Standard regression-based methods for measuring recovery after sport-related concussion

MICHAEL McCREA,[1,2] WILLIAM B. BARR,[3] KEVIN GUSKIEWICZ,[4,5,6]
CHRISTOPHER RANDOLPH,[7] STEPHEN W. MARSHALL,[6,8] ROBERT CANTU,[4,9]
JAMES A. ONATE,[10] AND JAMES P. KELLY[11]

[1]Neuroscience Center, Waukesha Memorial Hospital, Waukesha
[2]Department of Neurology, Medical College of Wisconsin, Milwaukee
[3]Departments of Neurology and Psychiatry, New York University School of Medicine, New York
[4]Department of Exercise and Sport Science, University of North Carolina at Chapel Hill, Chapel Hill
[5]Department of Orthopedics, University of North Carolina at Chapel Hill, Chapel Hill
[6]Injury Prevention Research Center, University of North Carolina at Chapel Hill, Chapel Hill
[7]Department of Neurology, Loyola University Medical School, Maywood
[8]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill
[9]Neurosurgery Service, Emerson Hospital, Concord
[10]Department of Rehabilitation Sciences, Athletic Training Program, Sargent College of Health and Rehabilitation Sciences, Boston University, Boston
[11]Department of Neurosurgery, University of Colorado Health Sciences Center, Denver

Abstract

Clinical decision making about an athlete's return to competition after concussion is hampered by a lack of systematic methods to measure recovery. We applied standard regression-based methods to statistically measure individual rates of impairment at several time points after concussion in college football players. Postconcussive symptoms, cognitive functioning, and balance were assessed in 94 players with concussion (based on American Academy of Neurology Criteria) and 56 noninjured controls during preseason baseline testing, and immediately, 3 hr, and 1, 2, 3, 5, and 7 days postinjury. Ninety-five percent of injured players exhibited acute concussion symptoms and impairment on cognitive or balance testing immediately after injury, which diminished to 4% who reported elevated symptoms on postinjury day 7. In addition, a small but clinically significant percentage of players who reported being symptom free by day 2 continued to be classified as impaired on the basis of objective balance and cognitive testing. These data suggest that neuropsychological testing may be of incremental utility to subjective symptom checklists in identifying the residual effects of sport-related concussion. The implementation of neuropsychological testing to detect subtle cognitive impairment is most useful once postconcussive symptoms have resolved. This management model is also supported by practical and other methodological considerations. (*JINS*, 2005, *11*, 58–69.)

Keywords: Brain injury, Brain concussion, Athletic injuries

INTRODUCTION

Sport-related concussion is now widely recognized as a major public health concern in the United States and worldwide (Aubry et al., 2002; Collins et al., 1999; Kelly, 1999). The Centers for Disease Control and Prevention report that approximately 300,000 sport-related concussions occur annually in the United States (Thurman et al., 1998), and recent data from the National Collegiate Athletic Association (NCAA) Injury Surveillance System reveal that concussion is amongst the most frequently observed injuries in collegiate ice hockey, football, and soccer (Dick, 2003). Epidemiological and prospective clinical studies estimate that approximately 3% to 8% of high school and collegiate football players sustain a concussion each season (Collins et al., 1999; Echemendia et al., 2001; Guskiewicz et al., 2000, 2003; Macciocchi et al., 1996; McCrea et al., 1997, 1998, 2002; Powell & Barber-Foss, 1999), and the rate of documented concussion in collegiate football has been on the rise in recent years (Covassin et al., 2003; Dick, 2003).

Sports medicine professionals consider diagnosing concussion and determining postinjury recovery among their greatest clinical challenges, especially when confronted with intense pressure to formulate a rapid assessment of injury and prompt decision about how soon an athlete can safely return to competition (Cantu, 1986; Kelly, 1999; Vastag, 2002). Despite a growing body of research, there remains little evidence-based guidance on how long it takes for an athlete to recover and deciding when it is safe for the athlete to return to play after a concussion. Existing guidelines largely stem from expert consensus, with limited scientific support from prospective studies (Aubry et al., 2002; Kelly, 1999; Vastag, 2002). As a result, sports medicine clinicians have historically relied on their experience and subjective observations to track postinjury recovery and guide clinical decision making about return-to-play.

Neuropsychologists have made a significant contribution to advancing the scientific understanding of the effects and expected recovery course after sport-related concussion (Echemendia & Julian, 2001). Several neuropsychological studies have reported estimates of recovery in symptoms and cognitive dysfunction ranging from several hours to several days (Collins et al., 1999, 2003; Echemendia et al., 2001; Erlanger et al., 2003; Hinton-Bayre et al., 1999; Lovell et al., 2003; Macciocchi et al., 1996; Maddocks & Saling, 1996; McCrea, 2001; McCrea et al., 1998, 2002). Interpretation of recovery data across studies has been hampered by a number of methodological factors, however, including small sample sizes, varied definitions of concussion, absence of an immediate injury assessment to measure the most acute effects, limited follow-up assessment of injured players, failure to establish a recovery endpoint, and lack of an appropriate control group. The findings from a recent 3-year prospective study demonstrated that collegiate football players, on average, followed a gradual course of recovery in symptoms, cognitive functioning, and postural stability over the first 5–7 days after concussion, but that 10% of injured players required more than 7 days to reach a full recovery (McCrea et al., 2003). The study focused on analysis of group effects, limiting the application of the findings to actual clinical decision making in individual cases.

There has been a growth of recent interest in defining the most clinically useful methods for detecting change in neuropsychological test scores in test–retest situations. The emphasis has been to move beyond analysis of group statistics toward the detection of change in individual participants. Much attention has focused on the use of the reliable change index (RCI) and standardized regression-based (SRB) methods, which are techniques initially developed and refined in studies of outcome from psychotherapy and surgical treatment. Conclusions drawn from recent reviews indicate that, while RCI methods may be easier to use in a clinical setting, SRB methods are demonstrated to be more accurate for detecting meaningful change as a result of their ability to correct for practice effects, regression to the mean, and the impact of baseline test performance . The SRB methods are thus recommended for use in research settings examining test–retest changes (Barr, 2002; Temkin et al., 1999).

Techniques to more precisely measure and characterize meaningful and reliable change in neurocognitive test performance in individual patients over time are especially intriguing in the case of sports concussion assessment. The prototypic sports concussion assessment model implemented in many professional, collegiate, and high school athletic programs incorporates a preseason baseline evaluation of all players that includes a clinical history, base rate symptom index, neurocognitive battery, and postural stability testing. Any player who sustains a concussion over the course of the season is then reevaluated with these measures at several time points, initially to determine the severity of their injury and then track their postinjury recovery. Clinical research programs also include matched, noninjured controls in the identical testing protocol. Despite the methodological advantages of this model, serial testing of injured athletes presents the clinician with the challenge of distinguishing between "real" change in individual test performance indicative of true recovery *versus* performance variability due to practice effects, measurement error, or random influence.

The current study sponsored by the NCAA applied a serial testing paradigm and regression-based methods to statistically define recovery and measure individual rates of impairment at several time points after concussion in collegiate football players.

METHODS

Research Participants

A total of 1631 football players from 15 National Collegiate Athletic Association (NCAA) Division I, II, and III member institutions were enrolled in one arm of a larger cohort study of the effects of sports-related concussion from 1999 through 2001. In total, 2410 player seasons were analyzed, as 779 players were enrolled for more than 1 year of the study. A case series of 94 injured players who sustained a concussion (5.76% of players, 3.90% of player seasons) were enrolled in an extensive injury assessment protocol. No player who sustained a concussion refused to participate or was excluded from the study protocol, but information on unidentified or unreported concussions was not available.

Fifty-six noninjured controls matched to injured participants on the basis of age, years of education, team, and baseline performance on concussion assessment measures were administered the identical protocol under the same conditions and retest intervals as injured players during the first year of the study. Limited resources did not allow enrollment of additional control participants in years 2 or 3 of the study, which had a minimal effect on matching characteristics for the complete study sample. Table 1 provides a comparison of demographic and medical history data for the injured and control groups.

Table 1. Concussion and control group characteristics

	Concussion (n = 94)		Control (n = 56)		Mean Diff.	t	p
	Mean	*SD*	*Mean*	*SD*			
Demographics:	73.50	2.94	72.75	3.23	.75	1.44	.151
Weight (lbs.)	235.26	46.88	218.50	46.19	16.76	2.11	.037*
Age (years)	20.04	1.36	19.20	1.45	.84	3.51	.001*
Academic year (collegiate)	2.78	1.18	2.02	1.23	.76	3.71	.001*
Self-reported history of:							
No. of previous concussions (past 7 years)	.58	.78	.39	.68	.19	1.47	.145
Range	0–5		0–3				
Any concussion (lifetime) (%)	43.2		30.4		$\chi^2 = 2.78$		.123
ADHD (%)	2.30		1.80		$\chi^2 = .034$		.854
Learning disability (%)	2.30		1.80		$\chi^2 = 1.64$		.440

Notes. *Statistically significant. ADHD = Attention Deficit Hyperactivity Disorder. LD = Learning Disability.

This study was approved by the Institutional Review Board (IRB) for protection of human research participants at the host institutions of the principal investigators. All participants granted written informed consent prior to enrollment in the study.

Study Design

All players underwent a preseason baseline evaluation on a battery of concussion assessment measures and extensive health questionnaire prior to their first year of participation in this study. Injured participants were identified and enrolled in the study protocol by a team physician or certified athletic trainer present on the sideline during an athletic contest or practice. *Concussion* was defined as an injury resulting from a blow to the head causing an alteration in mental status and one or more of the following symptoms prescribed by the American Academy of Neurology Guideline for Management of Sports Concussion (Practice Parameter, 1997): headache, nausea, vomiting, dizziness/balance problems, fatigue, trouble sleeping, drowsiness, sensitivity to light or noise, blurred vision, difficulty remembering, or difficulty concentrating (Kelly & Rosenberg, 1997). Loss of consciousness (LOC), posttraumatic amnesia (PTA) (e.g., inability to recall exiting the field, aspects of the examination, etc.), and retrograde amnesia (RGA) (e.g., in ability to recall aspects of the play, events prior to injury, score of the game, etc.) were documented immediately after injury.

All players identified by the team physician or certified athletic trainer as having sustained a concussion according to the study's injury criteria were tested with a Graded Symptom Checklist (GSC) (Lovell & Collins, 1998), the Standardized Assessment of Concussion (SAC) (McCrea et al., 2000), and the Balance Error Scoring System (BESS) (Guskiewicz et al., 2001) on the sideline immediately following injury. Follow-up testing on these measures was then conducted postgame/postpractice (2–3 hr after injury), and again on postinjury days 1, 2, 3, 5, and 7. The brief

neuropsychological test battery (see Table 2) was administered to assess neurocognitive functioning at baseline and on days 2 and 7 postinjury. A day-90 assessment point was also included in the original design, but significant attrition for both injured and control participants did not allow application of the standard regression-based methods employed in this study for this assessment point. Assessments were conducted by certified athletic trainers who were trained by the researchers and required to watch a training video on administration and scoring of all outcome measures used in the study.

Main Outcome Measures

A summary of the main outcome measures used in this study to assess postconcussive symptoms, cognitive functioning, and postural stability is provided in Table 2. Several studies on the effects of sport-related concussion have demonstrated the reliability and accuracy of the GSC (Lovell et al., 2003), SAC (Barr & McCrea, 2001; McCrea, 2001), BESS (Guskiewicz et al., 2001; Riemann et al., 1999; Riemann & Guskiewicz, 2000), and components of the neuropsychological test battery (Collins et al., 1999) in correctly classifying injured and noninjured participants after concussion.

Statistical Analysis

There was limited missing data, with 93% of data cells complete across all time points for all participants. To examine the potential effect of missing data on the results, we compared the baseline scores for the missing and nonmissing participants at every time point for all outcomes. The baseline scores did not differ between missing and nonmissing, suggesting that the data was missing at random (Diggle et al., 1994). As part of our previous analysis (McCrea et al., 2003), we also estimated the missing data using a single imputation model (Rubin, 1976, 1996; Schafer, 1997), based on time, participants status (injured *vs.* control), and

Table 2. Assessment measure characteristics

Measure	Functional domain	Description	Score range	Time to administer
Graded Symptom Checklist (GSC)	Postconcussive symptoms	Subject rates presence and severity of 17 symptoms (e.g., headache, dizziness, etc.)	0 (no symptoms)–6 (severe) Likert scale per item; total score range: 0–102; higher score indicates more severe symptoms	2–3 min
Standardized Assessment of Concussion (SAC)*	Cognitive functioning: - orientation - immediate & delayed memory - concentration Neurologic Screening: - strength - sensation - coordination	Brief neurocognitive assessment and neurologic screening; documentation of unconsciousness, posttraumatic amnesia, retrograde amnesia	Total score range: 0–30; lower score indicates more severe cognitive impairment	5 min
Balance Error Scoring System (BESS)	Postural stability	Noninstrumented, clinical assessment of postural stability in double leg, single leg, tandem stances on firm and foam surfaces	No defined range; test score equals total number of errors committed by the participant; higher score indicates more severe postural instability	5 min
Neuropsychological Test Battery*	Cognitive functioning: - attention - concentration - processing speed - mental flexibility - anterograde memory	Hopkins Verbal Learning Test Trail Making Test Part B Symbol Digit Modalities Test Stroop Color-Word Test Controlled Oral Word Association Test	Total score range based on individual measures; lower score indicates more severe impairment except for Trail Making Test (total time to complete)	25 min

*Alternate forms utilized to minimize practice effects practice effects from repeat testing on SAC and neuropsychological test battery.

baseline test performance, and obtained essentially identical results on separate analysis of the nonimputed and imputed data.

Group recovery data analyses have been published previously (McCrea et al., 2003), including illustrations of the natural course of recovery in symptoms, cognitive dysfunction, and balance problems following concussion. Those analyses utilized Generalized Estimating Equations (GEE) to examine the adjusted group mean differences on each assessment measure at each assessment point, while controlling for baseline test performance and other factors known to influence performance on specific measures (e.g., years of education on cognitive testing).

The current data analysis focused on individual rates of impairment, rather than generating group recovery curves on each of the study's main assessment measures. Classification of impairment at the individual case level was based on an empirical method using SRB indices for detection of significant change in test scores (McSweeny et al., 1993; Temkin et al., 1999). This method uses linear regression on baseline scores from the healthy control group to generate a formula for predicting follow-up scores at various time points. The resulting regression coefficient and the intercept of the regression line were used with the baseline score to compute a predicted score for each participant at Time 2 and at subsequent testing points. This approach provides an empirical method for detecting meaningful change while also providing correction for practice effects and regression to the mean.

Participants were considered to have undergone a meaningful change in test performance if the difference between the obtained and predicted score, divided by the standard error of prediction, was larger than a specific criterion value, translated to a 90% confidence interval (two-tailed, 5% chance of Type I error). Predictions for the GSC, BESS, and SAC were computed for all time points from the time of injury through day 7. Predicted scores for each of the neuropsychological tests were computed for days 2 and 7. Based on the conventional standard within neuropsychology and to minimize the rate of false positives, impairment on the neuropsychological test battery was defined as having significantly decreased scores on two or more tests. Conservatively assuming complete orthogonality among outcome measures, which is not likely the case due to shared group variance across measures, the expected rates of false positives (identifying a normal participant as impaired) ranged from 5% for a single measure to 15% for impairment on any one of three measures (e.g., impairment on the brief battery of the GSC, BESS, and SAC), and ranged up to 27.3% when adding the criterion of impairment on two of the seven measures in the neuropsychological test battery.

Frequencies were also generated for the percentage of participants who had reached a full symptom recovery based on GSC score, but continued to show impairment on the BESS, SAC, or the neuropsychological test battery on postinjury days 2 and 7, based on the respective SRB indices for each test.

Measures of sensitivity and specificity were computed for establishing each test's ability to distinguish between the injured and control groups. In this context, sensitivity (Se) refers to the probability that an injured participant will be identified as "abnormal" by a change in test performance. At time points subsequent to time of injury, sensitivity values indicate the probability that a player originally injured continued to be classified as "abnormal" according to at least one of the test measures. Specificity (Sp) refers to the probability that a control participant will be correctly classified as "normal" using the same method. Data were analyzed with SPSS 11.0 statistical software (SPSS, 1999).

RESULTS

Ninety-four players who sustained a concussion during football practice (56.8% of injuries) or games (43.2%) were studied. Most injuries were classified as either Grade 1 or 2 concussions according to the Cantu (Cantu, 1998) (98.6%), Colorado (Colorado Medical Society, 1991) (93.3%), and American Academy of Neurology (Practice Parameter, 1997) (93.2%) sports-concussion grading scales, based on our *post-hoc* review of injury characteristics. A small number of injured participants sustained LOC (6.4%; median duration 30 s). Additionally, a small percentage of participants exhibited PTA (19.1%; median duration 90 min) or RGA (7.4%; median duration 120 min). There was no LOC, PTA, or RGA associated with most injuries (77.8%). Ninety injured participants (96%) completed the assessment protocol through the day 7 assessment point.

Results from group (concussion *vs*. control) data analyses on the GSC, BESS, SAC, and neuropsychological test battery have been reported previously, including illustrations of the pattern of recovery in symptoms, cognitive functioning, and postural stability (McCrea et al., 2003). As context for interpretation of the current results from the SRB analysis, raw means for the concussion and control groups on the GSC, BESS, and SAC at baseline and each postinjury assessment point are provided in Table 3, and group data from the neuropsychological test battery are provided in Table 4.

Results from the current analyses focus on the rate of recovery by individual players with concussion. The percentages of injured participants and controls impaired at each assessment point on the GSC, SAC, BESS, and neuropsychological test battery, as defined for each test according to the SRB calculation, are presented in Figure 1. During the acute postinjury period, the highest rate of abnormality was observed on the GSC, with 89% of the sample reporting increased symptoms at the time of injury and 74% reporting symptoms 1 day later. There was a gradual decline in the percentage of participants with elevated GSC scores on ensuing days, with only 4% reporting significant symptoms 1 week after concussion. The base rate of common concussive symptoms in the control group was zero across all assessment points.

Table 3. GSC, SAC, and BESS data for concussion and control groups at baseline and postinjury assessment points

| | GSC | | | | SAC | | | | BESS | | | |
| | Concussion | | Control | | Concussion | | Control | | Concussion | | Control | |
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Baseline	1.97	4.94	.99	3.26	27.37	2.16	27.43	1.77	11.95	8.09	12.73	7.57
Time of concussion	20.60	10.58	.20	2.54	24.94	3.07	27.69	1.91	19.46	9.48	12.34	9.06
Postgame/ postpractice	16.73	11.86	.18	1.96	25.58	3.03	27.76	1.85	16.70	9.16	12.49	9.32
Day 1	12.25	12.52	.18	.69	26.25	2.79	27.96	1.65	14.18	8.04	11.96	8.11
Day 2	7.63	10.55	.06	.45	27.44	2.32	28.02	1.51	12.96	7.26	11.20	9.40
Day 3	6.03	10.26	.04	.44	27.57	2.46	27.96	1.64	12.31	7.80	11.29	7.71
Day 5	3.06	5.95	.04	.47	28.02	3.24	28.73	1.40	10.97	6.78	11.69	7.95
Day 7	1.27	3.37	.02	.46	28.41	1.85	28.37	3.39	9.67	6.88	10.93	8.21

GSC = Graded Symptom Checklist (Lovell & Collins, 1998); SAC = Standardized Assessment of Concussion (McCrea et al., 2000); BESS = Balance Error Scoring System (Guskiewicz et al., 2001).

Acute cognitive dysfunction, as measured by impairment on the SAC, was evident in 80% of the injured sample at the time of concussion. Persistent cognitive impairment was seen on the SAC in 31% of the injured sample on day 1, 23% on day 2, and 9% (essentially at control levels) on day 7. Statistically defined abnormality on the SAC ranged from only 5–9% of the control group across all assessment points. Similarly, examination of the neuropsychological test results indicates that 23% of concussed players were impaired on two or more measures on postinjury day 2 and 17% were impaired on day 7. Eight percent of control participants were impaired on two or more measures at day 2, and 9% on day 7. Analyses of individual neuropsychological test scores on days 2 and 7 indicate that the largest percentage of injured participants obtained abnormal test scores on measures of delayed recall and recognition memory (HVLT Delayed Recall and Recognition), cognitive processing speed (Trails B, Symbol Digit Modalities Test), and verbal fluency (COWAT).

Impairments in postural stability, as defined by poorer performance on the BESS, were seen in 36% of the injured participants immediately following concussion, compared to 5% of the control group. Twenty-four percent of injured participants remained impaired on the BESS on day 2, compared to 9% by day 7 postinjury. Table 5 presents the rates of impairment in concussion and control groups for the composite battery of brief measures (GSC, BESS, SAC) at all time points, and the addition of neuropsychological testing at postinjury days 2 and 7.

Sensitivity and specificity values are provided for each measure in Table 6. The results indicate that the GSC provided the most sensitive (Se = .89) and specific (Sp = 1.00) measure of abnormality at the time of injury. The specificity values remained at 1.00 at each time point thereafter, indicating that none of the controls exhibited a significant increase in self-reported symptoms at any time point. Sensitivity values for the BESS were highest at the time of injury (Se = .34). Specificity values for this instrument ranged from .91 to .97 across the various time points. A similar pattern of data was obtained with the SAC, with a peak sensitivity value of .80 at the time of injury and specificity values ranging from .89 to .98 through day 7. The neuropsychological test battery classified injured participants with sensitivity values of .23 and .19 at days 2 and 7, respectively. Specificity values were .93 and .91.

Table 4. Neuropsychological test data for concussion and control groups at baseline and postinjury days 2 and 7

| | Baseline | | | | Day 2 | | | | Day 7 | | | |
| | Concussion | | Control | | Concussion | | Control | | Concussion | | Control | |
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
HVLT Immediate Memory	25.03	4.36	25.31	4.05	24.42	4.52	25.42	4.37	25.56	4.26	26.17	4.59
HVLT Delayed Recall	8.61	2.18	9.15	2.13	8.11	2.68	8.84	2.48	8.50	2.65	10.29	5.88
HVLT Recognition	22.60	1.97	22.94	1.26	22.20	1.88	22.95	1.52	22.50	1.54	22.77	2.39
Trail Making Test, Part B	64.42	22.22	57.30	18.69	59.70	21.20	51.33	16.67	53.30	20.49	42.05	15.36
SDMT	55.56	11.61	58.90	12.19	55.05	11.21	62.23	13.76	55.00	10.35	61.65	13.35
Stroop CW Trial	47.21	9.23	48.66	9.75	47.94	11.68	51.50	9.61	53.33	11.34	55.86	10.77
COWAT	40.46	12.36	37.15	10.61	40.17	10.99	40.22	10.12	41.45	10.96	42.44	10.07

HVLT = Hopkins Verbal Learning Test (Shapiro et al., 1999); Trail Making Test (Reitan & Wolfson, 1985); SDMT = Symbol Digit Modalities Test (Smith, 1991); Stroop CW Trial = Color Word Trial (Golden, 1978); COWAT = Controlled Oral Word Association Test (Benton et al., 1983).
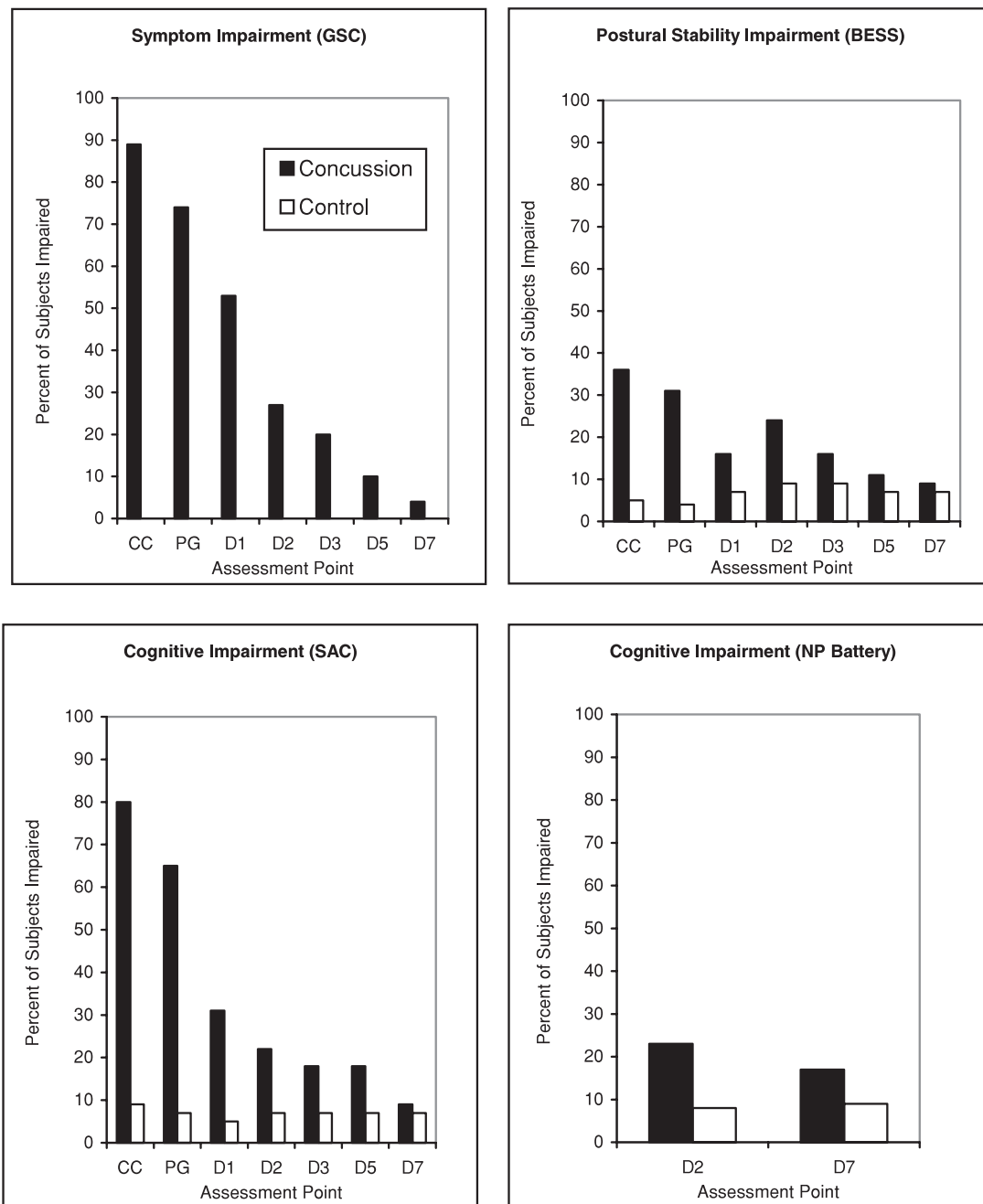
Fig. 1. Percentage of concussion and control participants classified as "impaired" from time of injury through day 7 on GSC, BESS, SAC and Neuropsychological Test Battery. GSC = Graded Symptom Checklist; BESS = Balance Error Scoring System; SAC = Standardized Assessment of Concussion; and NP Battery = Neuropsychological Test Battery. Assessment points: CC = time of concussion; PG = post-game/post-practice; D1 = postinjury day 1, D2 = postinjury day 2, etc.

An examination was made of sensitivity and specificity values for the entire battery of brief measures, defined as an abnormal score identified on either the GCS, BESS, or SAC. Again, sensitivity was highest at the time of injury, with 94% accuracy in classifying injured participants on the battery of brief measures. Specificity values ranged from .84 to .93 across the various time points. Inclusion of the neuropsychological test data barely increased the sensitivity of

the battery from .51 to .56 at day 2, with a decrease in specificity from .84 to .79. However, the neuropsychological test data more than doubled the sensitivity of the battery from .14 to .30 at day 7, accompanied by a modest increase in specificity from .79 to .86.

A small but clinically significant percentage of asymptomatic players continued to show impairment on the BESS, SAC, and neuropsychological testing on postinjury days 2

Table 5. Rates of impairment on battery of brief measures (GSC, BESS, SAC) and brief battery with neuropsychological testing

	GSC, BESS, SAC*				(GSC, BESS, SAC) + NP testing#			
	Concussion Abnormal		Control Abnormal		Concussion Abnormal		Control Abnormal	
	N	%	N	%	N	%	N	%
Time of Injury	89	95	6	11	—	—	—	—
Postgame/practice	81	86	6	11	—	—	—	—
Day 1	65	69	6	11	—	—	—	—
Day 2	48	51	9	16	53	56	12	21
Day 3	36	38	9	16	—	—	—	—
Day 5	24	26	7	13	—	—	—	—
Day 7	13	14	4	7	28	30	8	14

*"Abnormal" refers to impairment on any measure in brief battery (GSC, BESS, or SAC).
#"Abnormal" refers to impairment on any measure in brief battery or impairment on at least two neuropsychological measures.
GSC = Graded Symptom Checklist (Lovell & Collins, 1998); BESS = Balance Error Scoring System (Guskiewicz et al., 2001);
SAC = Standardized Assessment of Concussion (McCrea et al., 2000); NP Testing = Neuropsychological Test Battery.

and 7 (see Table 7). The highest percentage of players reporting to be completely symptom free were impaired on the BESS on day 2, with fewer impaired on cognitive measures. On day 7, the highest percentage of asymptomatic players were impaired on neuropsychological testing, while rates of impairment on the BESS and SAC for asymptomatic players were comparable to the control group.

DISCUSSION

The current study employed standard regression-based methods combined with a baseline and serial testing paradigm to statistically measure individual rates of cognitive and functional impairment in collegiate football players after sustaining a concussion. Nearly 90% of athletes with concussion reported an increase from their baseline rate of common concussive symptoms, and 80% showed a significant decline from baseline cognitive performance during the acute postinjury period. A significant percentage of concussed players also exhibited acute impairment on postural stability

testing. The percentage of injured athletes impaired on cognitive and balance testing gradually declined over the first several days postinjury, but still remained considerably higher than the comparative rate of statistically defined impairment in matched control participants. Less than 5% of injured athletes reported higher-than-base rate postconcussive symptoms by day 7 postinjury, and rates of impairment on brief screening measures of cognitive functioning and balance were comparable for injured and noninjured participants 1 week postinjury. Seventeen percent of concussed players continued to show impairment on neuropsychological testing 1 week after injury, compared to 9% of noninjured controls.

It is also important to note that functional recovery as measured by objective testing often lagged behind the resolution of postconcussive symptoms. Most concerning was our finding that 26% of concussed players who reported being symptom free, and who presumably were eager to return to competition 1 week after their injury, continued to show measurable impairment on standardized cognitive and

Table 6. Sensitivity (Sn) and specificity (Sp) for detecting impairment at postinjury time points

	Time of injury		Postgame		Day 1		Day 2		Day 3		Day 5		Day 7	
	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp
GSC	.89	1.00	.74	1.00	.53	1.00	.27	1.00	.20	1.00	.10	1.00	.04	1.00
BESS	.34	.91	.31	.96	.16	.93	.24	.91	.16	.91	.10	.93	.07	.95
SAC	.80	.91	.65	.93	.31	.95	.22	.89	.18	.93	.18	.93	.02	.98
Brief battery without NP testing	.94	.89	.86	.89	.69	.89	.51	.84	.38	.84	.26	.87	.14	.93
NP testing	—	—	—	—	—	—	.23	.93	—	—	—	—	.19	.91
Full battery with NP testing	—	—	—	—	—	—	.56	.79	—	—	—	—	.30	.86

Notes. Sensitivity values indicate the probability that a player originally injured continued to be correctly classified as "abnormal". Specificity (Sp) refers to the probability that a control participant will be correctly classified as "normal" using the same method.
Brief battery refers to GSC, BESS, and SAC. Full battery refers to brief battery plus neuropsychological testing.
GSC = Graded Symptom Checklist (Lovell & Collins, 1998); BESS = Balance Error Scoring System (Guskiewicz et al., 2001); SAC = Standardized Assessment of Concussion (McCrea et al., 2000); NP testing = neuropsychological test battery

Table 7. Percentage of asymptomatic participants after concussion (Sx-) and control participants classified as "Impaired" on postinjury days 2 and 7

	Day 2		Day 7	
	Sx—impaired (%) ($n = 68$)	Control impaired (%) ($n = 56$)	Sx—impaired (%) ($n = 85$)	Control impaired (%) ($n = 56$)
BESS	37	9	8	7
SAC	16	7	7	7
NP Battery	15	8	16	9

BESS = Balance Error Scoring System (Guskiewicz et al., 2001); SAC = Standardized Assessment of Concussion (McCrea et al., 2000); NP battery = neuropsychological test battery

balance testing, compared to 14% of controls. Neuropsychological testing was most sensitive to detecting cognitive impairment in otherwise asymptomatic players 1 week postinjury, with 16% impaired on at least two of the seven measures in the battery, compared to 9% of controls. These findings suggest that a period of cerebral dysfunction and vulnerability persists beyond resolution of subjective symptoms, which suggests the need for a thorough evaluation of all elements of recovery in making decisions about an athlete's readiness for a safe return to competition.

Advanced Methods for Measuring Recovery

This study used an alternative method of analysis to determine the proportion of injured participants exhibiting abnormally low scores on measures of symptoms, balance, and cognitive functioning, during the first week following concussion. We addressed this goal through the use of SRB methods developed in studies of epilepsy surgery outcomes, which is a different approach than those previously employed in studies examining recovery following sports concussion (Barr & McCrea, 2001; Barr, 2003; Hinton-Bayre et al., 1999). While the SRB method may be more difficult to use than the RCI, it offers several advantages in the research setting by taking into account differences in test scores at baseline, as well as controlling for other factors, including practice effects and regression to the mean. The SRB method also offers the potential to more closely investigate the influence of additional demographic variables on test–retest outcomes, which then allows the inclusion of an empirically derived demographic correction in the standard regression.

Our use of the two-tailed 90% confidence interval for detection of impairment is rather conservative, providing a 5% chance of false positives (i.e., misclassifying a normal participant as impaired) for any single measure. When looking at the full battery of measures (e.g., GSC, SAC, BESS, and Neuropsychological Testing Battery), the expected false positive rate increased to 27.3%, conservatively assuming complete orthogonality. Our control group data used to determine impairment in the concussion group fell within the expected range of false positives in nearly all instances. In a clinical setting, a less conservative confidence interval may be preferred to maximize the sensitivity of injury detection

at the cost of some degree of specificity, especially when the primary interest is protecting young sports participants from the ill effects of recurrent concussion (Barr & McCrea, 2001).

Implications for Sports Concussion Management

Prospective, empirical data on the rate and trajectory of individual recovery not only advance our understanding of the natural history of concussion, but also have more direct clinical implications specific to the management of sport-related head injuries. Defining "recovery" following sports concussion on the basis of findings from serial neuropsychological and other standardized testing has been the subject of great debate. Our previously reported group data analyses delineated a *typical* slope of recovery from concussion (McCrea et al., 2003), but did not provide a methodology for clinical decision making on an individual case basis. The more advanced SRB methods employed in this study revealed a higher percentage of players with concussion to still be impaired 1 week after injury than that previously suggested by group recovery data analyses. These data from the current study indicate that on most measures "return to baseline" does not necessarily reflect full recovery. The use of statistical models that empirically identify meaningful change while controlling for baseline test performance and practice effects on serial testing is essential to classifying impairment due to any condition. Further refinement and simplification of this approach is likely necessary to provide the sports medicine clinician with a user-friendly method to *measure* a player's level of recovery and readiness to return to play after concussion.

Relying primarily on a player's self-reported symptom recovery to guide injury management and return to play decision making after sport-related concussion is problematic. The sensitivity of symptom assessment based on a player's self-report is acceptable soon after injury, but diminishes after a few days. It is clear from the data on days 2 and 7, that even when players have experienced the resolution of symptoms as measured by the GSC, they may continue to exhibit deficits on objective measures of balance and cognitive functioning. A player's potential motivation to under-report symptoms in hopes of a more rapid return to

play also complicates matters. Those injured players who reported being symptom free but continued to exhibit mild impairment on standardized testing and formal neuropsychological testing 1 week postinjury may be at increased risk of recurrent or more severe injury (Guskiewicz et al., 2003) if returned to play based solely on their reported symptoms without objective recovery data to guide the clinician's decision making. Future studies may consider including a formal measure of symptom minimization response bias, and this forum may provide an interesting laboratory for neuropsychologists to compare this form of response bias to that of symptom exaggeration or malingering often encountered in other clinical or forensic settings.

A multidimensional model of sports concussion assessment is supported by our findings. Postconcussive symptoms, cognitive dysfunction, and postural instability are all common sequelae of concussion, but may manifest differently across individuals. Multidimensional assessment of all domains affected by concussion yields the greatest sensitivity in detecting injury and the best method for assessing postinjury recovery. The combined battery of brief screening instruments (total administration time < 15 min) that measured postconcussive symptoms, cognitive functioning, and balance provided a sensitive and specific means to detect and characterize concussion during the acute postinjury period, particularly on the sports sideline immediately after concussion. The SAC and neuropsychological test battery yielded similar rates of impairment on day 2, but neuropsychological testing provided a more sensitive measure of subtle cognitive dysfunction further out from injury, characterized by mild residual deficits in delayed recall memory, cognitive processing speed, and verbal fluency. Based on our findings, the combination of brief screening instruments appropriate for emergent use on the sideline and more extensive neuropsychological testing to assess recovery may be most appropriate for clinical use.

Given the frequency of impaired postural stability in the current study, balance testing is considered an important part of any assessment model to identify neurological dysfunction following concussion in athletes. Balance deficits, like postconcussion symptoms and cognitive impairment, may manifest and recover differently across athletes. As in the case with cognitive testing, obtaining baseline testing on clinical balance measures such as the BESS provides the greatest clinical accuracy in identifying the residual effects of concussion.

Inclusion of the neuropsychological test data only minimally increased the sensitivity and reduced the specificity of the battery of brief instruments within the first 2 days after concussion, but the addition of neuropsychological testing more than doubled the sensitivity and slightly increased the specificity of the brief battery at day 7. The relative value of neuropsychological testing is demonstrated by its superior sensitivity in detecting subtle neurocognitive impairment further out from injury, especially in players otherwise reporting to be completely symptom free, although there are some methodological considerations which may complicate the interpretation of these data. These data support the recommendation that neuropsychological testing be performed only when an injured participant reports being symptom free, which is also supported by practical and methodological considerations. Protocols that include neuropsychological testing at fixed intervals are appropriate for research intended to empirically track recovery, but are inappropriate in a clinical setting when a player is still symptomatic and should be withheld from competition regardless of the neuropsychological test results. Unnecessary serial testing, in addition to being burdensome to the athlete and medical staff, also introduces practice effects that may confound the interpretation of performance on subsequent testing (i.e., when a player reports being symptom free and is otherwise ready to return to play).

Study Limitations

It is possible that some players who sustained a concussion while enrolled in this study did not report their injury and were not identified by the team physician or certified athletic trainer. Recent reports (McCrea et al., 2004) suggest that the rate of concussion, even in well-controlled studies, is higher than that documented in the literature, due to a combination of players not recognizing the signs of injury or not reporting a concussion in order to continued uninterrupted sports participation.

The SRB method is subject to the reliability and validity of the main outcome measures used in this study, which have been supported by earlier studies on the effects of sports concussion. The availability of baseline data for injured and control participants on all concussion assessment measures also adds strength to the model by providing the most reliable means for detecting meaningful change in individual athletes. Although we incorporated most of the pencil-and-paper tests that have been routinely employed in sports concussion research, it is conceivable that a different test battery may have proven more or less sensitive to the lingering effects of concussion. There have as yet, however, been no direct comparisons of the sensitivity among pencil-and-paper or computer batteries being proposed for this purpose, and no conventional or computerized test battery has been established as the "gold standard".

The serial testing model employed in this study differs from previous SRB paradigms in epilepsy and cardiac surgery in the number of postevent testing sessions, particularly for the brief measures that were administered on several consecutive days after concussion. Further investigation is required to fully appreciate the effects of repeated neuropsychological testing, and the SRB approach at least provides one methodology do so. It is possible, for instance, that the control group is more readily able to benefit from practice on some neurocognitive measures, and this practice effect is likely to be attenuated in concussed players who are still encephalopathic, especially during the acute postinjury phase. Therefore, the normative group may not really be an appropriate control group to use for this purpose in predicting performance for any time point beyond

the first postinjury test session because the controls are likely to benefit from practice effects to a significantly greater extent than players who are suffering from the acute effects of concussion at the time of testing. This is likely to result in overclassifying injured players as being "impaired" even though they may have completely recovered. This issue cannot be overlooked, particularly when interpreting the comparative rates of impairment for injured and control participants beyond the acute phase on postinjury day 7. Future studies should consider utilizing a controlled, randomized design to compare cognitive and balance recovery patterns for concussed symptomatic athletes with concussed asymptomatic athletes to validate the current findings, and to substantiate our recommendation that testing be conducted once an athlete is reporting to be symptom free. Unfortunately, participant attrition disallowed application of our analysis to look at the lengthier course of recovery 90 days after concussion. Future studies should implement tightly controlled mechanisms for tracking injured players beyond the completion of the sports season to allow analysis of longer term outcomes.

Conclusion

In summary, sport-related concussion is characterized by a combination of measurable symptoms, cognitive dysfunction, and postural instability that follows a gradual course of recovery over several days in most cases. Relying solely on a player's self-reported recovery is not recommended, as a significant percentage of participants who report being completely symptom free after concussion may continue to show significant impairments on more sensitive standardized testing. Brief screening instruments are most sensitive and specific in accurately classifying injured and non-injured participants during the acute postinjury phase, while neuropsychological testing may be more sensitive to the subtle cognitive effects of concussion further out from injury, especially when a player otherwise reports being completely symptom free. Deferring neuropsychological testing until the injured player reports a full symptom recovery is recommended in a clinical setting, and for future studies investigating the incremental utility of neuropsychological testing in identifying the residual effects of sport-related concussion. Statistical models that empirically measure meaningful change on serial testing may provide a reliable benchmark for determining recovery and clinical decision making about an individual athlete's readiness to return to competition after sport-related concussion. These data provide an empirical base to be considered by sports governing bodies and expert panels responsible for developing practice guidelines for clinical decision making on return to competition after sport-related concussion.

REFERENCES

Aubry, M., Cantu, R., Dvorak, J., Graf-Baumann, T., Johnston, K., Kelly, J.P., Lovell, M., McCrory, P., Meeuwisse, W., & Schamasch, P. (2002). Concussion in sport group. Summary and agreement statement of the first International Conference on concussion in sport, Vienna 2001. *Physician and Sports Medicine*, *30*, 57–63.

Barr, W.B. (2002). Neuropsychological testing for assessment of treatment effects: Methodological issues. *CNS Spectrums*, *7*, 300–306.

Barr, W.B. (2003). Neuropsychological testing of high school athletes. Preliminary norms and test–retest indices. *Archives of Clinical Neuropsychology*, *18*, 91–101.

Barr, W.B. & McCrea, M. (2001). Sensitivity and specificity of standardized neurocognitive testing immediately following sports concussion. *Journal of the International Neuropsychological Society*, *7*, 693–702.

Benton, A.L., Hamsher, K., & Sivian, A.B. (1983). *Multilingual Aphasia Examination* (3rd ed.). Iowa City, IA: AJA Associates.

Cantu, R.C. (1986). Guidelines for return to contact sports after a cerebral concussion. *Physician and Sports Medicine*, *14*, 75–83.

Cantu, R.C. (1998). Return to play guidelines after a head injury. *Clinics in Sports Medicine*, *17*, 45–60.

Collins, M.W., Grindel, S.H., Lovell, M.R., Dede, D.E., Moser, D.J., Phalin, B.R., Nogle, S., Wasik, M., Cordry, D., Daugherty, K.M., Sears, S.F., Nicolette, G., Indelicato, P., & McKeag, D.B. (1999). Relationship between concussion and neuropsychological performance in college athletes. *Journal of the American Medical Association*, *282*, 964–970.

Collins, M.W., Lovell, M.R., & McKeag, D.B. (1999). Current issues in managing sports-related concussion. *Journal of the American Medical Association*, *282*, 2283–2285.

Collins, M.W., Field, M., Lovell, M.R., Iverson, G., Johnston, K.M., Maroon, J., & Fu, F.H. (2003). Relationship between postconcussion headache and neuropsychological test performance in high school athletes. *American Journal of Sport Medicine*, *31*, 168–173.

Colorado Medical Society (1991). *Report of the sports medicine committee: Guidelines for the management of concussion in sports*. Denver, CO: Colorado Medical Society.

Covassin, T., Swanik, C.B., & Sachs, M.L. (2003). Epidemiological considerations of concussions among intercollegiate athletes. *Applied Neuropsychology*, *10*, 12–22.

Dick, R. (2003). *National Collegiate Athletic Association (NCAA) Injury Survellience System 2002–2003*. Indianapolis, IN: NCAA.

Diggle, P.J., Liang, K.Y., & Zeger, S.L. (1994). *Analysis of longitudinal data*. Oxford, England: Oxford University Press.

Echemendia, R.J. & Julian, L.J. (2001). Mild traumatic brain injury in sports: Neuropsychology's contribution to a developing field. *Neuropsychological Reviews*, *11*, 69–88.

Echemendia, R.J., Putukian, M., Mackin, R.S., Julian, L., & Shoss, N. (2001). Neuropsychological test performance prior to and following sports-related mild traumatic brain injury. *Clinical Journal of Sport Medicine*, *11*, 23–31.

Erlanger, D., Kaushik, T., Cantu, R., Barth, J.T., Broshek, D.K., Freeman, J.R., & Webbe, F.M. (2003). Symptom-based assessment of the severity of a concussion. *Journal of Neurosurgery*, *98*, 477–484.

Golden, J.C. (1978). *Stroop Color and Word Test*. Chicago, Illinois: Stoelting Co.

Guskiewicz, K.M., McCrea, M., Marshall, S.W., Cantu, R.C., Randolph, C., Barr, W., Onate, J.A., & Kelly, J.P. (2003). Cumulative effects associated with recurrent concussion in collegiate football players: The NCAA Concussion Study. *Journal of the American Medical Association*, *290*, 2549–2555.

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*, 263–273.

Guskiewicz, K.M., Weaver, N.L., Padua, D.A., & Garrett, W.E. (2000). Epidemiology of concussion in collegiate and high school football players. *American Journal of Sport Medicine*, *28*, 642–650.

Hinton-Bayre, A.D., Geffen, G.M., Geffen, L.B., McFarland, K.A., & Friis, P. (1999). Concussion in contact sports: Reliable change indices of impairment and recovery. *Journal of Clinical and Experimental Neuropsychology*, *21*, 70–86.

Kelly, J. (1999). Traumatic brain injury and concussion in sports. *Journal of the American Medical Association*, *282*, 989–991.

Kelly, J.P. & Rosenberg, J.H. (1997). Diagnosis and management of concussion in sports. *Neurology*, *48*, 575–580.

Lovell, M.R. & Collins, M.W. (1998). Neuropsychological assessment of the college football player. *Journal of Head Trauma Rehabilitation*, *13*, 9–26.

Lovell, M.R., Collins, M.W., Iverson, G.L., Field, M., Maroon, J.C., Cantu, R., Podell, K., Powell, J.W., Belza, M., & Fu, F.H. (2003). Recovery from mild concussion in high school athletes. *Journal of Neurosurgery*, *98*, 296–301.

Macciocchi, S.N., Barth, J.T., Alves, W., Rimel, R.W., & Jane, J.A. (1996). Neuropsychological functioning and recovery after mild head injury in collegiate athletes. *Neurosurgery*, *39*, 510–514.

Maddocks, D. & Saling, M. (1996). Neuropsychological deficits following concussion. *Brain Injury*, *10*, 99–103.

McCrea, M. (2001). Standardized mental status testing on the sideline after sport-related concussion. *Journal of Athletic Training*, *36*, 274–279.

McCrea, M., Guskiewicz, K.M., Marshall, S., Barr, W.B., Randolph, C., Cantu, R.C, Onate, J.A., Yang, J., & Kelly, J.P. (2003). Acute effects and recovery time following concussion in collegiate football players: The NCAA Concussion Study. *Journal of the American Medical Association*, *290*, 2556–2563.

McCrea, M., Hammeke, T., Olsen, G., Leo, P., & Guskiewicz K. (2004). Unreported concussion in high school football players:

Implications for prevention. *Clinical Journal of Sport Medicine*, *14*, 13–17

McCrea, M., Kelly, J.P, Kluge, J., Ackley, B., & Randolph, C. (1997). Standardized assessment of concussion in football players. *Neurology*, *48*, 586–588.

McCrea, M., Kelly, J.P., Randolph, C., Cisler, R.A., & Berger, L. (2002). Immediate neurocognitive effects of concussion. *Neurosurgery*, *50*, 1032–1042.

McCrea, M., Kelly, J.P., Randolph, C., Kluge, J., Bartolic, E., Finn, G., & Baxter, B. (1998). Standardized assessment of concussion: On-site mental status evaluation of the athlete. *Journal of Head Trauma Rehabilitation*, *13*, 27–35.

McCrea, M., Randolph, C., & Kelly, J.P. (2000). *The Standardized Assessment of Concussion (SAC): Manual for Administration, Scoring and Interpretation* (2nd ed.). Waukesha, WI.

McSweeny, A.J., Naugle, R.I., & Chelune, G.J. (1993). T-scores for change: An illustration of a regression approach to depicting change in clinical neuropsychology. *The Clinical Neuropsychologist*, *7*, 300–312.

Powell, J.W. & Barber-Foss, K.D. (1999). Traumatic brain injury in high school athletes. *Journal of the American Medical Association*, *282*, 958–963.

Practice Parameter (1997). The management of concussion in sports (summary statement). Report of the quality standards committee. *Neurology*, *48*, 581–585.

Reitan, R.M. & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery*. Tucson, AZ: Neuropsychology Press.

Riemann, B.L. & Guskiewicz, K.M. (2000). Effects of mild head injury on postural stability as measured through clinical balance testing. *Journal Athletic Training*, *25*, 19–25.

Riemann, B.L., Guskiewicz, K.M., & Shields, E.W. (1999). Relationship between clinical and forceplate measures of postural stability. *Journal of Sports Rehabilitation*, *8*, 71–82.

Rubin, D.B. (1976). Inference and missing data. *Biometrika*, *63*, 581–592.

Rubin, D.B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, *91*, 473–489.

Schafer, J.L. (1997). *Analysis of incomplete multivariate data*. New York, NY: Chapman and Hall.

Shapiro, A.M., Benedict, R.H., Schretlen, D., & Brandt, J. (1999). Construct and concurrent validity of the Hopkins Verbal Learning Test–Revised. *Clinical Neuropsychologist*, *13*, 348–358.

Smith, A. (1991). *Symbol Digit Modalities Test*. Los Angeles, CA: Western Psychological Services.

SPSS (1999). *SPSS Advanced Models 10.0*. SPSS, Inc.

Temkin, N.R., Heaton, R.K., Grant, I., & Dikmen, S.S. (1999). Detecting significant change in neuropsychological test performance: A comparison of four models. *Journal of the International Neuropsychological Society*, *5*, 357–369.

Thurman, D., Branche, C., & Sniezek, J. (1998). The epidemiology of sports-related traumatic brain injuries in the United States: recent developments. *Journal of Head Trauma Rehabilitation*, *13*, 1–8.

Vastag, B. (2002). Football brain injuries draw increased scrutiny. *Journal of the American Medical Association*, *287*, 437–439.