Post-Concussion Acute Signs and Reliable Cognitive Decline in a Finnish Youth Ice Hockey Sample

Peltonen, Kati

2021-08

Peltonen , K , Vartiainen , M , Koskinen , S , Pertab , J , Laitala , T & Hokkanen , L 2021 , ' Post-Concussion Acute Signs and Reliable Cognitive Decline in a Finnish Youth Ice Hockey Sample ' , Archives of Clinical Neuropsychology , vol. 36 , no. 5 , pp. 757-766 . https://doi.org/10.1093/arclin/acaa10

http://hdl.handle.net/10138/338611
https://doi.org/10.1093/arclin/acaa108

unspecified
acceptedVersion

Post-Concussion Acute Signs and Reliable Cognitive Decline in a Finnish Youth Ice Hockey Sample

Peltonen, Kati; Department of Psychology and Logopedics, University of Helsinki, P.O. Box 21, 00014 Helsinki, Finland. E-mail: kati.peltonen@helsinki.fi, phone: +3582941911

Vartiainen, Matti; Department of Psychology and Logopedics, University of Helsinki, P.O. Box 21, 00014 Helsinki, Finland

Koskinen, Sanna; Department of Psychology and Logopedics, University of Helsinki, P.O. Box 21, 00014 Helsinki, Finland

Pertab, Jon; Neurosciences Institute, Intermountain Medical Center, 5171 Cottonwood St #810, Murray, UT 84107, USA

Laitala, Tiina; University of Turku, Institute of Biomedicine, FI-20014 University of Turku

Hokkanen, Laura; Department of Psychology and Logopedics, University of Helsinki, P.O. Box 21, 00014 Helsinki, Finland

Running title: Reliable change approach to post-concussion cognitive decline

Acknowledgements

COI: None.

Abstract

Objective In sports concussion research, the importance of an individualized approach incorporating neuropsychological assessment data has been emphasized. This study examined the impact of acute signs of concussion on post-injury cognitive functioning using reliable change methodology in a sample of Finnish, elite-level, youth ice hockey players.

Methods From a sample of 1 823 players (all male, 14 - 20 years-old) who completed pre-season baseline testing with the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT®) battery, two sub-groups were identified. A) 312 uninjured athletes, who completed baseline testing twice - one year apart. The scores were contrasted to calculate reliable change indices (*RCI*). B) from a subsample of 570 athletes participating in an intensive follow-up arm of the project, the analysis included 32 concussed athletes. The *RCIs* were determined for the five ImPACT composite scores and used in identifying athletes with declined performance three days post-injury.

Results Test-retest reliability ranged from .39 to .71. Athletes who had experienced acute loss of consciousness, amnesia or postural instability had increased odds for declines in two or more areas assessed by ImPACT (*Odds Ratio* = 7.67–8.00, $p < .05$). In contrast, acute disorientation or vacant look did not lead to cognitive change that met the reliable change threshold.

Conclusions The reliability coefficients and *RCIs* differed from those published earlier emphasizing the importance of national reference values. The presence of acute loss of consciousness, amnesia or postural instability may indicate more severe injury and predict the need for more intensive cognitive follow-up.

Keywords

Head injury, Traumatic brain injury, Practice effects, Reliable change, Assessment, Cross-cultural

Introduction

Consensus statements discussing the evaluation and management of post-concussion symptoms in sports populations highlight the importance of an individualized approach that incorporates neuropsychological assessment data (Dougan, Horswill & Geffen, 2013; McCrory et al., 2013). Due to variability in premorbid cognitive functioning between individual athletes, many organizations have adopted a preseason baseline cognitive testing protocol. Pre-season results can provide a more accurate estimate of postinjury cognitive change than would be gleaned from reliance on normative data alone (Van Kampen, Lovell, Pardini, Collins & Fu, 2006). Baseline testing is considered a valuable option for subpopulations with prior conditions affecting athletes' cognitive functioning, such as learning disability or ADHD (Littleton et al. 2015; Peltonen et al., 2018), and for athletes of different primary languages and educational backgrounds (Jones et al., 2014; Vartiainen et al., 2019). For these subpopulations, normative data may not be representative. Based on the most recent Consensus statement on concussion in sport, the baseline assessment is not recommended as a mandatory aspect of every assessment, but it may be useful for interpreting post-injury scores (McCrory et al., 2017).

The Immediate Postconcussion Assessment and Cognitive Testing (ImPACT®) battery is one of the most widely used cognitive assessment methods in Sports (Lovell, 2006). There are many studies supporting its reliability, sensitivity, and validity in the concussed population (Elbin, Schatz, & Covassin, 2011; Iverson, 2005; Maerlander et al., 2010; Schatz & Putz, 2006; Schatz & Sandel, 2013). However, only modest test–retest reliability and low sensitivity beyond a brief post-injury window have been reported (Broglio, Ferrara, Macciocchi, Baumgartner & Elliott, 2007; Nelson et al., 2016b) suggesting that the utility of ImPACT may be limited by its psychometric properties. In young athletes with rapidly developing brain functions the age-related differences in cognitive performance are still unknown (Harmon et al. 2019). There are also other possible modifiers that should be considered when interpreting the post-injury test results. For example, learning disability, attention deficit-hyperactivity disorder (ADHD) and concussion history can potentially skew post-injury comparison (Iverson et al. 2012, Nelson et al. 2016a, Peltonen et al. 2018).

Cultural, ethnic and linguistic influence on traditional paper pencil neuropsychological testing is widely recognized (Agranovich & Puente, 2007; Bure-Reyes et al., 2013; Cherner et al., 2007; Norman et al., 2011; Shuttleworth-Edwards et al., 2004). The ImPACT is developed in United States (Lovell 2004; Iverson, Lovell & Collins 2002) and there are large body of university-based research concerning the validity and reliability of the English language version of ImPACT (Maerlander et al. 2006; Schatz, Pardini, Lovell, Collins & Podell 2006; Schatz & Sandel 2013). In a cross-cultural study concerning English version of ImPACT, no difference was found in neurocognitive performance between athletes from South Africa and from Unites States suggesting the cultural equivalence of performance (Shuttleworth-Edwards et al., 2009). Although there are multiple language versions of ImPACT, there are limited studies that assess the adequacy of these translations or the cultural equivalence as well as validity of each test version. Significant variability has been observed in test-retest reliabilities among different linguistic subsamples in National Hockey League athletes, such as French-, Czech-, and Swedish speaking subsamples (Bruce, Echemendia, Meeuwisse, Comper, & Sisco, 2014). A subsequent study reported normative data for different language versions of ImPACT in a large sample (n= 4 780) of English, French, Swedish, Russian, Czech, Finnish, and German players (Echemendia et al., 2020). The authors concluded that symptom reporting, the number of concussions sustained, and neuropsychological test results varied significantly based on a player's language of origin (Echemendia et al., 2020). In a recent Finnish study of professional ice hockey players, performance on ImPACT was compared to other language subgroups (Vartiainen et al., 2019). The results suggested that English and Czech language samples differed from Finnish sample on the Visual Motor Speed and Reaction Time composites – highlighting the importance of using language based normative values (Vartiainen et al., 2019). While Visual Motor Speed and Reaction Time are measured in milliseconds, there are multiple possible factors affecting the performance. Potential variables explaining the cross-cultural differences include factors such as computer literacy, test-taking familiarity and attitudes towards working at speed at the expense of neatness (Shuttleworth-Edwards et al., 2009). Other influential factors might be the accuracy of translated test-taking instructions and the educational level of athletes. However, we assume there are no difference in educational level between the United States and Finland, while both countries have relatively advanced education, i.e. in both countries approximately 40 % of population has attained tertiary education (OECD, 2013).

Clinicians and researchers interested in individual cognitive differences between pre- and post-injury cognitive functioning rely on statistics to interpret the relevance of observed changes. The reliable change index (RCI), is a metric developed by Jacobson and Truax (1991) designed to assess whether the change between repeated measurements is statistically and clinically meaningful. The RCI provides a confidence interval used to calculate whether a change in scores is statistically significant, considering the measurement error surrounding test-retest difference scores. The reliable change methodology is especially useful in clinical small-sample studies (Zahra & Hedge, 2010). Many commonly used statistical methods in human research, like *t*-tests and analysis of variance (ANOVAs), are based on testing difference compared to the mean of the group (Zahra & Hedge, 2010). However, the reliable change index is used to estimate whether the observed change in an individual athlete is greater than variance that can be attributed to the limited reliability inherent in the measure itself.

In sport settings, athletes are usually assessed serially at the baseline and after injury. In serial assessment, practice effects can potentially skew interpretation of the results. The reliable change index formula can take into account the practice effect, when a control group is used for estimating the mean practice effect (Chelune, Naugle, Lüders, Sedlak & Awad, 1993). Iverson, Lovell and Collins (2003) used reliable change methodology (Chelune et al., 1993) in a study designed to assess the reliability and other psychometric properties of ImPACT version 2.0 in a sample of athletes based in the USA. Their healthy sample included 56 male and female athletes from different sports assessed twice in a seven day period. They applied the reliable change parameters gleaned from the 56 athletes to a second sample of 41 concussed amateur athletes who were assessed preseason and within 72 hours post-injury to identify athletes with cognitive deficits post-injury. Their study supported the reliability and validity of ImPACT in a concussed population (Iverson et al., 2003).

Despite decades of research, widely accepted and clinically useful metrics for grading the severity of concussion have remained elusive to the scientific community. There are several proposed grading systems available and most of these incorporate loss of consciousness and posttraumatic amnesia (both retrograde and anterograde) as important variables to consider in diagnosing concussion and estimating concussion severity (Cantu, 2001). In a meta-analysis, these traditional markers of concussion have been associated with neurocognitive deficits following sports-related concussion (Dougan et al., 2014). Recently the predictive value of LOC has been questioned on several grounds (Collins et al., 2003, Iverson et al., 2017, Lovell, Iverson, Collins, McKeag & Maroon, 1999). In addition to these traditional markers of concussion, other important acute sign of concussion is postural instability. Postural instability has been found to predict concussion diagnosis in a study using video analysis for verification (Echemendia et al., 2017). Researchers have focused on studying whether several acute concussion signs and symptoms predict the rate of recovery from concussion (Lau, Kontos, Collins, Mucha & Lovell, 2011; Lau, Lovell, Collins & Pardini, 2009; Iverson, 2007; Peltonen et al., 2020). Prognostic information such as this can be a useful component informing clinical decisions regarding return to play. Commonly utilized initial signs of concussion, which are usually observable events or behaviors, include loss of consciousness, amnesia, disorientation, postural instability and vacant look (McCrory et al., 2013).

The purpose of present study was to find the threshold for meaningful change in postinjury cognitive functioning in a Finnish sample. The first aim of the study was to estimate test-retest reliability and reliable change parameters in a healthy sample of ice hockey players who completed separate baseline assessments one year apart. The second aim was to validate the derived reliable change parameters by applying them to a pilot-sample of concussed athletes who underwent preseason baseline testing and were re-evaluated at three days post-injury. The third aim was to explore the association between cognitive change post-injury and acute signs of concussion – in this case - loss of consciousness, amnesia, disorientation, postural instability and vacant look. To our knowledge, this study was the first to explore one-year test-retest reliability and the reliable change parameters of a Finnish language version of ImPACT with a nationwide sample.

Methods

Participants

The data were collected as part of the xxxx “Heads in the Game” project (Peltonen et al., 2018; Peltonen et al., 2020). The subjects were recruited from players in elite-level junior-divisions. All seventeen ice hockey clubs with elite-level teams in Finland took part in the research project during seasons 2015-2016 and 2016-2017. Altogether 1 823 athletes participated in baseline assessments and a structured program to identify concussion and monitor recovery during the season. The baseline assessment included pre-season cognitive functioning (ImPACT and Standardized Assessment of Concussion, SAC), balance (Tandem walking and Balance Error Scoring System testing, BESS), and self-reported symptom (Post-Concussion Symptom Score, PCSS) assessments (Guskiewicz et al., 2013; ImPACT Applications Inc.).

The sample used to calculate reliable change index metrics was drawn from a subsample of 458 athletes who participated in two baseline assessments one year apart. ImPACT has not been in use in Finnish youth ice hockey before, so the athletes took part in the baseline evaluation for the first time. The baseline tests were administered in small groups (maximum 6 persons) across all clubs by the research team. Only elite-level teams and certain age groups participated in the study, of which the oldest and youngest age groups participated only once in the baseline evaluation. The participation was voluntary. Athletes were excluded from the sample if they had (1) a concussion during the season (n = 19), (2) a history of learning disability or ADD/ADHD (n = 32), (3) a history of concussion (n = 90), (4) a primary language that was not Finnish (n = 3), or (5) an invalid baseline testing (n = 2) suggested by the ImPACT validity criteria (Impulse Composite Score >30). After applying the exclusionary criteria, 312 athletes were included and comprised the healthy control sample. The athletes were all male, 13 - 20 years of age (M = 16.00, SD = 1.52) at the time of the first baseline assessment.

Four ice hockey clubs in Southern Finland were selected for a more intensive follow-up in the project. These clubs included a total of 570 adolescent and young adult athletes (12 – 21 years of age). The medical personnel of the clubs were instructed to contact the project team in the event of a suspected concussion. A total of 55 concussions in 52 athletes were reported during the two seasons 2015 - 2017 and they form the concussion sample for the second part of the study. Total of 20 athletes were excluded from the concussion sample because of pre-existing disorder (learning disabilities, ADHD or migraine) or missing post-injury data. Data from the remaining 32 subjects were included. The athletes were males, 14 - 20 years of age (M = 16.97, SD = 1.73) and they had 0 - 4 previous concussions (Mdn = 0).

Identification of concussion and the acute signs

A concussion is defined as a complex pathophysiological process affecting the brain, induced by biomechanical forces (McCrory et al., 2013). The medical personnel employed by each team (e.g. physician, physiotherapist, or first-aid personnel) identified and evaluated all suspected concussions acutely using the Sport Concussion Assessment Tool, 3rd Edition (SCAT3; Guskiewicz et al., 2013) at the time of injury. A concussion was identified after receiving a blow to head or body based upon the acute presentation of 1 or more somatic-, cognitive- and/or emotional symptoms (assessed with SCAT), physical signs (eg. loss of consciousness, amnesia or disorientation as identified by on-ice examination) or behavioral change (eg. irritability) (McCrory et al., 2013). The Sport Concussion Assessment Tool (SCAT) is commonly used to gather information about acute effects of concussion (Echemendia et al., 2017; Guskiewicz et al., 2013). The teams' medical personnel were trained by the research team to identify the acute signs of concussion as well as to identify and evaluate concussion. The teams' medical personnel carefully documented information of acute signs of concussion: loss of consciousness, amnesia, disorientation, postural instability and vacant look. Acute signs of concussion were evaluated through observation, direct questioning and by the athlete's self-report at time of injury.

Cognitive assessment

Baseline cognitive evaluation was completed before the 2015-2016, and 2016-2017 playing seasons. Concussed athletes were invited to participate in post-injury follow-up at three days (M = 3.19, Ra = 2 - 4) post-injury where the measures used in the baseline assessment were re-administered. Cognitive functioning was assessed with the ImPACT computerized neurocognitive test battery (Online version; ImPACT Applications Inc.). The battery consists of six individual test modules measuring attention, memory, reaction time, processing speed, learning, and executive functioning. ImPACT provides composite scores for Verbal and Visual Memory, Visual Motor Speed, Reaction Time, and Impulse Control, and also includes a Total Symptoms Score describing the severity of subjective symptoms. The athlete's age, number of previous concussions, medications, and medical history was collected as the background information. Athletes completed ImPACT in their native language - Finnish.

Statistical analysis

All analyses were performed using the IBM SPSS Statistics software version 25.0 (IBM Corp., Armonk, NY, USA). Two types of analysis, intraclass correlation coefficients (*ICCs*) and Pearson r correlations, were employed as measures of test reliability. A single measure two-way random effects analysis of variance was calculated to assess the consistency of testing at baseline 1 and one years later at baseline 2. To evaluate differences in scores between the two baseline assessments paired t-tests were conducted. The α-level for all analyses with the Bonferroni correction was set at $p < .01$. The reliable change methodology was used, because it allows the clinician to estimate measurement error concerning test-retest difference score.

Reliable change indices were calculated to assess whether a change between the repeated baseline assessments is statistically meaningful using either an 80 % or 90 % confidence interval. There are alternative means to calculate reliable change indices. For example, it is possible to use an estimated *Sdiff* (=√(2*[SEM_1^2]) by multiplying the squared baseline standard error of measurement by two (Jacobson & Truax, 1991). In this study we used the reliable change formula that calculates the *Sdiff* ($=\sqrt{[SEM_1^2]+[SEM_2^2]}$) from the baseline and retest standard error of measurement which takes into account the practice effect (Chelune et al., 1993).

To analyze the effect of the signs in the concussed sample, the 80 % confidence interval using the previously outlined calculation was used to calculate the reliable change. It was selected to minimize the possibility of type two errors which could potentially lead to premature return to play decisions. This was especially important considering the small concussed sample size. Several separate Chi-squared tests were conducted to determine if athletes displaying or reporting acute signs of concussion (loss of consciousness, amnesia, disorientation, postural instability or vacant look) differed in rate of cognitive decline from the athletes without the specific acute sign.

Ethics

The study was approved by the Ethical Committee of the Helsinki Uusimaa Hospital District. Each participant or, if the athlete was younger than 16 years of age, a parent/guardian have signed an informed consent. The study was conducted according to the Declaration of Helsinki.

Results

The descriptive statistics for healthy control sample tested twice are presented in Table 1. The *ICC* reliability indices ranged from 0.58 to 0.78 for the composite scores, and 0.43 for the Postconcussion Scale. Paired *t*-tests found significant differences ($p \leq .01$) in all composite scores as well as in Postconcussion Scale between the baseline 1 and baseline 2 assessments, with better performances and less reported symptoms in baseline 2. Visual Motor Speed yielded a large effect size, while effect sizes were moderate for Visual Memory and small for Verbal Memory, Reaction Time and Postconcussion Scale.

[INSERT TABLE 1 HERE]

Pearson *r* correlations between baseline assessments ranged from .39 to .71 (Table 2). Intraclass correlation coefficients reflected higher reliability than Pearson *r*, across all measures (Table 1 and 2). Reliable change indices were calculated for all composite scores and the Postconcussion Scale. *RCIs* at 80 % and 90 % confidence intervals, the standard error of measurements (*SEMs*) and standard error of differences (S_{diff}) for each of the ImPACT scales are presented in Table 2. The variation between the Jacobson & Truax (1991) method (lower line) and Chelune et al.'s (1993) method (top line) was marginal.

[INSERT TABLE 2 HERE]

The reliable change difference scores associated with the confidence intervals (80 % and 90 %) were applied to the original data (*n* = 312) and the rate of change is presented in Table 3. Reliable change confidence intervals showed that 17 % – 25 % of the composite scores and Postconcussion Scale scores fell outside the 80% confidence interval, while 5 % – 9 % of the scores were outside the 90 % confidence interval.

[INSERT TABLE 3 HERE]

[INSERT TABLE 4 HERE]

The number of scores that reliably changed for each subject was then computed in the concussed sample. The percentages of concussed sample that would be classified as reliably declined or improved based on the 80% and 90% confidence intervals are shown in Table 4. The 80 % confidence interval was selected to estimate the reliable decline as follows: Verbal Memory ≥ 14 points, Visual Memory ≥ 15 points, Visual Motor Speed ≥ 6 points, Reaction Time ≥ .10 s, and Postconcussion Scale ≥ 9 points. The percentages of athletes with concussions showing declines across the five composite scores were: no declines = 56.25 %, one decline = 25.00 %, two declines = 15.63 %, three declines = 0 %, four declines = 3.13 %, and five declines = 0 %. Athletes with concussions were much more likely to have two or more declines across the five composites than the subjects in healthy sample [$X^2(1) = 19.48$, $p = .00001$; *Odds Ratio* = 8.77, 95 % *CI* = 2.83–27.19].

Comparison between the groups with and without acute sign of concussion (loss of consciousness, amnesia, disorientation, postural instability or vacant look) revealed that the athletes with loss of consciousness, amnesia or postural instability were much more likely to have two or more declines across five composites than subjects without this sign (Table 5). In contrast, the group with acute disorientation and the group with acute vacant look did not differ in cognitive declines from the groups without these signs (Table 5).

[INSERT TABLE 5 HERE]

Discussion

This is the first study to report 1-year test–retest reliability for Finnish language version of ImPACT. Test-retest reliability (*r*) derived from healthy control sample of Finnish athletes ranged from .39 to .71. Reliable change confidence intervals were calculated, and reliable change indices (*RCI*) in the ImPACT composites were applied to 32 concussed athletes to assess their cognitive decline. Among the injured athletes, the acute signs of concussion predicted the presence of neurocognitive decline. Athletes having experienced acute loss of consciousness, amnesia, or postural instability were approximately eight times more likely to have two or more declines across the five ImPACT composites three days post-injury.

In a multilingual sample of professional ice hockey players Bruce et al. (2014) reported the following 1-year test-retest correlation coefficients (*r*) for the composite scores of ImPACT: Verbal Memory .46, Visual Memory .52, Reaction Time .64, Visual Motor Speed .76, and Total Symptom .49. The results of present study are in line with those, which may be due to same one-year time interval and homogenous samples of male ice hockey players. Compared to the reliability indices gleaned from previous research of Iverson et al. (2003) the correlation coefficients in the present study were considerably lower. In their study concerning ImPACT test-retest reliability, with 56 healthy students and an average retest interval of six days, the following correlation coefficients (*r*) for the composite scores were obtained: Verbal Memory .70, Visual Memory .67, Reaction Time .79, Visual Motor Speed .86, and Total Symptom .65 (Iverson et al., 2003). In the current study, the Postconcussion scale had the lowest test-retest reliability (.39) and Visual Motor Speed the highest reliability (.71) of all the composite scores. For clinical purposes, the one-year interval may be more useful than the one-week interval used by Iverson et al. (2003), as the vast majority of concussions occur months to years after baseline testing rather than within days. Also, the language version of ImPACT (English vs Finnish) may have an effect. Variable reliability values have been published depending on the time interval between the tests and subjects' linguistic background (Bruce et al., 2014; Tsushima et al., 2016). Bruce et al. (2014) concluded that the reliability of Visual Motor Speed and Reaction Time composites varied based on languages used (English, French, Czech and Swedish) from low to high, but were mostly marginal to adequate. However, test-retest reliabilities for Verbal and Visual memory were consistently low regardless of the language used. Possible influential factors explaining the differences in reliability estimates might be the lack of equivalence of these translations and the varying psychometric properties of each test version.

There is lack of comparative information available concerning the adequacy of the Finnish (or other ImPACT language versions) translation and the cultural adaption. There are, for example, differences in word length and word frequencies that may cause differences between language versions. Other possible influential factor might be differing educational systems. For example, Echemendia et al. (2020) discovered that there were cross-cultural differences in test performance especially in speed related tasks. They concluded that timelined tests are routine in North American educational systems (Echemendia & Julian, 2002) while other cultures may not share the same predilection to speed of performance (Ardila, 1995; Agranovich & Puente, 2007). Based on these findings, language-specific reliability estimates should be utilized in calculating and assessing post-injury reliable change indices in clinical settings.

Iverson et al. (2003) published the following quick reference values for reliable change in ImPACT for healthy athletes (n = 56) in 80 % confidence interval: Verbal Memory ≥ 9 points, Visual Memory ≥ 14 points, Reaction Time ≥ .06 seconds, Visual Motor Speed ≥ 3 points, and Postconcussion Scale ≥ 10 points. These reliable change values are derived from a sample of athletes from USA and cannot directly be adopted to Finnish sample as there are regional differences in neurocognitive performance (Echemendia et al. 2020, Vartiainen et al. 2019). In present Finnish sample the reliable decline was defined as follows: Verbal Memory ≥14 points, Visual Memory ≥15 points, Reaction Time ≥ .10 s, Visual Motor Speed ≥6 points, and Postconcussion Scale ≥9 points. In the Finnish sample larger difference between baseline and post-injury composite scores were required for reliable change than in the sample from USA (Iverson et al., 2003). For example, five points higher score change was required for reliable change in Verbal Memory performance in the Finnish sample compared to the sample from USA. In the current study larger sample size (312 versus 56) and longer test-retest interval (one year versus six days) was used to calculate the reference values increasing the reliability of the results compared to Iverson et al. (2003) study. These previously published reference values may have limited applicability to Finnish sample in real life test interval scenarios. If the reference values within the ImPACT test are derived from the US sample, they may be prone to generate false positives in detecting impairment for Finnish athletes. This, however, is a smaller risk in terms of athlete health compared to potential false negative findings. The findings of current study do not support the blind use of ImPACT with English norms in multilingual populations.

When the reliable change methodology was applied to concussed sample in current study, 44 % of athletes showed reliable declines across one or more of the five individual composite scores. Approximately 56 % of concussed athletes showed no decline in cognitive performance three days from the injury which is considerably higher than the proportion reported in the Iverson et al. (2003) study where only 24 % showed no cognitive decline. This may suggest that the present sample included milder cases than Iverson et al. Alternatively, the *RCI* reference values derived from the Finnish sample are less sensitive because they are based on values with higher variability (lower test-retest correlation and higher reliable change indices) than those with Iverson et al. In the present study, the most sensitive composite score detecting concussion was Reaction Time; performance in this domain declined for 22 % of the concussed athletes. Scores for Visual Memory and Postconcussion Scale declined in approximately 16 % of concussed cases.

Traditionally, loss of consciousness and amnesia have been essential markers of concussion in injury severity grading scales (Cantu, 2001). However, grading scales have been abandoned by the Concussion in Sports Group due to conflicting results about the predictive utility of these variables (Hon, Leung & Torres, 2019; McCrory et al., 2017), while for example brief loss of consciousness does not predict clinical course or long-term cognitive impairment (Hon, Leung & Torres, 2019; McCrory et al., 2017). Loss of consciousness, amnesia, disorientation, postural instability and vacant look are still considered the most important concussion markers indicating possible concussion diagnosis (Echemendia et al. 2018; Gardner, Kohler, Levi & Iverson, 2017; Gardner, Howell & Iverson, 2018; McCrory et al., 2013). In the present sample, athletes who experienced loss of consciousness, amnesia or postural instability showed an approximately eight-fold increased risk of decline in post-injury cognitive evaluation. Mixed results have been found in studies exploring the relationship between acute concussion signs and post-injury cognitive functioning (Iverson et al. 2017). For example, Collins et al. (2003) found that athletes demonstrating poor neurocognitive functioning at two days post-injury were over 10 times more likely ($p < .001$) to have exhibited retrograde amnesia and over four times more likely ($p < .013$) to have exhibited posttraumatic amnesia; conversely, loss of consciousness was not a significant predictor of post-injury cognitive functioning. The role of postural instability predicting post-injury cognitive deficits has been less highlighted in research literature before than the role of loss of consciousness and amnesia. Current study shows that the postural instability is an equally strong predictor of post-injury cognitive deficits as loss of consciousness and amnesia. Despite the lack of consistency between studies – the risks associated with premature return to play following injury warrant a conservative approach. The results of current study suggest that, the presence of acute loss of consciousness, amnesia or postural instability may indicate more severe injury and predict the need for extended cognitive follow-up. Additionally, these findings support the idea of neurocognitive and postural stability testing being included to the concussion management protocol (McCrory et al., 2013).

This study is not without limitations. We only examined a sample of male ice hockey players, which may limit the generalizability of the results to other athletic populations that use

ImPACT. While the sample of healthy athletes allowed for robust *RCI* calculations, the sample size for concussed athletes was small. Because assessed athletes were young with cognitive abilities still maturing, it is possible that this maturational process contributed to lower test-retest reliability estimates in the present sample - especially in the areas of speed-related cognitive abilities like reaction time and visual motor speed which have strong maturational patterns in people of this age (Peltonen et al., 2019). We chose to use the liberal invalidity criteria (Impulse Composite Score >30) based on previous research, although there are also other recommended validity criteria for ImPACT. Medical personnel were trained to identify the signs of concussion and to use SCAT assessment method at the sideline evaluation. However, no fidelity checks or interrater-reliability checks were performed in current study. Despite mentioned limitations, the selected study design reflects real world situations suggesting that practical real life-training and reporting systems can inform clinical decision making and advance understanding of concussion.

This was the first study to examine the reliability of Finnish version of ImPACT with nationwide sample of considerable size. Several areas of methodological strength support the conclusions drawn. The sample was quite homogenous, a strength in reliability testing, as all the athletes tested were young male ice hockey players. Baseline assessments were administered one year apart. This is considered to be long enough time interval to reflect the real world-situation, as the time between the preseason to post-concussion assessment is usually measured in years or months rather than in days (Bruce et al. 2014). Assessments one year apart also minimize the practice effect. Efforts to statistically control for practice effects were also employed. This study adds information about the cross-national differences on neurocognitive assessment and is the first to provide reliable change scores that have potential for clinical application in the Finnish population.

Funding

This work was supported by the Ministry of Social Affairs and Health Finland (grant number 201510103) and the Finnish Cultural Foundation (grant number 00190810).

References

Agranovich, A., & Puente, A. (2007). Do Russian and American normal adults perform similarly on neuropsychological tests? Preliminary findings on the relationship between culture and test performance. *Archives of Clinical Neuropsychology*, 22(3), 273–282. doi:10.1016/j.acn.2007.01.003

Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology*. 17(1), 143–150.

Broglio, S. P., Ferrara, M. S., Macciocchi, S. N., Baumgartner, T. A., & Elliott, R. (2007). Test-retest reliability of computerized concussion assessment programs. *Journal of Athletic Training*, 42(4), 509–514.

Broglio, S.P., Ferrara, M.S., Piland, S.G., & Anderson, R.B. (2006). Concussion history is not a predictor of computerized neurocognitive performance. *British Journal of Sports Medicine*, 40, 802–805.

Bruce, J., Echemendia, R., Meeuwisse, W., Comper, P., & Sisco, A. (2014). 1-year test–retest reliability of ImPACT in professional ice hockey players. *The Clinical Neuropsychologist*, 28(1), 14–25. doi:10.1080/13854046.2013.866272

Bure-Reyes, A., Hidalgo-Ruzzante, N., Vilar-Lopez, R., Gontier, J., Sanchez, L., Perez-Garcia, M., & Puente, A.E. (2013). Neuropsychological test performance of Spanish speakers: Is performance different across different Spanish-speaking groups? *Journal of Clinical and Experimental Neuropsychology*, 35, 404–412. doi:10.1080/13803395.2013.778232

Cantu, R.C. (2001). Posttraumatic (retrograde and anterograde) amnesia: Pathophysiology and implications in grading and safe return to play. *Journal of Athletic Training*, 36(3), 244–248.

Chelune, G.J., Naugle, R.I., Lüders, H., Sedlak, J., & Awad, I.A. (1993). Individual change after epilepsy surgery: practice effects and base-rate information. *Neuropsychology*, 7, 41–52.

Cherner, M., Suarez, P., Lazzaretto, D., Fortuny, L., Mindt, M., Dawes, S., … the HNRC group. (2007). Demographically corrected norms for the Brief Visuospatial Memory Test-revised and Hopkins Verbal Learning Test-revised in monolingual Spanish speakers from the

U.S.–Mexico border region. *Archives of Clinical Neuropsychology*, 22(3), 343–353. doi:10.1016/j.acn.2007. 01.009

Collins, M.W., Iverson, G.L., Lovell, M.R., McKeag, D.B., Norwig, J., & Maroon, J. (2003). On-field predictors of neuropsychological and symptom deficit following sports-related concussion. *Clinical Journal of Sport Medicine*, 13(4), 222–229. http://doi.org/10.1097/00042752-200307000-00005

D'Elia, L.F., Satz, P., Uchiyama, C.L., & White, T. (1996). Color Trails Test. Odessa, FL: Psychological Assessment Resources.

Dougan, B.K., Horswill, M.S., & Geffen, G.M. (2013). Do Injury Characteristics Predict the Severity of Acute Neuropsychological Deficits Following Sports-Related Concussion? A Meta-analysis. *Journal of the International Neuropsychological Society*, 20(01), 81–87. http://doi.org/10.1017/S1355617713001288

Echemendia, R.J., Bruce, J.M., Meeuwisse, W., Hutchison, M.G., Comper, P., & Aubry, M. (2018). Can visible signs predict concussion diagnosis in the National Hockey League? *British Journal of Sports Medicine*, 52, 1149–1154.

Echemendia, R. J., & Julian, L. (2002). Neuropsychological assessment of Latino/Hispanic children. In Ferraro, R. (Ed.), *Minority and cross-cultural aspects of neuropsychological assessment* (pp. 181–203). Lisse: Swets and Zeitlinger.

Echemendia, R.J., Meeuwisse, W., McCrory, P., Davis, G.A., Putukian, M., Leddy, J. … Herring, S. (2017). The Sport Concussion Assessment Tool 5th Edition (SCAT5): Background and rationale. *British Journal of Sports Medicine*, 51, 848–850. doi:10.1136/bjsports-2017-097506

Echemendia, R.J., Thelen,J., Meeuwisse, W., Hutchison, M.G., Rizos, J., Comper, P., & Bruce, J.M. (2020). Neuropsychological Assessment of Professional Ice Hockey Players: A Cross-Cultural Examination of Baseline Data Across Language Groups, *Archives of Clinical Neuropsychology*, acz077. https://doi.org/10.1093/arclin/acz077

Elbin, R.J., Schatz, P., & Covassin, T. (2011). One-year test-retest reliability of the online version of ImPACT in high school athletes. *American Journal of Sports Medicine*, 39, 2319–2324.

Galetta, K.M., Barrett, J., Allen, M., Madda, F., Delicata, D., Tennant, A.T. … Balcer, L.J. (2011). The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology,* 76, 1456–1462.

Galetta, K.M., Liu, M., Leong, D.F., Ventura, R.E., Galetta, S.L., & Balcer, L.J. (2016). The King-Devick test of rapid number naming for concussion detection: meta-analysis and systematic review of the literature. *Concussion*, 1(2), CNC8. doi: 10.2217/cnc.15.8

Gardner, A.J., Kohler, R.M.N., Levi, C.R., & Iverson, G.L. (2017). Usefulness of Video Review of Possible Concussions in National Youth Rugby League. *International Journal of Sports Medicine*, 38,71–75.

Gardner, A.J., Howell, D.R., & Iverson, G.L. (2018). A video review of multiple concussion signs in National Rugby League match play. *Sports Medicine – Open*, 4, 0–7.

Guskiewicz, K., Register-Mihalik, J., McCrory, P., McCrea, M., Johnston, K., Makdissi, M. … Meeuwisse, W. (2013). Evidence-based approach to revising the SCAT2: Introducing the SCAT3. *British Journal of Sports Medicine*, 47, 289–293. http://doi.org/10.1136/bjsports-2013-092225

Guskiewicz, K. M., Ross, S. E., & Marshall, S. W. (2001). Postural Stability and Neuropsychological Deficits After Concussion in Collegiate Athletes. *Journal of athletic training*, *36*(3), 263–273.

Harmon, K.G., Clugston, J.R., Dec, K., Hainline, B., Herring, S., Kane, S.F. … Roberts, W.O. (2019). American Medical Society for Sports Medicine position statement on concussion in sport. *British Journal of Sports Medicine*, 53, 213–225.

Hon, K.L., Leung, A.K.C., & Torres, A.R. (2019) Concussion: A Global Perspective, *Seminars in Pediatric Neurology*, 30, 117–127. https://doi.org/10.1016/j.spen.2019.03.017.

Iverson, G.L. (2005). Outcome from mild traumatic brain injury. *Current Opinion in Psychiatry*, 18, 301–317.

Iverson, G.L. (2007). Predicting slow recovery from sport-related concussion: the new simple-complex distinction. *Clinical Journal of Sport Medicine*, 17, 31–37.

Iverson, G.L., Echemedia, R.J., LaMarre, A.K., Brooks, B.L., & Gaetz, M.B. (2012). Possible lingering effects of multiple past concussions. *Rehabilitation Research and Practice*, 2012, 1–8. Article ID 316575.

Iverson, G.L., Gardner, A.J., Terry, D.P. Ponsford, J.L., Skills A.K., Broshek, D.K., & Solomon, G.S. (2017). Predictors of clinical recovery from concussion: a systematic review. *British Journal of Sports Medicine,* 51, 941–948.

Iverson, G.L., Lovell, M.R., & Collins, M.W. (2002). Immediate Postconcussion Assessment and Cognitive Testing (ImPACT) normative data version 2.0. Pittsburgh, PA: ImPACT Applications Inc.

Iverson, G.L., Lovell, M.R., & Collins, M.W. (2003). Interpreting Change on ImPACT Following Sport Concussion. *The Clinical Neuropsychologist*, 17(4), 460–467.

Jacobson, N.S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* ,59(1), 12–19.

Jones, N.S., Walter, K.D., Caplinger, R.,Wright, D., Raasch, W.G. & Young, C. (2014). Effect of education and language on baseline concussion screening tests in professional baseball players. *Clinical Journal of Sport Medicine*, 24(4), 284–288.

Lau, B., Lovell, M.-R., Collins, M.W., & Pardini, J. (2009). Neurocognitive and symptom predictors in recovery in high school athletes. *Clinical Journal of Sport Medicine*, 19(3), 216–221.

Lau, B.C., Kontos, A.P., Collins, M.W., Mucha, A., & Lovell, M.R. (2011). Which on-field signs/symptoms predict protracted recovery from sport-related concussion among high school football players? *The American Journal of Sports Medicine* ,39(11), 2311–2318. http://doi.org/10.1177/0363546511410655

Littleton, A. C., Schmidt, J. D., Register-Mihalik, J.K., Gioia, G.A., Waicus, K.M., Mihalik, J.P. & Guskiewicz, K. M. (2015). Effects of attention deficit hyperactivity disorder and stimulant medication on concussion symptom reporting and computerized neurocognitive test performance. *Archives of Clinical Neuropsychology*, 30(7), 683–693.

Lovell, M.R. (2004). ImPACT Version 2.0 Clinical User's Manual. Pittsburgh, PA: ImPACT Inc.

Lovell, M.R. (2006). The ImPACT Neuropsychological Test Battery. In R.J. Echemendia (Ed.), *Sports neuropsychology: Assessment and management of traumatic brain injury* (pp. 193–215). New York: Guilford Press.

Lovell, M.R., Collins, M.W., Iverson, G.L., Field, M., Maroon, J.C., Cantu, R. … Fu, F.H. (2003). Recovery from mild concussion in high school athletes. *Journal of Neurosurgery*, 98(2), 295–301. http://doi.org/10.3171/jns.2003.98.2.0296

Lovell, M.R., Iverson, G.L., Collins, M.W., McKeag, D., & Maroon, J.C. (1999). Does loss of consciousness predict neuropsychological decrements after concussion? *Clinical Journal of Sport Medicine*, 9(4), 193-198.

Maerlander, A., Flashman, L., Kessler, A., Kumbhani, S., Greenwald, R., Tosteson, T., & McAllister, T. (2010). Examination of the construct validity of ImPACT computerized test, traditional, and experimental neuropsychological measures. *The Clinical Neuropsychologist*, 24, 1309–1325.

McCrory, P., Meeuwisse, W.H., Aubry, M., Cantu, R.C., Dvořák, J., Echemendia, R.J. … Turner, M. (2013). Consensus Statement on Concussion in Sport: The 4th International Conference on Concussion in Sport Held in Zurich, November 2012. *British Journal of Sports Medicine*, 47, 250–258. http://doi.org/10.1016/j.pmrj.2013.02.012

McCrory P, Meeuwisse W, Dvorak J, Aubry, M., Bailes, J., Broglio, S. … Vos, P.E.(2017). Consensus statement on concussion in sport—the 5th international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine*, 51, 838–847.

Nelson, L., Guskiewicz, K., Marshall, S., Hammeke, T., Barr, W., Randolph, C., & McCrea, A. (2016a). Multiple self-reported concussions are more prevalent in athletes with ADHD and learning disability. *Clinical Journal of Sports Medicine*, 26(2), 120–127.

Nelson, L., LaRoche, A., Pfaller, A., Lerner, E., Barr, H. C., Guskiewicz, W., & McCrea, M. (2016b). Prospective, head-to-head study of three computerized neurocognitive assessment tools (CNTs): Reliability and validity for the assessment of sport-related concussion. *Journal of the International Neuropsychological Society*, 22, 24–37.

Norman, M., Moore, D., Taylor, M., Franklin, D., Cysique, L., Ake, C., … the HNRC Group. (2011). Demographically corrected norms for African Americans and Caucasians on the Hopkins verbal learning test-revised, brief visuospatial memory test-revised, Stroop color and word test, and Wisconsin card sorting test 64-card version. *Journal of Clinical and Experimental Neuropsychology*, 33, 793–804.

OECD (2013). Education at a Glance. Retrieved September 18, 2020, from http://www.oecd.org/edu/eag2013%20%28eng%29--FINAL%2020%20June%202013.pdf

Peltonen, K., Launes, J., Koskinen, S., Vartiainen, M., Pajunen, S., Pertab, J., Laitala, T. & Hokkanen, L. (2020). On-field signs of concussion predict deficits in cognitive functioning: Loss of consciousness, amnesia, and vacant look. *Translational Sports Medicine*, 00, 1–9. https://doi.org/10.1002/tsm2.179

Peltonen, K., Vartiainen, M., Laitala-Leinonen, T., Koskinen, S., Luoto, T., Pertab, J., & Hokkanen, L. (2018). Adolescent athletes with learning disability display atypical maturational trajectories on concussion baseline testing: Implications based on a Finnish sample, *Child Neuropsychology*, 25(3), 336–351, DOI: 10.1080/09297049.2018.1474865

Schatz, P., Pardini, J.E., Lovell, M.R., Collins, M.W., & Podell, K. (2006). Sensitivity and specificity of the ImPACT test battery for concussion in athletes. *Archives of Clinical Neuropsychology*;21(1):91–99.

Schatz, P., & Putz, B.O. (2006). Cross-validation of measures used for computer-based assessment of concussion. *Applied Neuropsychology*, 13, 151–159.

Schatz, P., & Sandel, N. (2013). Sensitivity and specificity of the online version of ImPACT in high school and collegiate athletes. *American Journal of Sports Medicine*, 41, 321–326.

Shuttleworth-Edwards, A.B., Kemp, R.D., Rust, A.L., Muirhead, J.G.L., Hartman, N.P., & Radloff, S.E. (2004). Cross-cultural effects on IQ test performance: A review and preliminary normative indications on WAISIII test performance. *Journal of Clinical and Experimental Neuropsychology*, 26(7), 903–920. doi:10.1080/13803390490510824

Tsushima, W.T., Siu, A.M., Pearce, A.M., Zhang, G., & Oshiro, R.S. (2016). Two-year Test–Retest Reliability of ImPACT in High School Athletes. *Archives of Clinical Neuropsychology*, 31, 105–111.

Van Kampen, D.A., Lovell, M.R., Pardini, J.E., Collins, M.W., & Fu, F.H. (2006). The ”Value Added” of Neurocognitive Testing After Sports-Related Concussion. *The American Journal of Sports Medicine*, 34(10), 1630–1635. http://doi.org/10.1177/0363546506288677

Vartiainen, M.V., Peltonen, K., Holm, A., Koskinen, S., Iverson, G.L., & Hokkanen, L. (2019). Preliminary normative study of ImPACT® in Finnish professional male ice hockey players. *Applied Neuropsychology: Adult*, 1–7. https://doi.org/10.1080/23279095.2019.1592175

Zahra, D. & Hedge, C. (2010). The Reliable Change Index: Why isn’t it more popular in academic psychology? *PsyPag Quarterly*, 76, 14–20.

Table 1. Test-retest reliability for healthy control group n=312 tested one year apart

Composite	*M (SD)*		ICC	ICC 95%		t^a	*p*
	Time 1	Time 2		Lower	Upper		
Verbal Memory	82.83 (9.92)	84.54 (10.71)	.59	.49	.67	2.73	.01
Visual Memory	70.99 (12.91)	75.13 (12.86)	.70	.59	.77	6.08	<.001
Visual Motor Speed	34.47 (5.93)	37.22 (6.38)	.78	.58	.87	10.31	<.001
Reaction Time	.63 (.08)	.60 (.07)	.58	.42	.69	-7.50	<.001
Postconcussion Scale	7.29 (8.42)	3.41 (4.31)	.43	.22	.57	-8.76	<.001

Notes: *ICC*=intraclass correlation coefficient; *CI*=confidence interval; *p*, significance of *t*.
[a] df = 311; Bonferroni-corrected α $p < .01$.

Table 2. Reliable change incides (*RCI*) in the healthy control group *n*=312.

Composite	*M (SD)*		*r*	SEM_1	SEM_2	S_{diff}	Confidence intervals	
	Time 1	Time 2					.80	.90
Verbal Memory	82.83 (9.92)	84.54 (10.71)	.43	7.49	8.09	11.02	14.11	18.07
						10.59	13.56	17.37
Visual Memory	70.99 (12.91)	75.13 (12.86)	.56	8.56	8.53	12.09	15.47	19.82
						12.11	15.50	19.86
Visual Motor Speed	34.47 (5.93)	37.22 (6.38)	.71	3.19	3.44	4.69	6.00	7.69
						4.52	5.78	7.41
Reaction Time	.63 (.08)	.60 (.07)	.46	.06	.05	.08	0.10	0.13
						.08	0.11	0.14
Postconcussion Scale	7.29 (8.42)	3.41 (4.31)	.39	6.58	3.37	7.39	9.46	12.12
						9.30	11.90	15.25

S_{diff}, standard error of difference scores based on Chelune et al (1993), upper row, $\sqrt{([SEM_1^2]+[SEM_2^2])}$; and Jacobson and Truax, (1991) lower row, $\sqrt{(2*[SEM_1^2])}$.

Table 3. Percentages of the control group (*n*=312) that would be classified as reliably improved or declined (Chelune et al. 1993) based on the 0.80 and 0.90 confidence intervals.

	.80 confidence interval			.90 confidence interval		
	Declined (%)	Improved (%)	Total (%)	Declined (%)	Improved (%)	Total (%)
Verbal Memory	7.7	9.3	17	2.6	3.5	6.1
Visual Memory	5.1	15.1	20.2	1.6	2.9	4.5
Visual Motor Speed	3.8	21.5	25.3	1.9	6.7	8.6
Reaction Time	4.2	13.5	17.7	1.3	4.8	6.1
Postconcussion Scale	1.0	19.2	20.2	0.0	8.3	8.3

Table 4. Percentages of the concussed sample (*n*=32) that would be classified as reliably improved or declined (Chelune et al. 1993) based on the 0.80 and 0.90 confidence intervals.

	.80 confidence interval			.90 confidence interval		
	Declined (%)	Improved (%)	Total (%)	Declined (%)	Improved (%)	Total (%)
Verbal Memory	9.4	6.3	15.7	9.4	3.1	12.5
Visual Memory	15.6	3.1	18.7	12.5	3.1	15.6
Visual Motor Speed	6.3	9.4	15.7	3.1	6.3	9.4
Reaction Time	21.9	3.1	25.0	18.8	3.1	21.9
Postconcussion Scale	15.6	9.4	25.0	15.6	3.1	18.7

Table 5.

Chi square and Odds ratio for having 2 or more declines on ImPACT in the concussion sample (*n*=32). Many athletes showed more than one sign.

Sign	Present (%)	Absent (%)	Chi square	Odds ratio (*OR*)	*OR* 95% *CI*
Loss of consciousness	6 (18.75)	26 (81.25)	4.73, p < .05	7.67, p < .05	1.04–56.77
Amnesia	6 (18.75)	26 (81.25)	4.73, p < .05	7.67, p < .05	1.04–56.77
Disorientation	22 (68.75)	10 (31.25)	.02, p = .90	.89, p = .45	.13–5.89
Postural instability	15 (46.88)	17 (53.13)	3.94, p < .05	8.00, p < .05	.81–78.83
Vacant look	14 (43.75)	18 (56.25)	1.58, p = .21	3.20, p = .11	.49–20.81