Reproducibility of computer based neuropsychological testing among Norwegian elite football players

T M Straume-Naesheim, T E Andersen, R Bahr

Background: Head injuries account for 4–22% of all football injuries. The rate of brain injuries is difficult to assess, due to the problem of defining and grading concussion. Thus computerised testing programs for cognitive function have been developed.

Objective: To assess the reliability of a computerised neuropsychological test battery (CogSport) among Norwegian professional football players.

Methods: Norwegian professional football league players (90.3% participation) performed two consecutive baseline CogSport tests before the 2004 season. CogSport consists of seven different subtasks: simple reaction time (SRT), choice reaction time (ChRT), congruent reaction time (CgRT), monitoring (MON), one-back (OBK), matching (Match) and learning (Learn).

Results: There was a small but significant improvement from repeated testing for the reaction time measurements of all seven subtasks (SRT: 0.7%, ChRT: 0.4%, CgRT: 1.2%, MON: 1.3%, OBK: 2.7%, Match: 2.0%, Learn: 1.1%). The coefficient of variation (CV) ranged from 1.0% to 2.7%; corresponding intraclass correlation coefficients ranged from 0.45 (0.34 to 0.55) to 0.79 (0.74 to 0.84). The standard deviation data showed higher CVs, ranging from 3.7% (Learn) to 14.2% (SRT). Thus, the variance decreased with increasing complexity of the task. The accuracy data displayed uniformly high CV (10.4–12.2) and corresponding low intraclass correlation coefficient (0.14 (0.01 to 0.26) to 0.31 (0.19 to 0.42)).

Conclusion: The reproducibility for the mean reaction time measures was excellent, but less good for measures of accuracy and consistency. Consecutive testing revealed a slight learning effect from test 1 to test 2, and double baseline testing is recommended to minimise this effect.

Football is the only contact sport that exposes a large number of participants to purposeful use of the head for controlling and advancing the ball. Based on a series of cross-sectional studies using neurological examinations, neuropsychological tests, computed tomography scans and electroencephalographic examinations on active and older retired Norwegian football players, Tysvaer postulated that heading the ball could lead to chronic brain injuries as seen in boxing. Since then, several cross-sectional studies have indicated that football can cause sustained measurable brain impairment, although not all studies have reported such a relation.

Head injuries account for 4–22% of all football injuries with a reported incidence during matches of 1.7 injuries per 1000 player hours. However, this figure incorporates all types of head injury, including facial fractures, concussions, lacerations, and eye injuries. The incidence of concussion has been estimated to be 0.5 injuries per 1000 match hours but is difficult to assess, due to the problem of defining and grading concussions. When using the traditional diagnostic criteria for concussions, which require loss of consciousness or amnesia, only a fraction of these are recognised as concussions. Trauma to the neck and/or head that is sufficient to cause facial fractures or lacerations, will potentially also cause damage to the brain, although this is easily overlooked because of the more visible injuries. Although most athletes with head injuries recover uneventfully following a single episode of concussion, repetitive mild head trauma may be implicated in the development of cumulative cognitive deterioration. Accurate monitoring of symptom resolution and cognitive recovery is therefore important to ensure the athlete’s safety and indicate whether the player should return to play or not.

The change of paradigm in the diagnosis and management of concussion has evoked the need for new diagnostic instruments within sports related head injuries. One item in such tests is deterioration in cognitive test performance. In the sports arena, changes in cognition following a concussion injury are conventionally determined by administering a battery of neuropsychological tests during the pre-season to establish a baseline for comparison after an injury. In studies using such a design, any changes from baseline are considered to be a consequence of the concussion injury.

In the past decade, computerised cognitive function testing programs have been developed—for example, CogSport (CogState Ltd, Melbourne, Australia), ImPACT (ImPACT Inc., Pittsburgh, PA), ANAM (Automated Neuropsychological Assessment Metrics; developed by the US Department of Defense), CRI by HEADMINDER (concussion resolution index; Headminder Inc., New York). The conventional paper and pencil tests were designed primarily for assessment of cognitive dysfunction caused by neuronal or psychiatric disorders and not for the assessment of mild changes in cognitive function over time. Therefore, these tests often have poor psychometric properties for serial studies including a limited range of possible scores, floor and ceiling effect(s), learning effects, and poor test–retest reliability. Computerised testing using infinitely variable test paradigms may overcome these concerns.

Makdissi et al. compared the sensitivity of the CogSport test and conventional paper and pencil tests to detect cognitive changes following mild concussion in a cohort of elite players from the Australian Football League by comparing baseline tests with post-injury tests. Their data suggested that computerised tests may be particularly sensitive to the cognitive consequences of sports related concussions, and...
also that conventional neuropsychological tests do not show this sensitivity in athletes with mild concussion. Similar findings have recently been reported in studies on high school athletes with head injuries using ImPACT.17–19

Computer based cognitive tests have many advantages over paper and pencil tests that may allow them to detect subtle impairments such as those expected to occur in mildly concussed athletes.20 In general, repeated tests of healthy adults in different age groups have shown that computer based tests are reliable21,22; although there is a learning effect between test 1 and 2, this effect seems to decrease after the first two tests.22

The test properties of the CogSport test, a computer based neuropsychological evaluation tool widely used in football concussion management, have not been assessed by independent researchers, nor has it been examined among elite athletes. Therefore, the objective of this study was to evaluate the CogSport test by investigating the reproducibility of two consecutive baseline tests in a cohort of elite football players.

METHODS
We invited all the 14 clubs of the Norwegian professional male football league (Tippeligaen) with their A-squad contract players (about 300) to participate in the study; 289 players (90%) agreed to take part. Written informed consent was obtained from all participants and the project was approved by the Regional Ethics Committee for Southern Norway.

Neuropsychological testing
The teams meet for the La Manga Cup and pre-season training camp in February/March every year at the Norwegian Football Association training centre in La Manga, Spain. We conducted testing among 13 of the 14 teams of Tippeligaen at La Manga prior to the 2004 season in a test lab set up in the residential complex, Los Lomas II. Trained personnel administered and supervised the neuropsychological testing, and the tests were completed by the players in groups of three in the same quiet room to allow efficient data collection. The last team was tested at its home field in Norway two weeks later under similar standardised conditions. There is no time difference between Spain and Norway and the testing was performed at the same time of day with the same person instructing and supervising the test for each team.

We used the computer based neuropsychological test CogSport (versions 2.2.0 and 2.2.1). Norwegian speaking players were tested with the Norwegian language version of the test, where instructions for each subtask were in Norwegian, and the rest of the players used the English language version. The test has been described in detail elsewhere.13,23,24 The stimulus for all tasks consists of playing cards and responses are given using the keyboard. The d key indicates “no” and the k key “yes” (vice versa for left handed players). These are the only keys used throughout the whole test.

The CogSport test battery includes seven subtasks testing different cognitive brain functions (table 1). All subtasks

<table>
<thead>
<tr>
<th>Test</th>
<th>Description</th>
<th>Cognitive function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple reaction time</td>
<td>A single card was presented face down in the centre of the screen. The player instructed to press “yes” whenever the card turned face-up. Fifteen trials were presented and the test was repeated three times, at the beginning, in the middle, and at the end of the battery. All other tests were presented just once.</td>
<td>Motor function</td>
</tr>
<tr>
<td>Choice reaction time</td>
<td>This test used the same stimuli as above, but the player was now instructed to indicate whether the card was red by pressing “yes” or “no”.</td>
<td>Decision making</td>
</tr>
<tr>
<td>Congruent reaction time</td>
<td>Two cards were presented and the player had to indicate if they were same colour or not by pressing “yes” or “no”</td>
<td>Simple attention</td>
</tr>
<tr>
<td>Monitoring</td>
<td>Five cards moved simultaneously on the screen and the player was instructed to press “yes” as soon as one card moved outside a predefined area.</td>
<td>Divided attention</td>
</tr>
<tr>
<td>One-back</td>
<td>The player was instructed to indicate whether a new card was identical to or different from the last by pressing “yes” or “no”.</td>
<td>Working memory</td>
</tr>
<tr>
<td>Matching</td>
<td>Six card pairs were presented at the top of the screen and the player had to decide whether a pair presented at the bottom of the screen matched any of the above.</td>
<td>Complex attention</td>
</tr>
<tr>
<td>Learning</td>
<td>Incidental learning: this followed immediately after the matching task, and was identical to that task except that the six pairs were turned face down.</td>
<td>Learning and memory</td>
</tr>
<tr>
<td>Associate learning</td>
<td>Similar to the matching task, however, the pairs were turned face down when the player correctly indicated a matching pair presented at the bottom of the screen.</td>
<td></td>
</tr>
</tbody>
</table>

Data are the mean difference of the log_{10} of the reaction times (ms) and arcsine of the per cent correct responses for the accuracy data for test 1 and 2 with the corresponding 95% confidence intervals (CI) in parenthesis. Test results were compared using paired Student’s t test.

Table 1 Description of the seven CogSport subtasks and their assumed corresponding cognitive function

Table 2 Comparison between the results from test 1 and test 2 for the main CogSport outcome measures

<table>
<thead>
<tr>
<th>Subtask</th>
<th>Mean difference (95% CI)</th>
<th>Improvement (%)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple reaction time</td>
<td>-0.016 (-0.020 to -0.012)</td>
<td>-0.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Choice reaction time</td>
<td>-0.009 (-0.013 to -0.003)</td>
<td>-0.4</td>
<td>0.004</td>
</tr>
<tr>
<td>Congruent reaction time</td>
<td>-0.034 (-0.040 to -0.027)</td>
<td>-1.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Monitoring, reaction time</td>
<td>-0.031 (-0.043 to -0.0194)</td>
<td>-1.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>One-back, reaction time</td>
<td>-0.074 (-0.083 to -0.066)</td>
<td>-2.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Matching, reaction time</td>
<td>-0.062 (-0.077 to -0.0519)</td>
<td>-2.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Learning, reaction time</td>
<td>-0.032 (-0.039 to -0.025)</td>
<td>-1.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>One-back, accuracy</td>
<td>0.066 (0.036 to 0.096)</td>
<td>5.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Matching, accuracy</td>
<td>-0.014 (-0.044 to 0.016)</td>
<td>-1.1</td>
<td>0.300</td>
</tr>
<tr>
<td>Learning, accuracy</td>
<td>0.039 (0.019 to 0.058)</td>
<td>3.8</td>
<td>0.001</td>
</tr>
</tbody>
</table>
include between 15 and 40 trials, and the data are reported by the CogSport program as mean reaction times with corresponding standard deviations for all subtasks, accompanied by accuracy data for all tasks except simple reaction time and monitoring. Anticipatory responses (reaction times <100 ms) and abnormally slow responses (reaction times >3500 ms) are recorded as errors and excluded from the analyses. Accuracy data are calculated as the number of true positive responses divided by the number of trials.

The computer program sends a test report by e-mail to the test supervisor with basic analyses of the result. In addition, the test report includes an estimate of whether the player’s performance meets the minimum requirements of the test with regard to alertness throughout the test and the plausibility of whether they understood the instructions or not. This built-in decision is based on the variability of performance on simple reaction time and a threshold value for accuracy on the three final tasks. If a player has more than 40 incorrect responses on one task, the test is stopped.

We used SPSS version 11 for the statistical analyses and its two way random single measure model for calculating the intraclass correlation coefficients. Paired Student’s t tests were used to investigate significant differences and any directional trends between the groups.

RESULTS

Demographics

Of the 289 players (96.3%) who agreed to participate in the study, 18 did not report for testing, leaving us with 271 (90.3%) players who underwent two consecutive neuropsychological tests. However, due to technical problems with some tests (unrelated to test performance), the number of players with dual tests decreased to 247. In addition, 15 tests did not fulfil the minimum requirements set by the computer program and therefore could not be included in the analyses. Thus, a total of 232 players (83% Norwegians, 8% Scandinavians (with no problems in understanding Norwegian), and 9% from other countries (mainly European)) were included in the study. The mean (SD) age of the investigated group was 25.7 (4.6) years (range 17–35); 87.5% were right and 12.5% left handed; 62.9% had completed secondary education (that is, high school), and 36.6% had a tertiary level of education (that is, college or beyond). The demographic characteristics of excluded group did not differ significantly in any way from the included group.

Reproducibility

There was a significant improvement in the CogSport subtasks from test 1 to test 2, ranging between 0.4% and 2.7% for the log10 transformed reaction time measures (table 2, fig 1). The improvement in reaction time was slightly higher for the more complex tasks compared with the simpler ones (table 2). The accuracy data for the more complex subtasks (one-back and learning) also indicated a better performance (higher percentage of correct responses) in test 2, except for matching (table 2, fig 2).

Data analysis

Reliability and correlation studies of CogSport on young adults recommend the mean reaction time for all seven subtasks and accuracy data from the three final tasks (one-back, matching, learning) as the main outcome measures. Our data analysis therefore focused on these 10 measures. Before all calculations, the mean reaction times and standard deviation data were log10 transformed and the accuracy data were arcsine transformed to obtain a more normal distribution.

Reproducibility analyses were performed using the method error (ME), calculated as the standard deviation (SD) of the mean difference between test 1 and 2 divided by the square root of the number of tests performed: ME = SDmean diff/[2]. From the ME we calculated the coefficient of variation (CV), which quantifies the variation between each measurement as a percentage of the joined mean: CV = ME/(Xmean + Xmean)/2. These calculations were done for all outcome measures supplied by the test. We also calculated the intraclass correlation coefficient for the same measures. The intraclass correlation coefficient is defined as the ratio of the “true” variance, or the variance between subjects (S2b), relative to the total variance given by the variance between subjects adding the variance within subjects (S2w). The intraclass correlation coefficient ranges from 0 to 1, and from the equation it its simplest form (S2b / (S2b + S2w)), we see that when the variation within the subjects (that is, a player’s test score on two consecutive tests) moves towards 0, the intraclass correlation coefficient approaches 1 indicating good reproducibility.

We used SPSS version 11 for the statistical analyses and its method error (ME), calculated as the standard deviation (SD) of the mean difference between test 1 and 2 divided by the square root of the number of tests performed: ME = SDmean diff/[2]. From the ME we calculated the coefficient of variation (CV), which quantifies the variation between each measurement as a percentage of the joined mean: CV = ME/(Xmean + Xmean)/2. These calculations were done for all outcome measures supplied by the test. We also calculated the intraclass correlation coefficient for the same measures. The intraclass correlation coefficient is defined as the ratio of the “true” variance, or the variance between subjects (S2b), relative to the total variance given by the variance between subjects adding the variance within subjects (S2w). The intraclass correlation coefficient ranges from 0 to 1, and from the equation it its simplest form (S2b / (S2b + S2w)), we see that when the variation within the subjects (that is, a player’s test score on two consecutive tests) moves towards 0, the intraclass correlation coefficient approaches 1 indicating good reproducibility.

We used SPSS version 11 for the statistical analyses and its method error (ME), calculated as the standard deviation (SD) of the mean difference between test 1 and 2 divided by the square root of the number of tests performed: ME = SDmean diff/[2]. From the ME we calculated the coefficient of variation (CV), which quantifies the variation between each measurement as a percentage of the joined mean: CV = ME/(Xmean + Xmean)/2. These calculations were done for all outcome measures supplied by the test. We also calculated the intraclass correlation coefficient for the same measures. The intraclass correlation coefficient is defined as the ratio of the “true” variance, or the variance between subjects (S2b), relative to the total variance given by the variance between subjects adding the variance within subjects (S2w). The intraclass correlation coefficient ranges from 0 to 1, and from the equation it its simplest form (S2b / (S2b + S2w)), we see that when the variation within the subjects (that is, a player’s test score on two consecutive tests) moves towards 0, the intraclass correlation coefficient approaches 1 indicating good reproducibility.

We used SPSS version 11 for the statistical analyses and its two way random single measure model for calculating the intraclass correlation coefficients. Paired Student’s t tests were used to investigate significant differences and any directional trends between the groups.
DISCUSSION

This is the first study to examine the test properties of a computer-based neuropsychological test battery performed by an independent research group. The main finding was that the day to day reproducibility for the mean reaction time measures was excellent in a large cohort of professional football players, but that the accuracy and consistency measures were less reliable. We also observed a slight learning effect from the first to the second test. Thus our results are in accordance with those of recent studies examining the reliability of computerised neuropsychological tests among healthy young adults and elderly people. Collie et al assessed the reliability of CogSport by serial testing at a one hour and a one week interval 60 young volunteers recruited through advertisements around university campuses in Melbourne, Australia. Elite athletes are select individuals, who may differ from this group in many different ways, including background characteristics such as education level and socioeconomic status. However, even more important is that superior neurocognitive skills may be one of the selection criteria to become an elite footballer. In fact, a closer look at the reaction time data of Collie et al’s 60 volunteers reveals that they were considerably slower than the footballers on all subtasks. The reproducibility of the CogSport test on elite athletes has not been thoroughly investigated before. The apparent difference between regular
controls and elite athletes illustrates the need to develop appropriate reference data in populations of elite athletes, and supports the practice of individual baseline testing in the elite as a basis for the management of concussion.

The CV ranged from 1.0% to 2.7% for the mean reaction time measures and all values under 5% must be considered as good. Collie et al.,\textsuperscript{20} in their study on 60 healthy non-athletic young volunteers, reported intraclass correlation coefficients for the reaction time measurement higher than 0.69 for all of the four tested subtasks. Except for simple reaction time, the results were similar when comparing the test–retest results with both the one hour and the one week interval between the tests. The intraclass correlation coefficient for the mean reaction time measures from our material were within the same range. Reaction time measures have been shown to provide the most sensitive index of cognitive changes following a head injury,\textsuperscript{27} which in part is due to the fact that they are highly reproducible, as indicated in both our study and previous studies on other study populations.\textsuperscript{16 20 22 28} In contrast, of the consistency measures only the standard deviations for the most complex tasks (matching and learning) were within this limit. Although there was a uniform trend of less variation on test 2, the reproducibility data imply that these measures are unlikely to be helpful for follow up evaluations. The simpler tasks were the least consistent and one may speculate that the lack of complexity in these tasks causes the player to lose focus during the task. In the Cogsport testing program, simple reaction time testing is repeated three times during the session, which may exaggerate this effect.

In our study, the accuracy data showed inadequate reproducibility and the highest improvement from test 1 to test 2. The ceiling effect found on both one-back and matching may also make these less suitable as outcome measures, even with a dual baseline setting. Previous analyses using this computerised battery have shown ceiling effects for all accuracy data except matching and learning,\textsuperscript{25} but our results indicate that this is also the case for matching.

It should be noted that intraclass correlation coefficients must be interpreted with caution. From the simplified equation for the intraclass correlation coefficient ($S_{ww}/(S_{b} + S_{ww})$) it is evident that data of a homogeneous group (that is, where the between-subjects variability ($S_{b}$) is small compared with the within-subjects variability ($S_{ww}$)) will produce a poorer intraclass correlation coefficient than data of a heterogenous group (that is, with high between-subject variability with respect to the within-subject variability), even if within-subject variability is exactly the same for the two groups. It is therefore recommended not to compare directly the intraclass correlation coefficients from different study populations without knowing the variance within the tested groups.\textsuperscript{10} We have therefore, as recommended,\textsuperscript{25} also presented the test–retest coefficients of variation, which are independent of test result range and therefore can be compared directly between studies. It should be noted that, compared with the performance data reported by Collie et al.,\textsuperscript{20} our footballers displayed both faster mean reaction times and a more homogeneous performance. When this is taken into consideration, a comparison of the test–retest intraclass correlation coefficients with Collie et al indicates that the reproducibility of the mean reaction time measures may be even better among elite footballers than non-athletic controls. In a one year follow up of 84 elite Australian Rules footballers, the test–retest coefficients of variation were not reported.\textsuperscript{31}

In our group of professional football players, there was a significant improvement from test 1 to test 2 for the mean reaction time measures on all subtasks of CogSport. Collie et al found a similar practise effect when a group of elderly volunteers (mean age 64 (8) years) performed four consecutive CogSport tests in three hours.\textsuperscript{27} Whereas our professional football players tended to display a more pronounced practise effect when the tasks became more complicated, Collie et al’s elderly volunteers showed an opposite trend. More relevant is a comparison with elite Australian Rules footballers, and, as mentioned above, 84 of these were tested after an injury-free season (the exact timeframe was not stated) without displaying any significant differences in performance since baseline for any of the subtasks of CogSport (for the final two tasks, matching and learning, accuracy data were presented instead of mean reaction times). A practise test was conducted before the baseline test, but it is not clear if this was done for the follow-up as well (either in full or shortened).

Since we performed only two tests, we are not in a position to say whether the practise effect will decrease with further testing. However, Falleti et al followed 26 young volunteers who performed four different baseline CogSport tests on three different days, where the first two tests were performed on the same day with a two hour break in between, and the first was discarded. In the three remaining baseline tests (time intervals not stated) there were no differences in performance on reaction time or accuracy measures.\textsuperscript{25}

Another aspect of the mean reaction time measurements, which became evident on Bland–Altman plots, was that the improvement from test 1 to 2 was not evenly distributed. The players with the slowest mean reaction times improved the most, and on some subtasks those with the fastest mean reaction times were actually slower on test 2. Such regression towards the mean has also been described by Erlanger on simple and choice reaction time measures from a similar computerised neuropsychological test package from HEADMINDER.\textsuperscript{12}

Due to the practise effect, we agree with previous studies conducted on other populations that the test requires a dual baseline, where the first test is discarded.\textsuperscript{22 23} Whether this procedure should be used in follow up testing if more than a couple of weeks have passed since the baseline testing, needs further investigation. One problem with a dual baseline tests is that the test becomes more time consuming and there is a risk that the player will lose their focus. A large study of patient with head injuries found that effort explained 53% of...
the variance in neuropsychological test performance (in comparison, educational level accounted for only 11% and age only 4%). It has to be noted that the group used the old paper and pencil test, which implies a different testing setting. The results can therefore not be transferred directly to computerised testing. Nevertheless, the issue of including some kind of effort measure when conducting a neuropsychological test was recently stressed at the Second International Symposium on Concussion in Sport in Prague. In conclusion, the reproducibility for the mean reaction time measures was excellent in the cohort on professional footballers included in the present study. However, the accuracy and consistency measures were less reliable, and may therefore be less sensitive as outcome measures in post-concussion management. Consecutive testing revealed a slight learning effect from test 1 to test 2, and dual baseline testing with rejection of the first test is recommended to minimise this effect.

ACKNOWLEDGEMENTS
This study was paid for by a grant from FIFA. In addition, financial support came from the Oslo Sports Trauma Research Center, which has been established at the Norwegian University of Sport and Physical Education through generous grants from the Eastern Norway Regional Health Authority, the Royal Norwegian Ministry of Culture and Church Affairs, the Norwegian Olympic Committee and Confederation of Sports, Norsk Tipping AS, and Pfizer AS. A special thanks to Jiri Dvorak and Astrid Junge from the Oslo Sports Trauma and Confederation of Sports, Norsk Tipping AS, and Pfizer AS. Support came from the Oslo Sports Trauma Research Center, which has been established at the Norwegian University of Sport and Physical Education through generous grants from the Eastern Norway Regional Health Authority, the Royal Norwegian Ministry of Culture and Church Affairs, the Norwegian Olympic Committee and Confederation of Sports, Norsk Tipping AS, and Pfizer AS. A special thanks to Jiri Dvorak and Astrid Junge from the FIFA-Medical Assessment and Research Centre (F-MARC) for their collaboration on developing the study protocol and Alex Collie for technical support. The authors thank Jostein and Grete Steene-Johannessen for test supervision, Ingar Holme and Lars Bo Anderson for statistical assistance, and the players, team physicians, physiotherapists, and coaches for their cooperation.

Authors’ affiliations
T M Strømme-Naesheim, T E Andersen, Oslo Sports Trauma and Research Center, Oslo, Norway
R Bahr, Norwegian University of Sport and Physical Education, Oslo, Norway

Competing interests: none declared

REFERENCES
Reproducibility of computer based neuropsychological testing among Norwegian elite football players

T M Straume-Naesheim, T E Andersen and R Bahr

doi: 10.1136/bjsm.2005.019620

Updated information and services can be found at:
http://bjsm.bmj.com/content/39/suppl_1/i64

These include:

References
This article cites 30 articles, 9 of which you can access for free at:
http://bjsm.bmj.com/content/39/suppl_1/i64#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/