alpha level of p < .05 using SPSS statistical software (SPSS, 2003). Effect Size was reported as either a correlation coefficient for the discriminant analysis, Cohen’s d for ANOVA, or Partial Eta² for MANOVA.

### 2. Results
Demographic variables were analyzed to establish between-group homogeneity, with no differences noted between concussion history groups on age, education, handedness, history of special education, or diagnosis of learning disability. Males were significantly more likely [χ²(1) = 18.2, p = .001, φ = .36] to be in the concussion group (79%) than the control group (44%), although this is consistent with the available literature, in that many sports traditionally played by males (especially football) have significantly higher rates of concussion per athletic exposure (Powell & Barber-Foss, 1999). Athletes in the concussion group were significantly younger than athletes in the control group (16.5 versus 17.3) [F(1,136) = 5.35; p = .022; d = .38] While none of the controls had a history of concussion, 83% of those in the concussed group had a history of one previous concussion, and 17% had a history of two or more concussions. Between groups analyses of demographic data are provided in Table 3.
A multivariate analysis of variance (MANOVA) was performed with concussion group as the independent variable and the five ImPACT composite scores and the symptom scale score as the dependent variables. Hotelling’s Trace revealed a significant multivariate effect of concussion group on cognitive performance ($F(6,131) = 16.6; p = .001$). Univariate ANOVAs revealed significant effects of concussion group on Verbal Memory ($F(1,136) = 32.4; p = .001$), Visual Memory ($F(1, 136) = 34.9; p = .001$), reaction time ($F(1, 136) = 43.6; p = .001$), Processing Speed ($F(1, 136) = 61.1; p = .001$), and symptom scale scores ($F(1, 136) = 39.6; p = .001$), but no effect of concussion group on Impulse Control scores ($F(1, 136) = 0.3; p = .87$) (see Table 4 for Effect Size). Age did not emerge as a covariate, and did not account for a significant amount of between-group variance ($F(6,131) = 1.58; p = .16; Partial Eta^2 = .07$).

A stepwise discriminant analysis was conducted with the total score on the post-concussion symptom checklist and the five ImPACT composite scores. One discriminant function identified post-concussion checklist scores, Processing Speed composite, Visual Memory composite, and Impulse Control composite as significant factors ($\chi^2(4) = 74.4, p = .0001$), with 85.5% of cases correctly classified. Eighty-two percent of participants in the concussed group and 89.4% of participants in the non-concussed group were correctly classified. Means and standard deviations for the variables in the equation are provided in Table 4 and the classification matrix is provided in Table 5. The Eigenvalue for these data (0.742) suggested that the discriminating power of the function was quite high, with a canonical correlation of 0.653. The significance of the discriminant function and the indices of power are shown in Tables 5 and 6, respectively. Standardized and canonical correlation coefficients are provided in Table 7.

Using the classification results of the DFA, the combined sensitivity of ImPACT and the symptom score (or the probability that a test result will be positive when a concussion is present) is 81.9%, and the specificity (or the probability that a test result will be negative when no concussion is present) is 89.4%.
that a test result will be negative when a concussion is not present) is 89.4%. The positive likelihood ratio (PLR—ratio between the probability of a positive test result given the presence of a concussion and the probability of a positive test result given the absence of a concussion) is 7.73:1. The negative likelihood ratio (NLR—ratio between the probability of a negative test result given the presence of a concussion and the probability of a negative test result given the absence of concussion is .20:1. The positive predictive value (PPV—probability that a concussion is present when the test is positive) is 89.4% and the negative predictive value (NPV—probability that a concussion is not present when the test is negative) is 81.9%.

3. Discussion

Athletic participation is a daily activity for many youth, adolescents, and young adults, placing them at risk for sports-related concussion. Multiple sports-related concussions have been shown to result in impaired neurocognitive functioning post-injury (Collins, Lovell, et al., 1999), decreased performance on baseline testing (Moser & Schatz, 2002; Moser, Schatz, & Jordan, 2005), and place the athlete at increased risk for more severe on-field markers of concussion, such as loss of consciousness, anterograde amnesia, and confusion (Collins et al., 2002). Neuropsychological baseline assessment paradigms facilitate the detection and management of mild neurocognitive changes in athletes who have sustained a concussion (Schatz & Zillmer, 2003), and computerized assessment of sports-related concussion offers unique advantages to the athlete, athletic trainer, team physician, consulting neuropsychologist, coach, and the athlete’s family (McKeever & Schatz, 2003).

The current study demonstrates that the ImPACT computerized test battery is both a sensitive and specific instrument for the assessment of the neurocognitive and neurobehavioral sequelae of concussion. More importantly, as a whole, ImPACT’s PPV/NPV and PLR/NLR are very high, so poor performance on the composite score (either relative to baseline or compared to normative sample) yields a very strong likelihood of reflecting concussion. This can only serve to improve our ability to diagnose and subsequently treat sports concussion.

The difficult part of sports-concussion has always been having highly sensitive but accurate techniques to detect the presence of concussion, given that standard techniques of a neurological examination, neuroimaging, and electrophysiological techniques are notoriously poor at detecting concussions. It appears that ImPACT is clearly sensitive and specific in detecting sports concussions, at least relative to healthy controls.

It is not perfectly clear what the relatively weaker specificity in this study reflects. Not all concussed athletes show symptoms or cognitive deficits after a concussion, especially given that some were tested up to 72 h after the concussion. It is conceivable that some of the concussed athletes had “recovered” enough at the time of testing that they truly were back to baseline. This plausible scenario would artificially lower the sensitivity.

Our results show that ImPACT provides post-injury cognitive and symptom data that can assist a practitioner in making safer return to play decisions. Using the neuropsychological data provided by ImPACT alone, 85% of cases were correctly classified. In this study, cognitive impairments in Visual Memory, Processing Speed, and Impulse Control along with symptom status effectively classified most of the concussed and control athletes. When used appropriately, by a trained neuropsychologist and in conjunction with a thorough clinical interview, the utility of this instrument is likely to be further enhanced. Therefore, ImPACT can serve as an effective tool in the concussion management process. While our results are very strong at the group level, our findings do not directly address decision making at the individual level. For example, what combination of post-concussion symptom scores and ImPACT composite score changes are needed to identify a concussion? This will be the focus of future research using ImPACT.

### Table 7

<table>
<thead>
<tr>
<th>Factor</th>
<th>Standardized coefficients</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Memory composite</td>
<td>0.347</td>
<td>0.588</td>
</tr>
<tr>
<td>Processing Speed composite</td>
<td>0.648</td>
<td>0.778</td>
</tr>
<tr>
<td>Impulse Control composite</td>
<td>0.361</td>
<td>-0.016</td>
</tr>
<tr>
<td>Post-Concussion Symptom Scale</td>
<td>-0.475</td>
<td>-0.626</td>
</tr>
</tbody>
</table>
This study is not without its limitations. While post-concussion evaluations were performed on a prospective “as needed” basis, the current study was retrospective in nature. This retrospective “concussed versus athletic control” design yielded groups with significantly different ages, genders, and history of concussion. The concussion group was comprised of a significantly greater percentage of males, and was significantly (although less than 1 year) younger. In spite of the fact that age did not emerge as a significant predictor of between-group variance, a prospective truly matched control design would be more appropriate. While all but the Impulse Control InPACT composite score contributed to a multivariate between-groups difference (on MANOVA), Impulse Control emerged (along with Visual Memory, Processing Speed, and Symptom Scale scores) as variables that contributed towards discriminating between groups. It appears that shared variance among correlated predictor variables may have “cancelled out” certain predictor variables (such as Verbal Memory) and thus allowing other variables (such as Impulse Control) to contribute unique variance to the discriminant analysis. Concussed athletes participated exclusively in “at-risk” contact sports and had a history of concussion, while non-concussed athletes participated in low risk non-contact sports and had no history of concussion. Participants in the concussion group participated primarily in football, which in part explains why there were more males in this group.

ImPACT offers a thorough assessment of changes in cognitive functioning and symptom status following concussion, which is consistent with Concussion in Sport group recommendations that neuropsychological assessment become an integral aspect of concussion diagnosis and management (Aubry et al., 2002; McCrory et al., 2005), and that athletes should not return to competition until they are asymptomatic. Based on the recommendations of the Concussion in Sport groups and the current findings, it is recommended that any athlete participating in contact sport receive a baseline neurocognitive evaluation of some sort, whether computerized or traditional paper and pencil testing, dependent on the resources and preferences of the athlete’s program. This will not only improve the ability of sports medicine personnel to manage recovery and return to play decisions but also provide objective criteria for determining when an athlete is ready to return to competition.

In summary, the current study highlights the value of neuropsychological assessment in the decision making process of return to play for concussed athletes. The use of ImPACT as a tool for monitoring recovery and recovery of function is strongly recommended. It is also important to note that while this study focused on college football athletes, the findings are likely to be applicable to other contact sports as well. Further studies are needed to validate these findings in other populations of athletes.

References


