

Validity of the Immediate Post Concussion Assessment and Cognitive Testing (ImPACT)

Bara Alsalaheen¹ · Kayla Stockdale¹ · Dana Pechumer¹ · Steven P. Broglio²

Published online: 12 April 2016
© Springer International Publishing Switzerland 2016

Abstract

Background The immediate post concussion assessment and cognitive testing (ImPACT) is the most widely used concussion assessment tool. Despite its popularity, it is unclear if validation studies for the ImPACT test covered all aspects of validity to support its widespread use in research and clinical practice.

Objective The purpose of this report is to review literature surrounding the validity and the utility of the ImPACT test.

Data sources and appraisal A systematic review of relevant studies in PubMed, CINAHL, and PsycINFO was carried out. Studies were evaluated using the STROBE (strengthening the reporting of observational studies in epidemiology) or the STARD (standards for reporting of diagnostic accuracy) criteria.

Results The literature search yielded 5968 studies. Sixty-nine studies met the inclusion criteria and were included in the qualitative review. Although the convergent validity of ImPACT was supported, evidence of discriminant and predictive validity, diagnostic accuracy and responsiveness was inconclusive. The utility of the ImPACT test after acute symptom resolution was sparse. The review found many factors influenced the validity and utility of ImPACT scores.

Conclusion Clinicians must consider the benefit of ImPACT testing for their patients on a case-by-case scenario and must take the psychometric properties of the test into account when interpreting results.

Electronic supplementary material The online version of this article (doi:10.1007/s40279-016-0532-y) contains supplementary material, which is available to authorized users.

✉ Bara Alsalaheen
alsalahe@umflint.edu

Kayla Stockdale
stockkay@umflint.edu

Dana Pechumer
dpechume@umflint.edu

Steven P. Broglio
Broglio@umich.edu

¹ Physical Therapy Department, University of Michigan-Flint, 2157 William S. White Building, 303 E. Kearsley Street, Flint, MI 48502-1950, USA

² Neurotrauma Research Laboratory, University of Michigan, Ann Arbor, MI, USA

Key Points

The Immediate Post concussion Assessment and Cognitive Testing (ImPACT) test is one of the most widely used concussion assessment tools. This study reviews the validity of the ImPACT test, and potential factors that affect its validity in clinical practice.

The current review concluded that the convergence validity of the ImPACT test is supported. However, limited evidence exists to support its predictive validity and responsiveness.

The validity and utility of the ImPACT test in clinical practice is affected by testing environments, instructions, exertion, and sleep prior to the test. Additionally, 10–35 % of athletes were able to intentionally underperform on the test.

1 Introduction

Despite extensive concussion-related clinical research in the last decade, concussion diagnosis continues to be based on the subjective opinion of the medical provider administering the clinical examination. To support the examination process, many consensus statements emphasize the multifaceted nature of concussion assessment that includes neurocognitive assessment, postural stability assessment, and self-reported symptoms [1, 2]. Over the last 20 years, computerized neurocognitive assessments have replaced the traditional paper and pencil tests for concussion, and are considered the “cornerstone” of concussion assessment [1]. This is evidenced by a survey of athletic trainers in the National Collegiate Athletic Association (NCAA) in the USA, in which 90 % utilized neurocognitive testing in the assessment of concussion [3]. Additionally, baseline neurocognitive testing is mandatory in the National Football League and the National Hockey League. In the absence of baseline scores, post-concussion performance is compared to age-validated reference values. Despite the growing popularity of neurocognitive testing, some remain skeptical about its validity and clinical utility in concussion management [4, 5].

Since its inception in the 1990s, the Immediate Post Concussion Assessment and Cognitive Testing (ImPACT) has been the most widely used neurocognitive testing program in concussion management [6]. Eighty-nine percent of the NCAA athletic trainers who utilized neurocognitive testing reported using the ImPACT test [3]. Additionally, the ImPACT is used for assessment of concussions in sport and non-sport injuries of all ages and levels of play [7, 8]. ImPACT is often used to quantify cognitive declines and recovery in the days and weeks following concussion and has been used as the standard by which other evaluative tools are measured. The ImPACT test reports four composite scores including verbal memory, visual memory, visual-motor processing speed, and reaction time [9]. As concussion results in a constellation of signs and symptoms that affect cerebral functioning [1], a clear understanding of the validity literature germane to diagnosis, prognosis, and management of concussion is needed to improve clinical care.

The current standards for validation process requires evidence from multiple areas including the content, criterion, construct, response process, internal structure, and diagnostic accuracy [10]. Additionally, other factors can affect the validity of an outcome measure in practice and must be examined; these factors include responsiveness to intervention, cultural and language equivalence, administration environment, and effects of effort and fatigue on test performance. While several investigations have looked

at some of these components, it is unclear if the validity of the ImPACT has been demonstrated for all the above-mentioned domains [4]. Therefore, the purpose of this review is to describe and evaluate studies that examined the ImPACT validity, to classify the retrieved studies according to the type of validity examined, and to describe the validity metrics reported for the ImPACT battery.

2 Methods

2.1 Search Strategy

An initial electronic literature search of published studies between January 1999 and November 2014 was completed. An updated search was completed on November 2015 for potential studies published since the original search. Studies published before 1999 were not included in this search as an earlier version of the test did not have separate output scores for verbal and visual memory. The searched databases included PubMed, CINAHL, and PsycINFO. The current search was a part of a larger project aiming to examine the reliability, validity, and clinical utility of the ImPACT test in the management of concussion. Therefore, the search terms were identical to a previously published part of this project [11]. The search was completed using the following search terms: “ImPACT OR immediate post-concussion assessment and cognitive test OR impact testing OR neurocognitive testing OR neurocognitive OR neuropsychological testing OR neuropsychological” AND “concussion OR mTBI OR mild traumatic brain injury OR post concussive syndrome OR mild head injury OR closed head injury” [11]. For both initial and updated searches, the search filters of English language publications and studies that included human subjects were applied. Review articles, abstracts, case studies, editorials, and gray literature were excluded from the analysis. Gray literature was excluded because it does not often include enough information detailing the methods and outcomes of the study. Therefore, its quality cannot be assessed using the tools employed in this review. Additionally, a hand search of the reference lists of included studies and an electronic search on the ImPACT test website were performed.

2.2 Study Selection

Studies were included if participants completed the ImPACT test with or without other measures, and if studies could be classified into one or more validity types detailed below. Since the ImPACT performance is reported through four composite scores, studies describing the ImPACT test modules or subscales are not used in clinical practice and,

therefore, were excluded from this review. Studies that did not separately report verbal and visual memory (ImPACT 1.0) were also excluded. Additionally, studies examining the ImPACT reliability were excluded since they were reviewed elsewhere [11].

2.3 Data Extraction and Classification

Two reviewers (DP and KS) identified potential studies following independent review of the titles and the abstracts. The same two reviewers completed an independent review of potential studies, extracted the data using a piloted Excel spreadsheet, and independently classified the selected studies according to the type of validity examined. Disagreement on the extracted data or study classification was resolved by consensus. If disagreement remained, a third reviewer (BA) was consulted to resolve the disagreement. The studies were classified into diagnostic accuracy validity, construct validity or criterion-related validity studies based on the classification described by Portney and Watkins (Table 1) [10]. The lack of a “gold standard” test for concussion assessment created a unique situation for validating the ImPACT test. When the ImPACT test was validated against other tests with known psychometric limitations, these studies were classified as construct validity (convergent and discriminant) studies rather than criterion-related validity. The correlation coefficients for construct validity studies were interpreted according to the guidelines proposed by Portney and Watkins where correlation coefficients exceeding 0.75 indicate good correlation; coefficients between 0.50 and 0.75 indicate moderate correlation; coefficients between 0.25 and 0.50 represent fair correlation; and coefficients ≤ 0.25 represent little or no correlation [10]. Additionally, we reviewed studies that examined factors affecting the validity and utility of the ImPACT test and are not classified in the above-mentioned

categories. These factors include test-taking environment, invalid baseline scores and “sandbagging” (i.e., intentionally underperforming), learning disability (LD) or attention deficit hyperactivity disorder (ADHD), cultural equivalence, alternate test form equivalence, and the effect of physical exertion on the test. Furthermore, the utility of the test after the resolution of symptoms was reviewed.

2.4 Assessment of Reporting Quality

The reporting quality of each study was assessed using the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) instrument [12]. The STROBE instrument assesses the strength of reporting of observational studies in epidemiology. It addresses 22 fundamental aspects of the methods and reporting of observational studies. Each aspect was assigned a numerical value of one when explicitly described, and zero if inadequately described or absent. As such, a total score out of 22 was reported to reflect the overall reporting quality of each study. Higher STROBE scores reflect a better reporting quality [12].

The studies examining the diagnostic accuracy of ImPACT were evaluated for the reporting quality using the STARD instrument (Standards for the Reporting of Diagnostic accuracy studies) [13]. The STARD instrument consists of 25 items with each item assigned a numerical value of one when explicitly described, and zero if inadequately described or absent. Therefore, a maximum score of 25 was reported to reflect the overall reporting quality of each study. Higher STARD scores reflect a better reporting quality for diagnostic accuracy studies. Reporting quality assessment was performed independently by two reviewers (DP and KS). Disagreement was resolved by consensus. If disagreement remained, a third reviewer (BA) was consulted. Reviewers demonstrated excellent inter-rater

Table 1 Summary and description of validity types considered for this review

Validity type	Question
Diagnostic accuracy validity	What is the sensitivity, specificity, and the predictive values of the ImPACT test to diagnose concussion?
Construct validity	The known groups methods: does the ImPACT test document the known differences between participants with recent concussion and participants without concussion?
	Convergent validity: does the ImPACT test correlate highly or yield similar results as other measures believed to measure the same construct?
	Discriminant (divergent) validity: does the ImPACT test exhibit low correlations with measures believed to assess different characteristics?
Criterion related validity	Factor analysis: does the construct measured by ImPACT contain one or more underlying dimensions, or different theoretical components?
	Predictive validity: can the ImPACT test be used to predict a future outcome or criterion score?
	Concurrent validity: does the ImPACT test demonstrate the ability to test for concussion assessment when administered at approximately the same time?

ImPACT Immediate Post Concussion Assessment and Cognitive Testing

reliability in reporting quality scoring as indicated by the two-way mixed, single measures intraclass correlation coefficient [$ICC_{3,1} = 0.96$ (0.94, 0.98)].

3 Results

3.1 Search Yield

The initial electronic and manual search yielded 5968 articles. Duplicates were removed ($n = 321$) and an independent review of the abstracts excluded 5566 records, leaving 81 articles for full-text review. After reviewing the 81 articles and reapplying the same inclusion criteria, an additional 12 articles were excluded. Excluded studies used version 1.0 of the test [14–17], did not contain cognitive performance scores from the ImPACT test [18–23], or used the ImPACT subscales [24, 25] (Fig. 1). Therefore, 69 studies were included in the qualitative review.

Twenty-six studies examined construct validity whereas eight and four studies examined criterion-related validity and diagnostic accuracy, respectively. Additionally, 38 studies examined miscellaneous validity and utility considerations for the ImPACT test (Table 2).

3.2 Reporting Qualities of the Studies

The reporting quality for the reviewed studies ranged from moderate to high as indicated by STROBE scores ranging from 14 to 21. For the diagnostic accuracy studies, the

reporting quality ranged from low to moderate (13 to 17) on the STARD. Participants in all validity studies varied in age, sex, concussion history, testing environment, feedback, and time since injury. Additionally, methodological considerations varied across studies and are detailed throughout Sect. 4 and the Electronic Supplementary Material.

3.3 Construct Validity

3.3.1 Convergent and Discriminant Validity

Convergence among the ImPACT composite scores was examined. The ImPACT scores were additionally validated against other cognitive measures or other non-cognitive evaluative measures used in patients with concussion (Electronic Supplementary Material, Table S1). Six studies examined the inter-correlations between ImPACT composite scores [26–31]. In those studies, the relationship between the memory composite scores (verbal and visual) ranged from little to moderate in seven out of the eight analyses performed ($|r| = 0.22$ – 0.58 , $p < 0.001$). Furthermore, the relationship between reaction time and processing speed was statistically significant in all studies with little to moderate relationship ($|r| = 0.24$ – 0.56 , $p < 0.05$). When considering the relationships between memory scores (verbal or visual) to processing speed, correlation scores ranged from moderate to good ($|r| 0.29$ – 0.89 , $p \leq 0.04$), and the relationship between memory scores to reaction time ranged from little to fair ($|r| = 0.13$ – 0.30 , $p < 0.05$) (Table 3).

Fig. 1 Review yield. *ImPACT* The Immediate Post Concussion Assessment and Cognitive Testing

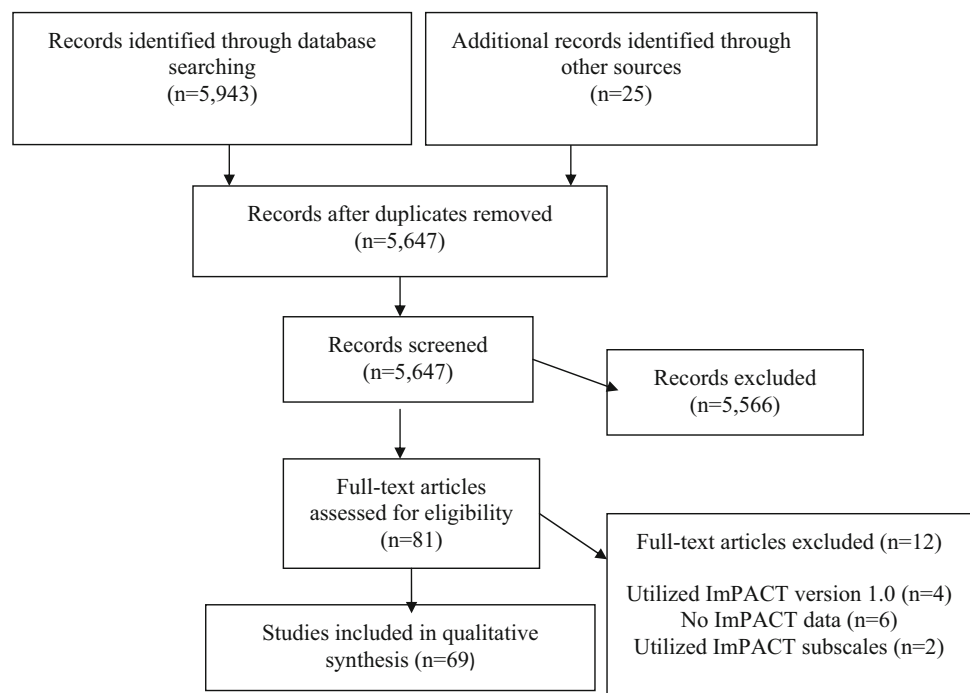


Table 2 Research yield and reporting quality for the reviewed studies

Construct validity	Convergent and discriminant validity: Allen et al. [28] (18), Alsalaheen et al. [36] (15) Barlow et al. [43] (16), Broglio et al. [39] (16), Fedor et al. [30] (17), Iverson et al. [32] (16), Kontos et al. [40] (17), Maerlender et al. [33] (18), Maerlender et al. [27] (18), Parker et al. [35] (18), Resch et al. [31] (18), Sandel et al. [41] (17), Sasaki et al. [38] (19), Schatz et al. [29] (18), Solomon et al. [26] (17), Tjarks et al. [34] (17), Vernau et al. [37] (14) Known groups methods: Iverson [50]. (20), Lau et al. [51] (17), Lovell et al. [49] (17), Ponsford et al. [44] (20), Rieger et al. [45] (17), Shores et al. [46] (21), Solomon et al. [47] (18), Solomon et al. [48] (17) Factor structure: Allen et al. [28] (18), Schatz et al. [52] (20)
Criterion related validity	Predictive validity: Barlow et al. [43] (16), Lau et al. [53] (18), Lau et al. [54] (21), Meehan et al. [56] (19), Ponsford et al. [44] (20), Sandel et al. [41] (17), Solomon et al. [47] (18), Wiebe et al. [55] (20)
Diagnostic accuracy validity	Broglio et al. [57] (14), Gardner et al. [58] (17), Van Kampen et al. [59] (14), Schatz et al. [60] (13)
Other validity considerations	Utility of ImPACT in the absence of symptoms: Broglio et al. [61] (14), Fazio et al. [62] (17) Invalid baselines, “sandbagging,” group and motivation influence: Erdal [70] (17), Kuhn et al. [73] (18), Lichtenstein et al. [66] (18), Moser et al. [67] (19), Robinowitz et al. [72] (15) Schatz et al. [71] (17), Szabo et al. [68] (17), Schatz et al. [65], (18) Learning disability/ADHD: Brooks et al. [76] (16), Elbin et al. [74] (19), Solomon et al. [47] (18), Solomon et al. [48] (17), Zuckerman et al. [75] (18) Cultural equivalence: Blake et al. [81] (19), Kontos et al. [79] (20), Ott et al. [80] (18), Shuttleworth-Edwards et al. [77] (17), Tsushima et al. [78] (16) Effects of sport season, competition and exertion on ImPACT performance: Covassin et al. [83] (16), Kontos et al. [88] (17), Majerske et al. [89] (20), McAllister et al. [86] (19), McGrath et al. [90] (17), Miller et al. [84] (17), Mulligan et al. [82] (18), Munce et al. [85] (18), Shuttleworth-Edwards et al. [87] (20) Effect of sleep: McClure et al. [91] (19), Sufrinko et al. [92] (18) Equivalence of alternate forms of the test: Resch et al. [31] (18) Utility of comparison to normative data versus individual baseline: Schatz et al. [93] (19) ImPACT as an outcome measure to assess the effects of newer technologies and intervention: Collins et al. [98] (19), Gagnon et al. [96] (18) Mihalik et al. [97] (17), Moser et al. [94] (18), Reddy et al. [95] (17)

ImPACT Immediate Post Concussion Assessment and Cognitive Testing, *ADHD* attention deficit hyperactivity disorder

Six studies examined the convergence of ImPACT scores to various cognitive outcome measures representing many cognitive domains such as speed, executive function, attention, verbal memory, visual memory, and working memory. With the exception of Iverson et al. [32], all studies examined these associations in healthy samples [26–29, 33]. Out of 29 correlation analyses completed with verbal memory, 16 significant little to fair correlations were observed ($|r| = 0.21–0.46$, $p < 0.05$). Out of 29 correlating analyses completed with visual memory, a significant little to moderate correlation was observed in nine analyses ($|r| = 0.20–0.59$, $p < 0.05$). Eighteen significant fair to moderate correlations (out of 31 analyses) were observed between visual motor processing speed and cognitive measures ($|r| = 0.26–0.70$, $p \leq 0.048$). Reaction time exhibited significant fair to moderate correlations with other cognitive measures in 14 out of the 32 completed analyses ($|r| = 0.26–0.64$, $p \leq 0.040$) (see Electronic Supplementary Material, Table S2).

Nine studies examined the relationship between ImPACT and other concussion evaluative measures including other computerized cognitive tests (e.g., Axon Sports, Concussion Resolution Index) [29], balance and

ocular motor function [34–37], and imaging [38]. Furthermore, the convergent validity of ImPACT was examined against self-reported symptoms [36, 39], depression [40], and perceived recovery [41]. ImPACT demonstrated a fair relationship with perceived recovery ($|r| = 0.30–0.44$, $p \leq 0.003$) [41]; however, the relationship between ImPACT scores to depression and cognitive symptoms was inconsistent [42]. A little to fair relationship existed between ImPACT scores to most balance measures ($|r| = 0.20–0.50$). Furthermore, ImPACT scores demonstrated a fair to moderate relationship with the King-Devick test of visual motor performance ($|r| = 0.44–0.70$), but not with advanced imaging after concussion ($p > 0.05$) (see Electronic Supplementary Material, Table S3, for additional details). Barlow et al. [43] examined the relationship between change in composite scores and change in other evaluative measures in patients with concussion, reporting a fair relationship between change in verbal memory to change in visual memory ($r = 0.31$, $p < 0.001$) and Balance Error Scoring System ($r = 0.374$, $p = <0.001$). However, there was no association between changes in memory and speed composites after concussion.

Table 3 Convergent validity among ImPACT composite scores

Study	Comparison	Visual memory	Processing speed	Reaction time
Maerlender et al. [27]	Verbal memory	$r = 0.36, p < 0.001^*$	$r = 0.29, p = 0.04^*$	$r = 0.06, p = 0.68$
Allen et al. [28]		$r = -0.58, p < 0.01^*$	$r = 0.89, p < 0.01^*$	$r = 0.30, p < 0.01^*$
Schatz et al. [29]		$r = 0.34, p = NS$	–	–
Solomon et al. [26]		$r = 0.39, p < 0.001^*$	$r = 0.30, p < 0.001^*$	$r = -0.07, p = 0.39$
Fedor et al. [30]		$r = 0.47, p < 0.001^*$	$r = 0.29, p < 0.001^*$	$r = -0.013, p < 0.05^*$
Resch et al. [31]		Form 1: $r = 0.22, p < 0.01^*$ Form 2: $r = 0.37, p < 0.001^*$ Form 3: $r = 0.33, p < 0.01^*$	Form 1: $r = 0.07, p = NS$ Form 2: $r = 0.15, p = NS$ Form 3: $r = 0.20, p = NS$	Form 1: $r = -0.05, p = NS$ Form 2: $r = 0.14, p = NS$ Form 3: $r = 0.17, p = NS$
Maerlender et al. [27]	Visual memory		$r = 0.28, p = 0.04^*$	$r = -0.12, p = 0.12$
Allen et al. [28]			$r = -0.88, p < 0.01^*$	$r = -0.12, p = NS$
Solomon et al. [26]			$r = 0.36, p < 0.001^*$	$r = -0.26, p < 0.001^*$
Fedor et al. [30]			$r = 0.33, p < 0.001^*$	$r = -0.17, p < 0.01^*$
Resch et al. [31]			Form 1: $r = 0.03, p = NS$ Form 2: $r = 0.21, p = NS$ Form 3: $r = 0.18, p = NS$	Form 1: $r = -0.29, p < 0.001^*$ Form 2: $r = -0.09, p = NS$ Form 3: $r = -0.07, p = NS$
Maerlender et al. [27]		Processing speed		
Allen et al. [28]				$r = 0.24, p < 0.05^*$
Solomon et al. [26]				$r = -0.45, p < 0.001^*$
Fedor et al. [30]				$r = -0.32, p < 0.001^*$
Resch et al. [31]				Form 1: $r = 0.33, p < 0.001^*$ Form 2: $r = -0.40, p < 0.001^*$ Form 3: $r = -0.56, p < 0.001^*$

ImPACT Immediate Post Concussion Assessment and Cognitive Testing, NS non-specified

* Statistically significant (i.e., < 0.05)

3.3.2 Known Groups Methods Validity

Several studies examined the utility of the ImPACT test in documenting the differences between groups based on injury type (concussion vs. other injuries), presence of symptoms (symptomatic vs. asymptomatic), recovery times, and history of medically treated headache.

Three studies examined if there was a difference in ImPACT scores between patients with concussion and other orthopedic injuries [44–46]. Patients with concussion demonstrated worse visual memory scores completed within 3–7 days of injury assessment in two of three studies [44, 45]. However, only one study demonstrated significant differences in verbal memory, processing speed, and reaction time at initial assessment completed in the emergency department [46]. In two studies [44, 45], ImPACT was administered at 3 months after injury; participants with concussion in both studies exhibited significantly worse visual memory, but no difference was observed for reaction time, processing speed, and verbal memory.

Athletes who sought treatment for headache prior to baseline testing ($n = 22$) exhibited worse visual memory (mean = 72.05, SD = 14.77) compared to athletes not

seeking treatment [($n = 137$), (mean = 79.42, SD = 12.32), $p = 0.012$] [47]. No differences were observed in reaction time, processing speed, or verbal memory based on headache history. Furthermore, Solomon et al. [48] revealed no differences in ImPACT scores based on treatment for headache.

Of the 60 cheerleaders who denied an increase in symptoms after concussion, 20 (33 %) exhibited a reliable decline in at least one composite score leading to identification of 33 % more concussions within the first week of injury [49].

When athletes were retrospectively classified based on recovery times into simple and complex recovery [50, 51], the complex group exhibited worse initial visual memory and processing speed, but not verbal memory [50, 51]. Differences on initial reaction time were inconclusive [50, 51].

Taken together, visual memory demonstrated validity in discriminating between patients with concussion compared to orthopedic injuries, correctly identifying patients with protracted recovery, and patients with a history of headache prior to injury. However, the remaining ImPACT scores did not demonstrate validity in differentiating between the same groups.

3.3.3 Factor Analysis

Allen et al. [28] and Schatz et al. [52] completed a factor analysis of ImPACT based on the test subscales and the four composite scores, respectively. Allen et al. concluded that a five-factor structure explains 69 % of the variance in ImPACT performance [28]. These five factors (i.e., choice efficiency, verbal and visual memory, inhibitory cognitive abilities, visual processing abilities, and color match total commissions) did not correspond with the conventional composite scores used by ImPACT [28]. Schatz et al. [52] reported that 72.5 % of variance in ImPACT composite scores was explained by the two factors speed and memory. In both studies [28, 52], the distinction between verbal and visual memory was not supported. Collectively, these results suggest that the current composite scores of ImPACT are slightly dissimilar to factors identified through factor analysis, suggesting a shared variance and raising a concern about the discriminative validity of ImPACT composite scores to measure tangible cognitive constructs.

3.4 Criterion-Related Validity

The validity of ImPACT to predict the course of recovery, presence of future impairments or symptoms, the total symptom score, and perception of recovery has been examined. Lau et al. [51] demonstrated that ImPACT scores have a predictive utility in distinguishing between simple and complex concussion according to a classification that is no longer in use ($F = 2.69$, $p = 0.04$). Additionally, ImPACT demonstrated a sensitivity of 53.2 % and specificity of 75.4 % in classifying short (≤ 14 days, $n = 58$) and long recovery (> 14 days, $n = 50$) [53]. Across the selected sensitivity levels, verbal memory consistently demonstrated the highest specificity (10.5–22.8 %) in identifying patients with protracted recovery (> 14 days), whereas visual memory was least specific (1.8–3.5 %) [54]. Weibe et al. [55] demonstrated that mean initial ImPACT composite scores less than the 39th percentile correctly classified 76.2–88.1 % of participants on the presence of neurocognitive deficits at a subsequent evaluation performed 2 weeks after concussion [55]. At 1 month post-injury, however, ImPACT performance was not predictive of the total symptom scores [47] or persistent symptoms [56]. Furthermore, ImPACT was not predictive of the development of post-concussion syndrome [43], and it only accounted for 28 % of variance of perceived recovery [41]. In sum, the ImPACT does not appear to have a prognostic utility in predicting persistent concussion symptoms or the course of recovery.

3.5 Diagnostic Accuracy Validity of ImPACT

Four studies examined the diagnostic accuracy of the ImPACT test within the first 72 h after concussion against four different reference standards (Table 4) [57–60]. A reliable change in at least one of the four composite scores was used as a positive finding in two studies [57, 59], providing a sensitivity of 62.5–83 %. The sensitivity of individual composite scores ranged from 29.3 to 75.6 % [58], but inclusion of the Post-Concussion Symptom Scale raises the overall sensitivity to (79.2–81.9 %) [57, 60]. The diagnostic accuracy of the ImPACT within 72 h of concussion appears to be supported when post-concussion symptoms are evaluated simultaneously. The benefit of the ImPACT test depends on its ability to document cognitive declines after the resolution of symptoms. Although some studies investigated ImPACT cognitive performance after the resolution of symptoms [61, 62], diagnostic accuracy metrics such as sensitivity and specificity beyond the resolution of symptoms are sparse [63].

3.6 Other Considerations

3.6.1 Utility of ImPACT Test after the Resolution of Symptoms

ImPACT has also been applied following symptom resolution whereby Fazio et al. [62] demonstrated that patients who were symptomatic after concussion exhibited worse performance on all ImPACT composite scores compared to asymptomatic post-concussion and control groups. A separate investigation demonstrated 38 % of asymptomatic post-concussion college athletes continued to have a cognitive decline in at least one ImPACT composite score [61]. However, those identified as having ongoing declines from baseline may have been false-positive findings as demonstrated in other works [63, 64].

3.6.2 Invalid Baselines, “Sandbagging,” Group and Motivation Influence

A number of investigations have explored if age, sex, test-taking environment, history of concussion, or attention deficit disorder (ADD)/learning disability (LD) affect the proportion of invalid baseline scores with the overall percentages of invalid baseline scores for each ImPACT composite score ranging from 2.7 to 11.1 % [39, 65–69]. Younger athletes (10–12 years) have a significantly higher proportion of invalid baseline scores (7 %) compared to older athletes (2.7 %) [66, 69]. Furthermore, athletes with ADD/LD have a significantly higher proportion of invalid baseline scores (13.2 %) compared to athletes without

Table 4 Study characteristics and diagnostic accuracy metrics for the ImPACT test

Study	Number of participants (MF)	Age, years [mean (SD or range)]	Sport participation	Time since concussion, mean (SD)	Reference standard	Sensitivity, %	Specificity, %	Positive predictive value, %	Negative predictive value, %
Broglio et al. [57]	24 (not reported)	University athletes	Athletes	Day 1 post injury	RCI	79.2 ^a 62.5 ^b	NS	NS	NS
Gardner et al. [58]	Concussion: 87/0 Control: 46 Control: 41	Concussion: 24.20 (4.20) Control: 24.28 (3.61)	Rugby	72 h post injury, 46.44 (13.31) h	WAIS-III PSI ^c	VerbM: 65.9 VisM: 78.0 RT: 29.3 PS: 75.6	VerbM: 43.1 VisM: 39.2 RT: 84.3 PS: 56.9	VerbM: 50.0 VisM: 50.8 RT: 29.3 PS: 75.6	VerbM: 63.2 VisM: 69.0 RT: 59.7 PS: 56.9
VanKampen et al. [59]	Concussion: 122 (100/22) Control: 70 (33/37)	Concussion: 16.6 (12–27) Control: 17.3 (14–22)	Football Soccer Basketball Swimmers Track Other	Within 2 days	On-field diagnosis. Decline is defined by the RCIs	83.0	NS	83.0	70.0
Schatz et al. [60]	138 (86/52)	Concussion: 16.5 (2.3) Control: 17.3 (1.7)	Football Soccer Ice hockey Field hockey Basketball Other contact Non-contact	Within 72 h post injury	Symptom scale	81.9 ^a	89.4 ^a	89.4 ^a	81.9 ^a

VerbM verbal memory, VisM visual memory, RT reaction time, PS processing speed, NS not specified, WAIS-III PSI Wechsler Adult Intelligence Scale-Third edition, Processing Speed Index, RCI Reliable Change Index, ImPACT Immediate Post Concussion Assessment and Cognitive Testing, M male, F female, SD standard deviation

^a With symptom scores

^b Without symptom scores

^c Utilizing cutoff score of 0.818 for verbal memory, 0.71 for visual memory, <0.49 for reaction time, 39.2 for processing speed

ADD/LD (4.1 %) [66]. Similar findings were reported by Schatz et al. [67], with individuals completing the test in groups exhibiting worse performance on all ImPACT scores and a larger proportion of invalid baseline scores (8.5 %) compared to participants taking the test individually (0.3 %), ($\chi^2 = 12.1, p = 0.001$). However, a separate study evaluating group size (small vs. large) reported the proportion of participants obtaining invalid baseline scores did not vary [66]. Additionally, the proportion of invalid baseline scores did not vary based on sex or previous history of concussion [66].

Three studies examined the success rate of intentional sandbagging of ImPACT scores without being detected by the predetermined cutoff validity indicators [68, 70, 71]. Using built-in cutoffs, the proportion of participants who were successfully able to sandbag their scores ranged from 10.6 to 35 % [68, 70, 71]. Using the impulse control composite score alone, only 35 % of naïve sandbaggers and none of the coached sandbaggers were identified [71]. To counter potential sandbagging, informational sessions to the athletes prior to the completion of ImPACT have been implemented, but did not reduce the percentage of invalid baseline scores [68]. However, when the ImPACT test was re-administered to athletes with invalid baseline scores, the success rate of achieving a valid performance ranged between 62.5 and 87.5 % [68, 69].

When participants were rated based on their motivational level when taking the baseline ImPACT test, participants with high motivation exhibited better performance on all ImPACT scores compared to the participants with low motivation [72]. Additionally, participants completing the test in a supervised setting demonstrated significantly better reaction time and processing speed compared to unsupervised participants [73].

These findings suggest that practitioners should standardize testing environments and instructions to reduce distractions, particularly when the test is administered in a group setting. Moreover, practitioners should examine baseline scores for invalid performance or intentional underperforming and should consider re-administering the test when suboptimal performance is suspected.

3.6.3 Learning Disability/ADHD

Two studies [74, 75] examined the effect of LD ($n = 486$), ADHD ($n = 1144$), LD/ADHD ($n = 216$), and control ($n = 985$) on the baseline ImPACT scores. Both studies concluded that adolescent healthy participants exhibited significantly better performance on all ImPACT composite scores compared to participants with an LD, ADHD, or LD/ADHD history. Similarly, a point biserial correlational analysis showed significant relationships between LD/ADD and verbal memory ($r = -0.161, p = .04$) and

visual memory ($r = -0.236, p = .003$), but not processing speed ($r = -0.069, p = 0.39$) and reaction time ($r = 0.035, p = 0.66$) [47]. This is in contrast to Solomon et al. [48] ($n = 89$) who found that the presence of ADHD ($n = 6$) and LD ($n = 9$) had no discernible effect on the ImPACT scores. When participants were matched by sex, age, and years of education, Brooks et al. [76] demonstrated that female adolescents with attention problems performed worse on visual motor processing speed ($p < 0.001$), and male adolescents with attention problems exhibited worse reaction times ($p = 0.005$).

The findings of a possible relationship between LD/ADHD warrant a detailed assessment of its premorbid effects on baseline ImPACT scores and a cautious interpretation of ImPACT test by a neuropsychologist or trained personnel.

3.6.4 Cultural Equivalence

Two studies examined the equivalence of ImPACT scores in South African [77] and Hawaiian athletes [78]. Independent t-tests revealed significant differences in all ImPACT composite scores between US ($n = 9640$) and South African ($n = 1617$) athletes for all age groups (11–13, 14–16, and 16–18 years). However, these differences were not consistent, in which one group was always outperforming the other. Additionally, Hawaiian athletes ($n = 751$) demonstrated a comparable performance to the US mainland athletes in all baseline ImPACT scores [78]. The effect of race/ethnicity on ImPACT scores has also been examined in Kontos et al. [79], who demonstrated no differences in baseline scores between Caucasian ($n = 48$) and African American ($n = 48$) players measured at baseline, and 2 and 7 days after concussion.

When examining the interaction between linguistic abilities and test-taking language, Ott et al. [80] concluded that a language effect existed on all ImPACT composite scores ($p < .001$). More specifically, Spanish-speaking participants completing the test in Spanish ($n = 2087$) exhibited poorer performance on all composite scores than Spanish-speaking ($n = 9733$) and English-speaking ($n = 11,955$) athletes completing the test in English [80]. Furthermore, when the test was completed in English, Spanish-speaking athletes performed poorer than English-speaking athletes on verbal memory, visual motor speed, and reaction time [80]. Furthermore, Blake et al. [81] examined the effect of language of test administration in a sample of bilingual participants who took the test in English and in Spanish. Participants exhibited significantly better English language performance in verbal memory and processing speed, but not visual memory or reaction time [81].

In summary, it appears the ImPACT test demonstrates cultural equivalence. However, the language of test

administration in bilingual athletes appears to influence the results. Therefore, bilingual athletes must take the baseline and any subsequent testing in the same language.

3.6.5 Effects of Sport Season, Competition, and Exertion on ImPACT Performance

The acute effects of football competition and exertion on ImPACT were examined. When examined within 48 h of a collegiate football competition, 39.9 % of participants demonstrated a reliable decline in visual memory, 35.7 % in reaction time, 25 % in processing speed, and 14.3 % in verbal memory [82]. When ImPACT was completed within 15 min of maximum physical exertion, a significant change was observed for verbal memory only [83].

The effects of participating in collegiate rugby and adolescent football on ImPACT performance were examined by Miller et al. [84] and Munce et al. [85]. Both studies reported no detrimental effects of sport season on ImPACT performance when post-season ImPACT scores were compared to pre-season scores [84, 85]. Another investigation reported no significant differences in ImPACT scores between contact and non-contact athletes [86], but rugby players exhibited worse post-season-processing speed compared to their pre-season scores [87]. Additionally, the heading exposure (low, moderate, and high exposure) of soccer players had no reported effect on ImPACT performance in adolescents [88].

The effect of exertion on post-concussion ImPACT performance was also examined. Majerske et al. [89] demonstrated that participants engaged in activities of higher intensity in the 30 days following concussion demonstrated slower recovery of visual memory and reaction time. Furthermore, McGrath et al. [90] reported that 27.7 % of the post-concussion participants who were symptomatic and returned to baseline ImPACT scores demonstrated a reliable decline in at least one ImPACT composite score after maximal exertion.

The findings of this review indicate that physical exertion prior to the test is associated with poorer performance and must be considered before administering the test. Additionally, investigations evaluating the effects of a season of play in a contact and collision sport are mixed. It should be noted, however, that the ImPACT test was developed to measure and document the known large cognitive declines associated with concussion, not the possible small declines associated with multiple subclinical head traumas over the course of a contact-sport season.

3.6.6 Effects of Sleep

Mclure et al. [91] reported that athletes sleeping fewer than 7 h ($n = 678$) performed significantly worse on verbal

memory, visual memory, and reaction time compared to athletes with greater than 7 h of sleep ($n = 3008$) prior to the baseline ImPACT testing. Sufrinko et al. [92] reported that athletes with sleeping difficulties ($n = 34$) did not exhibit baseline differences on any ImPACT scores compared to athletes without sleeping difficulties ($n = 231$). However, athletes with pre-existing difficulty sleeping exhibited worse verbal memory 2 days after injury, and worse reaction times at 5–7 and 10–14 days after injury. No differences between groups were observed for visual memory or processing speed [92].

3.6.7 Equivalence of Alternate Forms of the Test

Resch et al. [31] demonstrated that verbal memory in form 1 is non-equivalent to verbal memory in forms 2, 3, and 4; additionally, verbal memory is non-equivalent between forms 2 and 4. Visual memory was not equivalent between forms 1 and 3. For reaction time and processing speed, forms 3 and 4 were non-equivalent [31].

3.6.8 Utility of Comparison to Normative Data versus Individual Baseline

Schatz et al. [93] examined the varying classification of post-concussion performance when compared to baseline performance and normative scores. They concluded that the above-average athletes were correctly classified as impaired when compared to their own baseline scores. However, 52–54 % of “above-average” athletes were misclassified when compared to normative values of ImPACT [93].

3.6.9 ImPACT as an Outcome Measure to Assess the Effects of Intervention

Moser et al. [94] and Camiolo Reddy et al. [95] utilized ImPACT to examine the effects of prescribed rest and amantadine, respectively, on patient recovery after concussion. In both cases, a significant improvement was observed in ImPACT scores post-intervention. In adolescents with protracted recovery after concussion who were treated by multimodal physical therapy intervention, patients exhibited improvement in processing speed only [96].

Mihalik et al. [97] utilized ImPACT to examine if using mouth guards mitigate the acute cognitive effects of concussion and demonstrated that the ImPACT performance did not differ among athletes with and without a mouth-guard at the time of injury. Additionally, wearing a Revolution helmet did not reduce the neurocognitive decrement after concussion compared to standard helmets [98].

4 Discussion

The findings of this review demonstrate that the convergent validity of the ImPACT test has been well studied and is strong. The evidence to support the discriminant validity, criterion related validity, diagnostic accuracy validity, and responsiveness, however, is sparse or inconclusive. Additionally, other factors such as testing environment, exertion, invalid baseline scores, and sandbagging appear to threaten the validity of the ImPACT results.

Studies examining the diagnostic accuracy of ImPACT were all completed within 72 h of concussion. Therefore, the diagnostic accuracy of the ImPACT test beyond this time point remains unclear. In two studies [57, 59], concussion was determined based on scores exceeding reliable change criteria on at least one ImPACT composite score. These same criteria (i.e., exceeding reliable change) identified 33–38 % of the concussed sample despite no longer reporting concussion symptoms [61, 62]. It is unclear however, if all of the athletes had continued cognitive declines or if some were falsely identified. Indeed, 22–46 % of healthy individuals tested twice demonstrated a “reliable change” in at least one ImPACT composite score without sustaining a concussion [11]. Future studies examining the diagnostic accuracy of ImPACT must therefore consider and control for healthy participants demonstrating a similar change.

Although the majority of the studies support the convergence among memory scores (verbal and visual) and among speed scores (reaction time and processing speed), the relationship was inconsistent. For instance, faster visual motor-processing speed was correlated with both slower and faster reaction times [27, 28]. When examined against each other and against other cognitive measures, the convergence of ImPACT composite scores was generally supported. However, examination of the correlation matrix (Electronic Supplementary Material, Table S2) revealed that many of the cognitive tests exhibited significant correlations with multiple ImPACT composite scores, suggesting a shared variance. Therefore, the discriminant validity of the ImPACT composite scores is not fully supported. Some investigators speculated that the ImPACT scores are useful to capture specific neurocognitive deficits after concussion and a higher number of abnormal composite scores indicate a more severe concussion [59]. However, the findings of shared variance between ImPACT composite scores refute this claim and caution against its use for differentiation between specific neurocognitive constructs [27]. Furthermore, the majority of the studies examined the convergence of ImPACT in healthy participants. Therefore, it is unclear if the same pattern of convergence exists in patients with concussion [99].

The fair to moderate relationship between ImPACT and other post-concussion evaluative measures supports the multifaceted effects of the injury. However, the lack of relationship between ImPACT and the AxonSports cognitive test was unexpected given that both tests aim to measure the effects of concussion on cognitive functioning. As concussion effects become better delineated after the acute period, the correlation between ImPACT and other measures may decrease or disappear [100].

Multiple forms of ImPACT were designed to overcome the practice effects encountered by traditional paper-and-pencil neurocognitive testing. Given that multiple forms of ImPACT are commonly administered serially to document concussion's effects on cognitive functioning, non-equivalence between alternate forms of the test may have led to false-positive or -negative findings and may partially explain the fluctuating reliability findings of ImPACT discussed elsewhere [11, 101]. The ImPACT defaults to form 1 of the test during baseline testing, whereas the practitioner can determine the test form upon subsequent post-injury administration. Therefore, clinicians must consider non-equivalence when interpreting the change between subsequent administrations. For instance, Resch et al. [31] suggested that if a clinician administers form 2 of the test after a suspected concussion, they should rely less on the findings of verbal memory given its non-equivalence with baseline testing (i.e., form 1), and they should place more emphasis on the remaining three composite scores. Furthermore, given the non-equivalence between test forms and the rate of false positives, we postulate that a decline in at least two ImPACT composite scores to be used as a diagnostic criterion for documenting a reliable cognitive decline after concussion. These suggestions are further supported by the findings of Nelson et al. [63], who found that the rate of false positives reduced to 4.8–10.8 % when a decline in two ImPACT scores was used as a minimum diagnostic criterion compared to a false-positive rate of 29.6–42.7 % when a decline in one ImPACT score was utilized.

Despite recommendations to determine sample size of correlation studies based on power analysis [10, 102, 103], and to adjust for the multiplicity of correlations [104], none of the convergence studies justified sample size, and only one study [36] adjusted for the multiplicity of comparisons. The studies reviewed here included a range of 29–323 participants, reporting 4–93 correlation analyses within each study. Therefore, many of the reviewed studies may have been susceptible to spurious or “false-positive” correlations. Given this statistical uncertainty, the width of the correlation coefficients confidence intervals is a critical value in interpreting results [103]. Despite this, none of the reviewed studies reported the correlation coefficient confidence interval.

Although the factor structure of ImPACT demonstrates the construct validity of the test based on two–five factors, these factors do not correspond with the current four composite scores, indicating a need to reconsider the current composite score structure. Additionally, the ImPACT factors may vary when examined in patients with concussion compared to healthy participants, as demonstrated by one investigation showing the factor structure of the post-concussion symptom scale differed in patients with concussion compared with healthy participants [105].

Younger individuals, individuals with LD/ADHD, and individuals completing the test in a group setting appear to have higher proportion of invalid baseline scores. Although these may be true invalid baseline scores, they may reflect a true performance in the LD/ADHD group, or a lack of understanding of the test instruction among younger individuals or individuals taking the test in a group. Despite the built-in validity indicators, 10–35 % of athletes were able to successfully sandbag their scores without being detected; therefore, cautious interpretation of baseline scores is warranted.

Despite the limitations of ImPACT highlighted in this review, it remains the most widely used concussion assessment tool. Many of the threats to ImPACT validity can be addressed with a standardized best clinical practice. For example, clinicians must standardize testing environments, testing instructions, and closely examine the baseline scores for invalid baseline scores suggesting a poor understanding of testing instructions or a sandbagged performance. The potential diagnostic and prognostic benefit of ImPACT test in patients with concussion needs to be considered on case-by-case bases and in conjunction with other recommended assessments. Additionally, appropriately trained personnel must conduct a cautious interpretation of ImPACT scores with special emphasis on the interpretation of change scores between sequential administrations of the test. As for return-to-play decisions, clinicians must utilize a multifaceted assessment approach that emphasizes the number and pattern of self-reported symptoms, motor control assessment, and an appropriate cognitive evaluation [1, 2]. When an athlete does not present with objective impairments and is asymptomatic at rest, full return to play is considered after a stepwise progression of activities emphasizing gradual increase in physical and cognitive demands while monitoring symptoms of reoccurrence or emergence.

While this review aimed to summarize the literature surrounding the validity of the ImPACT, it is limited by the quantity and the quality of the reviewed studies. Many of the validity studies were completed with healthy participants, therefore its generalizability to patients with concussion is unclear. Participants in some studies were not

matched for sports, sex, or age. A history of LD/ADHD was self-reported and is subject to recall bias.

5 Conclusion

Although ImPACT demonstrates convergent validity as a cognitive measure after concussion, evidence for its discriminant and predictive validity, its diagnostic accuracy, and its utility after the symptom resolution is sparse or inconclusive. As such, cautious implementation and interpretation of the test scores in clinical practice is warranted as several factors appear to threaten the validity of ImPACT scores. These findings highlight the role of appropriate medical professionals, with training specific to sport concussion, in the injury assessment, diagnosis, and management process.

Compliance with Ethical Standards

Funding No sources of funding were used in the preparation of this review.

Conflict of interest Bara Alsalaheen, Kayla Stockdale, Dana Pechumer, and Steven Broglio declare that they have no conflicts of interest relevant to the content of this review.

References

1. McCrory P, Meeuwisse WH, Aubry M, et al. Consensus statement on concussion in sport: the 4th International Conference on Concussion in Sport held in Zurich, November 2012. *Br J Sports Med.* 2013;47:250–8.
2. Broglio SP, Cantu RC, Gioia GA, et al. National Athletic Trainers' Association position statement: management of sport concussion. *J Athl Train.* 2014;49:245–65.
3. Covassin T, Elbin R 3rd, Stiller-Ostrowski JL. Current sport-related concussion teaching and clinical practices of sports medicine professionals. *J Athl Train.* 2009;44:400–4.
4. Randolph C, McCrea M, Barr WB. Is neuropsychological testing useful in the management of sport-related concussion? *J Athl Train.* 2005;40:139–52.
5. Mayers LB, Redick TS. Clinical utility of ImPACT assessment for postconcussion return-to-play counseling: psychometric issues. *J Clin Exp Neuropsychol.* 2012;34:235–42.
6. Maroon JC, Lovell MR, Norwig J, et al. Cerebral concussion in athletes: evaluation and neuropsychological testing. *Neurosurgery.* 2000;47:659–69.
7. Cole WR, Arrieux JP, Schwab K, et al. Test–retest reliability of four computerized neurocognitive assessment tools in an active duty military population. *Arch Clin Neuropsychol.* 2013;28:732–42.
8. Meehan WP 3rd, d'Hemecourt P, Collins CL, et al. Computerized neurocognitive testing for the management of sport-related concussions. *Pediatrics.* 2012;129:38–44.
9. Maroon JC, Field M, Lovell M, et al. The evaluation of athletes with cerebral concussion. *Clin Neurosurg.* 2002;49:319–32.
10. Portney L, Watkins M. *Foundations of clinical research: applications to practice.* Upper Saddle River: Prentice-Hall, Inc.; 2009.

11. Alsalaheen B, Stockdale K, Pechumer D, et al. Measurement error in the immediate postconcussion assessment and cognitive testing (ImPACT): systematic review. *J Head Trauma Rehabil.* 2015;. doi:10.1097/HTR.0000000000000175.
12. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med.* 2007;147:573–7.
13. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med.* 2003;138:W1–12.
14. Collins MW, Field M, Lovell MR, et al. Relationship between postconcussion headache and neuropsychological test performance in high school athletes. *Am J Sports Med.* 2003;31:168–73.
15. Lovell MR, Collins MW, Iverson GL, et al. Grade 1 or “ding” concussions in high school athletes. *Am J Sports Med.* 2004;32:47–54.
16. Colvin AC, Mullen J, Lovell MR, et al. The role of concussion history and gender in recovery from soccer-related concussion. *Am J Sports Med.* 2009;37:1699–704.
17. Bruce JM, Echemendia RJ. History of multiple self-reported concussions is not associated with reduced cognitive abilities. *Neurosurgery.* 2009;64:100–6.
18. Echlin PS, Tator CH, Cusimano MD, et al. Return to play after an initial or recurrent concussion in a prospective study of physician-observed junior ice hockey concussions: implications for return to play after a concussion. *Neurosurg Focus.* 2010;29:E5.
19. Echlin PS, Tator CH, Cusimano MD, et al. A prospective study of physician-observed concussions during junior ice hockey: implications for incidence rates. *Neurosurg Focus.* 2010;29:E4.
20. Guskiewicz KM, Marshall SW, Broglio SP, et al. No evidence of impaired neurocognitive performance in collegiate soccer players. *Am J Sports Med.* 2002;30:157–62.
21. Ponsford J, Cameron P, Fitzgerald M, et al. Predictors of post-concussive symptoms 3 months after mild traumatic brain injury. *Neuropsychology.* 2012;26:304–13.
22. Pellman EJ, Lovell MR, Viano DC, et al. Concussion in professional football: neuropsychological testing—part 6. *Neurosurgery.* 2004;55:1290–303.
23. Peterson CL, Ferrara MS, Mrazik M, et al. Evaluation of neuropsychological domain scores and postural stability following cerebral concussion in sports. *Clin J Sports Med.* 2003;13:230–7.
24. Schatz P, Sandel N. Sensitivity and specificity of the online version of ImPACT in high school and collegiate athletes. *Am J Sports Med.* 2013;41:321–6.
25. Kontos AP, Huppert TJ, Beluk NH, et al. Brain activation during neurocognitive testing using functional near-infrared spectroscopy in patients following concussion compared to healthy controls. *Brain Imaging Behav.* 2014;8:621–34.
26. Solomon GS, Kuhn A. Relationship between concussion history and neurocognitive test performance in National Football League draft picks. *Am J Sports Med.* 2014;42:934–9.
27. Maerlender A, Flashman L, Kessler A, et al. Discriminant construct validity of ImPACT: a companion study. *Clin Neuropsychol.* 2013;27:290–9.
28. Allen BJ, Gfeller JD. The immediate post-concussion assessment and cognitive testing battery and traditional neuropsychological measures: a construct and concurrent validity study. *Brain Inj.* 2011;25:179–91.
29. Schatz P, Putz BO. Cross-validation of measures used for computer-based assessment of concussion. *Appl Neuropsychol.* 2006;13:151–9.
30. Fedor A, Gunstad J. Higher BMI is associated with reduced cognitive performance in division I athletes. *Obes Facts.* 2013;6:185–92.
31. Resch JE, Macciocchi S, Ferrara MS. Preliminary evidence of equivalence of alternate forms of the ImPACT. *Clin Neuropsychol.* 2013;27:1265–80.
32. Iverson GL, Lovell MR, Collins MW. Validity of ImPACT for measuring processing speed following sports-related concussion. *J Clin Exp Neuropsychol.* 2005;27:683–9.
33. Maerlender A, Flashman L, Kessler A, et al. Examination of the construct validity of ImPACT computerized test, traditional, and experimental neuropsychological measures. *Clin Neuropsychol.* 2010;24:1309–25.
34. Tjarks BJ, Dorman JC, Valentine VD, et al. Comparison and utility of King-Devick and ImPACT(R) composite scores in adolescent concussion patients. *J Neurol Sci.* 2013;334:148–53.
35. Parker TM, Osternig LR, van Donkelaar P, et al. Recovery of cognitive and dynamic motor function following concussion. *Br J Sports Med.* 2007;41:868–73.
36. Alsalaheen B, Whitney SL, Marchetti GF, et al. Relationship between cognitive assessment and balance measures in adolescents referred for vestibular physical therapy after concussion. *Clin J Sports Med.* 2015;26:46–52.
37. Vernau BT, Grady MF, Goodman A, et al. Oculomotor and neurocognitive assessment of youth ice hockey players: baseline associations and observations after concussion. *Dev Neuropsychol.* 2015;40:7–11.
38. Sasaki T, Pasternak O, Mayinger M, et al. Hockey Concussion Education Project, Part 3. White matter microstructure in ice hockey players with a history of concussion: a diffusion tensor imaging study. *J Neurosurg.* 2014;120:882–90.
39. Broglio SP, Sosnoff JJ, Ferrara MS. The relationship of athlete-reported concussion symptoms and objective measures of neurocognitive function and postural control. *Clin J Sports Med.* 2009;19:377–82.
40. Kontos AP, Covassin T, Elbin RJ, et al. Depression and neurocognitive performance after concussion among male and female high school and collegiate athletes. *Arch Phys Med Rehabil.* 2012;93:1751–6.
41. Sandel NK, Lovell MR, Kegel NE, et al. The relationship of symptoms and neurocognitive performance to perceived recovery from sports-related concussion among adolescent athletes. *Appl Neuropsychol Child.* 2013;2:64–9.
42. Covassin T, Elbin RJ 3rd, Larson E, et al. Sex and age differences in depression and baseline sport-related concussion neurocognitive performance and symptoms. *Clin J Sports Med.* 2012;22:98–104.
43. Barlow M, Schlabach D, Peiffer J, et al. Differences in change scores and the predictive validity of three commonly used measures following concussion in the middle school and high school aged population. *Int J Sports Phys Ther.* 2011;6:150.
44. Ponsford J, Cameron P, Fitzgerald M, et al. Long-term outcomes after uncomplicated mild traumatic brain injury: a comparison with trauma controls. *J Neurotrauma.* 2011;28:937–46.
45. Rieger BP, Lewandowski LJ, Callahan JM, et al. A prospective study of symptoms and neurocognitive outcomes in youth with concussion vs orthopaedic injuries. *Brain Inj.* 2013;27:169–78.
46. Shores EA, Lammel A, Hullick C, et al. The diagnostic accuracy of the Revised Westmead PTA Scale as an adjunct to the Glasgow Coma Scale in the early identification of cognitive impairment in patients with mild traumatic brain injury. *J Neurol Neurosurg Psychiatry.* 2008;79:1100–6.
47. Solomon GS, Haase RF. Biopsychosocial characteristics and neurocognitive test performance in National Football League players: an initial assessment. *Arch Clin Neuropsychol.* 2008;23:563–77.
48. Solomon GS, Haase RF, Kuhn A. The relationship among neurocognitive performances and biopsychosocial characteristics of

- elite National Football League draft picks: an exploratory investigation. *Arch Clin Neuropsychol*. 2013;28:9–20.
49. Lovell MR, Solomon GS. Neurocognitive test performance and symptom reporting in cheerleaders with concussions. *J Pediatr*. 2013;163(1192–1195):e1191.
 50. Iverson G. Predicting slow recovery from sport-related concussion: the new simple-complex distinction. *Clin J Sports Med*. 2007;17:31–7.
 51. Lau B, Lovell MR, Collins MW, et al. Neurocognitive and symptom predictors of recovery in high school athletes. *Clin J Sports Med*. 2009;19:216–21.
 52. Schatz P, Maerlender A. A two-factor theory for concussion assessment using ImPACT: memory and speed. *Arch Clin Neuropsychol*. 2013;28:791–7.
 53. Lau BC, Collins MW, Lovell MR. Sensitivity and specificity of subacute computerized neurocognitive testing and symptom evaluation in predicting outcomes after sports-related concussion. *Am J Sports Med*. 2011;39:1209–16.
 54. Lau BC, Collins MW, Lovell. Cutoff scores in neurocognitive testing and symptom clusters that predict protracted recovery from concussions in high school athletes. *Neurosurgery*. 2012;70:371–9 (**discussion 379**).
 55. Wiebe DJ, Collins MW, Nance ML. Identification and validation of prognostic criteria for persistence of mild traumatic brain injury-related impairment in the pediatric patient. *Pediatr Emerg Care*. 2012;28:498–502.
 56. Meehan WP 3rd, Mannix RC, Stracciolini A, et al. Symptom severity predicts prolonged recovery after sport-related concussion, but age and amnesia do not. *J Pediatr*. 2013;163:721–5.
 57. Broglio SP, Macciocchi SN, Ferrara MS. Sensitivity of the concussion assessment battery. *Neurosurgery*. 2007;60:1050–7.
 58. Gardner A, Shores EA, Batchelor J, et al. Diagnostic efficiency of ImPACT and CogSport in concussed rugby union players who have not undergone baseline neurocognitive testing. *Appl Neuropsychol Adult*. 2012;19:90–7.
 59. Van Kampen DA, Lovell MR, Pardini JE, et al. The “value added” of neurocognitive testing after sports-related concussion. *Am J Sports Med*. 2006;34:1630–5.
 60. Schatz P, Pardini JE, Lovell MR, et al. Sensitivity and specificity of the ImPACT test battery for concussion in athletes. *Arch Clin Neuropsychol*. 2006;21:91–9.
 61. Broglio SP, Macciocchi SN, Ferrara MS. Neurocognitive performance of concussed athletes when symptom free. *J Athl Train*. 2007;42:504–8.
 62. Fazio VC, Lovell MR, Pardini JE, et al. The relation between post concussion symptoms and neurocognitive performance in concussed athletes. *Neurorehabilitation*. 2007;22:207–16.
 63. Nelson LD, LaRoche AA, Pfaller AY, et al. Prospective, head-to-head study of three computerized neurocognitive assessment tools (CNTs): reliability and validity for the assessment of sport-related concussion. *J Int Neuropsychol Soc*. 2016;22:24–37.
 64. Broglio SP, Ferrara MS, Macciocchi SN, et al. Test-retest reliability of computerized concussion assessment programs. *J Athl Train*. 2007;42:509–14.
 65. Schatz P, Moser RS, Solomon GS, et al. Prevalence of invalid computerized baseline neurocognitive test results in high school and collegiate athletes. *J Athl Train*. 2012;47:289–96.
 66. Lichtenstein JD, Moser RS, Schatz P. Age and test setting affect the prevalence of invalid baseline scores on neurocognitive tests. *Am J Sports Med*. 2014;42:479–84.
 67. Moser RS, Schatz P, Neidzowski K, et al. Group versus individual administration affects baseline neurocognitive test performance. *Am J Sports Med*. 2011;39:2325–30.
 68. Szabo AJ, Alosco ML, Fedor A, et al. Invalid performance and the ImPACT in national collegiate athletic association division I football players. *J Athl Train*. 2013;48:851–5.
 69. Schatz P, Kelley T, Ott SD, et al. Utility of repeated assessment after invalid baseline neurocognitive test performance. *J Athl Train*. 2014;49:659–64.
 70. Erdal K. Neuropsychological testing for sports-related concussion: how athletes can sandbag their baseline testing without detection. *Arch Clin Neuropsychol*. 2012;27:473–9.
 71. Schatz P, Glatts C. “Sandbagging” baseline test performance on ImPACT, without detection, is more difficult than it appears. *Arch Clin Neuropsychol*. 2013;28:236–44.
 72. Rabinowitz AR, Merritt VC, Arnett PA. The return-to-play incentive and the effect of motivation on neuropsychological test-performance: implications for baseline concussion testing. *Dev Neuropsychol*. 2015;40:29–33.
 73. Kuhn AW, Solomon GS. Supervision and computerized neurocognitive baseline test performance in high school athletes: an initial investigation. *J Athl Train*. 2014;49:800–5.
 74. Elbin RJ, Kontos AP, Kegel N, et al. Individual and combined effects of LD and ADHD on computerized neurocognitive concussion test performance: evidence for separate norms. *Arch Clin Neuropsychol*. 2013;28:476–84.
 75. Zuckerman SL, Lee YM, Odom MJ, et al. Baseline neurocognitive scores in athletes with attention deficit-spectrum disorders and/or learning disability. *J Neurosurg Pediatr*. 2013;12:103–9.
 76. Brooks BL, Iverson GL, Atkins JE, et al. Sex differences and self-reported attention problems during baseline concussion testing. *Appl Neuropsychol Child*. 2015;. doi:[10.1080/21622965.2014.1003066](https://doi.org/10.1080/21622965.2014.1003066):1–8.
 77. Shuttleworth-Edwards AB, Whitefield-Alexander VJ, Radloff SE, et al. Computerized neuropsychological profiles of South African versus US athletes: a basis for commentary on cross-cultural norming issues in the sports concussion arena. *Phys Sportsmed*. 2009;37:52–9.
 78. Tsushima WT, Oshiro R, Zimbra D. Neuropsychological test performance of Hawai'i high school athletes: Hawai'i ImPACT normative data. *Hawaii Med J*. 2008;67:93–5.
 79. Kontos AP, Elbin RJ 3rd, Covassin T, et al. Exploring differences in computerized neurocognitive concussion testing between African American and White athletes. *Arch Clin Neuropsychol*. 2010;25:734–44.
 80. Ott S, Schatz P, Solomon G, et al. Neurocognitive performance and symptom profiles of Spanish-speaking Hispanic athletes on the ImPACT test. *Arch Clin Neuropsychol*. 2014;29:152–63.
 81. Lehman Blake M, Ott S, Villanyi E, et al. Influence of language of administration on ImPACT performance by bilingual Spanish-English college students. *Arch Clin Neuropsychol*. 2015;30:302–9.
 82. Mulligan I, Boland M, Payette J. Prevalence of neurocognitive and balance deficits in collegiate aged football players without clinically diagnosed concussion. *J Orthop Sports Phys Ther*. 2012;42:625–32.
 83. Covassin T, Weiss L, Powell J, et al. Effects of a maximal exercise test on neurocognitive function. *Br J Sports Med*. 2007;41:370–4.
 84. Miller JR, Adamson GJ, Pink MM, et al. Comparison of pre-season, midseason, and postseason neurocognitive scores in uninjured collegiate football players. *Am J Sports Med*. 2007;35:1284–8.
 85. Munce TA, Dorman JC, Odney TO, et al. Effects of youth football on selected clinical measures of neurologic function: a pilot study. *J Child Neurol*. 2014;29:1601–7.
 86. McAllister TW, Flashman LA, Maerlender A, et al. Cognitive effects of one season of head impacts in a cohort of collegiate contact sport athletes. *Neurology*. 2012;78:1777–84.
 87. Shuttleworth-Edwards AB, Smith I, Radloff SE. Neurocognitive vulnerability amongst university rugby players versus noncontact sport controls. *J Clin Exp Neuropsychol*. 2008;30:870–84.

88. Kontos AP, Dolese A, Elbin RJ, et al. Relationship of soccer heading to computerized neurocognitive performance and symptoms among female and male youth soccer players. *Brain Inj.* 2011;25:1234–41.
89. Majerske CW, Mihalik JP, Ren D, et al. Concussion in sports: postconcussive activity levels, symptoms, and neurocognitive performance. *J Athl Train.* 2008;43:265–74.
90. McGrath N, Dinn WM, Collins MW, et al. Post-exertion neurocognitive test failure among student-athletes following concussion. *Brain Inj.* 2013;27:103–13.
91. McClure DJ, Zuckerman SL, Kutscher SJ, et al. Baseline neurocognitive testing in sports-related concussions: the importance of a prior night's sleep. *Am J Sports Med.* 2014;42:472–8.
92. Sufrinko A, Pearce K, Elbin RJ, et al. The effect of preinjury sleep difficulties on neurocognitive impairment and symptoms after sport-related concussion. *Am J Sports Med.* 2015;43:830–8.
93. Schatz P, Robertshaw S. Comparing post-concussive neurocognitive test data to normative data presents risks for underclassifying “above average” athletes. *Arch Clin Neuropsychol.* 2014;29:625–32.
94. Moser RS, Glatts C, Schatz P. Efficacy of immediate and delayed cognitive and physical rest for treatment of sports-related concussion. *J Pediatr.* 2012;161:922–6.
95. Reddy CC, Collins M, Lovell M, et al. Efficacy of amantadine treatment on symptoms and neurocognitive performance among adolescents following sports-related concussion. *J Head Trauma Rehabil.* 2013;28:260–5.
96. Gagnon I, Galli C, Friedman D, et al. Active rehabilitation for children who are slow to recover following sport-related concussion. *Brain Inj.* 2009;23:956–64.
97. Mihalik JP, McCaffrey MA, Rivera EM, et al. Effectiveness of mouthguards in reducing neurocognitive deficits following sports-related cerebral concussion. *Dent Traumatol.* 2007;23:14–20.
98. Collins M, Lovell MR, Iverson GL, et al. Examining concussion rates and return to play in high school football players wearing newer helmet technology: a three-year prospective cohort study. *Neurosurgery.* 2006;58:275–86.
99. Delis DC, Jacobson M, Bondi MW, et al. The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: lessons from memory assessment. *J Int Neuropsychol Soc.* 2003;9:936–46.
100. Collins MW, Kontos AP, Reynolds E, et al. A comprehensive, targeted approach to the clinical care of athletes following sport-related concussion. *Knee Surg Sports Traumatol Arthrosc.* 2014;22:235–46.
101. Resch J, Driscoll A, McCaffrey N, et al. ImPact test-retest reliability: reliably unreliable? *J Athl Train.* 2013;48:506–11.
102. Suresh K, Chandrashekar S. Sample size estimation and power analysis for clinical research studies. *J Hum Reprod Sci.* 2012;5:7–13.
103. Moinester M, Gottfried R. Sample size estimation for correlations with pre-specified confidence interval. *Quant Methods Psychol.* 2014;10:124–30.
104. Curtin F, Schulz P. Multiple correlations and Bonferroni's correction. *Biol Psychiatry.* 1998;44:775–7.
105. Kontos AP, Elbin RJ, Schatz P, et al. A revised factor structure for the post-concussion symptom scale: baseline and postcon-cussion factors. *Am J Sports Med.* 2012;40:2375–84.