Cognitive complaints in patients after whiplash injury: the impact of malingering

B Schmand, J Lindeboom, S Schagen, R Heijt, T Koene, H L Hamburger

Abstract

Objectives—The validity of memory and concentration complaints that are often reported after a whiplash trauma is controversial. The prevalence of malingering or underperformance in post-whiplash patients, and its impact on their cognitive test results were studied.

Methods—The Amsterdam short term memory (ASTM) test, a recently developed malingering test, was used as well as a series of conventional memory and concentration tests. The study sample was a highly selected group of patients, who were examined either as part of a litigation procedure (n=36) or in the normal routine of an outpatient clinic (n=72).

Results—The prevalence of underperformance, as defined by a positive score on the malingering test, was 61% (95% CI: 45-77) in the context of litigation, and 29% (95% CI: 18-40) in the outpatient clinic (p=0.003). Furthermore, the scores on the memory and concentration test of malingering post-whiplash patients (n=43) and non-malingering post-whiplash patients (n=65) were compared with the scores of patients with closed head injury (n=20) and normal controls (n=46). The malingering post-whiplash patients scored as low as the patients with closed head injury on most tests.

Conclusions—The prevalence of malingering or cognitive underperformance in late post-whiplash patients is substantial, particularly in litigation contexts. It is not warranted to explain the mild cognitive disorders of whiplash patients in terms of brain damage, as some authors have done. The cognitive complaints of non-malingering post-whiplash patients are more likely a result of chronic pain, chronic fatigue, or depression.

Keywords: whiplash injury; neuropsychological tests; cognition disorders; malingering

The status of cognitive complaints in patients with a chronic pain syndrome after whiplash injury of the neck is uncertain. Apart from pain in the head, neck and shoulders, and increased fatigability, these patients often complain of forgetfulness and poor concentration. Although these studies may be criticised on methodological grounds, their findings have been ascribed either to damage of basal frontal and brain stem structures, to intracranial vasomotor dysregulation, to adverse effects of medication, or to the debilitating influence of chronic pain. Some neurologists, on the other hand, emphasise the lack of evidence of physical damage, and suggest that many post-whiplash patients exaggerate or even simulate their complaints, especially when litigation is involved. These accusations are strengthened by recent reports of regional differences with respect to incidence of the syndrome and insurance claim behaviour.

In view of these suggestions, neuropsychologists should attempt to rule out the possibility of malingering when studying post-whiplash patients. Until now, this issue has been neglected, which is worrying as it is notoriously difficult to detect malingering from the results of common neuropsychological tests; it is quite easy to fabricate an abnormal, yet credible test profile, even for people who are not initiated in neuropsychology. The issue is especially important when neuropsychological tests are applied in litigation procedures. Moreover, in the context of whiplash, we need not only to be able to detect plain malingering (intentionally faking of symptoms), but also more subtle forms of suboptimal cognitive performance—for example, of underperformance as a consequence of the patients’ need to get recognition for their complaints.

The aim of the present study was twofold. Firstly, we wanted to investigate the extent of malingering in post-whiplash patients. Secondly, if we could separate malingers from non-malingering post-whiplash patients, we would be in a better position to appreciate whether, and if so to what extent, these patients are cognitively affected. Therefore, we applied a test for the detection of malingering (and other types of suboptimal performance) to a group of post-whiplash patients as part of a battery of memory and concentration tests. All patients were seen either in an outpatient clinic or during a litigation procedure. We expected that the prevalence of underperformance would be higher in the litigation than in the non-litigation context. Next, to study the impact of malingering on the cognitive profile, the post-whiplash patients were divided into subgroups of malingering and non-malingering patients, and their scores on conventional memory and concentration tests were compared with scores of patients with closed head injury, and normal controls. We expected that malingers would perform worse on these tests than non-malingering post-whiplash patients, and that the scores of the last group would lie somewhere between those of patients with closed head injury and controls.
Table 1 Demographic characteristics of non-malingering and malingering patients after whiplash, patients after closed head injury, and normal controls

<table>
<thead>
<tr>
<th></th>
<th>Whiplash non-malingering (n=65)</th>
<th>Whiplash malingering (n=41)</th>
<th>Closed head injury (n=20)</th>
<th>Normal controls (n=46)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/female</td>
<td>31/34</td>
<td>14/29</td>
<td>12/8</td>
<td>17/29</td>
<td>0.14 (χ²)</td>
</tr>
<tr>
<td>Age (mean (SD))</td>
<td>37.2 (10.3)</td>
<td>41.0 (11.1)</td>
<td>37.7 (14.3)</td>
<td>33.9 (14.3)</td>
<td>0.06 (ANOVA)</td>
</tr>
<tr>
<td>Education (mean (SD))</td>
<td>5.2 (1.1)</td>
<td>4.9 (1.3)</td>
<td>5.2 (1.2)</td>
<td>4.7 (1.4)</td>
<td>0.12 (ANOVA)</td>
</tr>
<tr>
<td>ASTMIQ (mean (SD))</td>
<td>98.8 (3.9)</td>
<td>94.8 (13.1)</td>
<td>97.3 (4.4)</td>
<td>96.4 (14.2)</td>
<td>0.59 (ANOVA)</td>
</tr>
</tbody>
</table>

*Level of education is scored on a 7 point scale: 1 = primary school not finished; 2 = primary school (6 y); 5 = intermediate professional school (about 10 y); 7 = university (about 17 y). ASTM = Amsterdam short term memory test; DART = Dutch adult reading test.

Methods

SUBJECTS

The study was conducted in the psychological departments of a university hospital and a teaching hospital in Amsterdam. We examined 108 consecutive neurological outpatients with a late post-whiplash syndrome, 20 patients with closed head injury, and 46 controls.

The post-whiplash patients were referred for neuropsychological evaluation because of memory or concentration problems, either as part of a litigation procedure (n=36), or as part of the neuropsychological evaluation in the participating outpatient clinics (n=72). Although the second group was referred within the health care system, only 12 patients were not involved in a damage claim or in a workmen’s compensation claim. All patients had a cervical acceleration-deceleration trauma, mostly in traffic accidents (67% rear end collisions). Exclusion criteria were any kind of head injury, loss of consciousness, radiographic findings indicating fractures or dislocations in the upper spinal cord, psychiatric disorders, and alcohol or substance misuse. All patients complained of pain in the neck, head, shoulders, or arms within the first two days after the accident. The patients satisfied the Quebec classification criteria of whiplash-associated disorders, grades I–III. The mean interval since the accident was 24 (SD 22) months (range three months to 12 years). This interval was longer in litigation subjects (32 (SD 15) months) than in clinical routine subjects (20 (SD 23) months; p<0.0001, Mann-Whitney test). Thirty one per cent used analgesics, 7% benzodiazepines, 2% antidepressants, and 20% some combination of these medications. The remaining 40% used no medication. Forty four per cent were not working and received either a permanent disability benefit (19%), or a temporary sickness benefit (25%). Twenty five per cent had made some arrangement to reduce their work load (for example, part time, or reduced number of tasks), and the remaining 31% continued working as before the accident. Mean age of the post-whiplash patients was 38.9 (SD 10.7) years. Litigation subjects were significantly older (42.4 (SD 10.8) years, than non-litigation subjects (36.8 (SD 10.3) years, p=0.01, t test). There were no significant sub-group differences in sex or educational level.

Data from two control groups were used: 20 patients with memory and concentration disorders due to severe closed head injury, and 46 healthy control subjects. The patients with closed head injury had been included in a previous validation study of the malingering test. They had been admitted to hospital after closed head injury with loss of consciousness ranging from 15 minutes to 13 weeks (median two weeks). Their mean Glasgow coma scale score at admission was 9.3 (SD 3.5); mean duration of post-traumatic amnesia was 35 (SD 36) days. All patients with closed head injury had subjective memory complaints at the time of testing, confirmed by a relative, and corroborated by abnormal memory test scores. The patients with closed head injury in this study were affected by their cognitive disorders to such an extent that they had been unable to resume their work. Mean interval since injury was 3.8 years.

The normal control subjects were chosen from a panel of research volunteers consisting mainly of hospital personnel and their friends and relatives. This panel represents a wide range of professions. Subjects with psychiatric disorders or alcohol or other substance misuse were excluded from all groups. Table 1 shows their demographic characteristics.

PROCEDURE

The post-whiplash patients underwent a standard neuropsychological examination, including an interview of about 45 minutes in which a psychological history was taken, followed by a cognitive test programme lasting two to three hours (with a 10–15 minute break). The patients with closed head injury and controls received a test programme of similar duration.

TESTS

The test programmes differed in the two participating centres and were tailored to the individual post-whiplash patients. However, an overlapping core battery consisted of the following tests.

The Amsterdam short term memory test (ASTM) for the detection of malingering. This test has been constructed using a “symptom validity testing” paradigm. The test consists of 30 items and two practice items. In each item the subject is presented with five printed words from the same semantic category (for example, Holland, France, Belgium, England, Germany), which he has to read aloud and try to remember. Then he is distracted with a simple written addition or subtraction task (for example, 27+15=), which he has to solve mentally. Finally, five words from the same semantic category as before are presented. The subject has to indicate the three words that were also presented in the first series (for example: Russia, France, Germany, Greece,
Whiplash and malingering

The army individual test battery was used.20 The trail making test part A and part B from Thescoreisthetimetocompletioninseconds.

is a measure of perceptual interference, re-

meaningofeachcolourword. TheStrooptest

The colour of the ink is di-

glesofcard2,andnamethecolourofthe-

wordsofcard1,namethecolouroftherectan-

glewordsofcard3, as quickly as possible. The task is to read aloud the colour printed colour words, 100 coloured rectangles, 100 colour printed colour words respectively. The task is to write digits under nine arbitrary symbols as quickly as possible during 90 seconds. At the top of the test sheet is a printed key that pairs each symbol with a digit. The substitution test is considered to be a test of visual scanning, manual response speed, visuomotor coordination, and sustained attention.19 Scores are age corrected t scores (population mean 50 (SD 10)).

Dutch adult reading test (DART)14 15

The Dutch adult reading test (DART) is the Dutch adaptation of the national adult reading test (NART),16 a short reading test for the estimation of premorbid verbal IQ (population mean=100 (SD 15)).

Verbal fluency17

The verbal fluency test consists of naming animals, and professions, during one minute each. Raw scores are transformed into age corrected t scores (population mean 50 (SD 10)).

Symbol digit modalities test

The symbol digit modalities test or substitution test from the Dutch version of the Wechsler adult intelligence scale (WAIS) was used.18 The task is to write digits under nine arbitrary symbols as quickly as possible during 90 seconds. At the top of the test sheet is a printed key that pairs each symbol with a digit. The substitution test is considered to be a test of visual scanning, manual response speed, visuomotor coordination, and sustained attention.19 Scores are age corrected t scores (population mean 50 (SD 10)).

Trail making test

The trail making test part A and part B from the army individual test battery was used.20 The task is to connect numbers (part A), and to connect numbers alternating with letters (part B) on a sheet of paper. This is a test of visual scanning, visuomotor and conceptual tracking, mental flexibility, and motor speed.17 The score is the time to completion in seconds.

Stroop test21

This test consists of three cards with 100 black printed colour words, 100 coloured rectangles, and 100 colour printed colour words respectively. The task is to read aloud the colour words of card 1, name the colour of the rectangles of card 2, and name the colour of ink of the colour words of card 3, as quickly as possible. The colour of the ink is different from the meaning of each colour word. The Stroop test is a measure of perceptual interference, response inhibition, and selective attention.19 The score is the time to completion in seconds.

Auditory verbal learning test (AVLT)22

The subject has to memorise a series of 15 unrelated concrete nouns in five learning trials. After an interval of 20 minutes he has to recall the words, followed by recognition of the 15 items between 15 distractor words. Raw scores are used. The (theoretically) maximum learning score is 75, maximum recall score is 15, maximum recognition score is 30.

Logical memory (story recall) of the Rivermead behavioural memory test23

A 21 item news message is read to the subject, who repeats as many items as he can remember. After a 15 minute interval he is asked to recall the message again. The score is the number of items recalled.

More elaborate descriptions and references for most of these tests are given by Lezak.16

Analyses

Firstly, the frequency of subnormal scores (<86) on the ASTM malingering test was established in the post-whiplash patients. The prevalences of underperformance in litigation and non-litigation subjects were compared by $\chi^2$ test (with Yates’ correction). Secondly, the total whiplash group was divided into subgroups of malingering (ASTM score <86) and non-malingering post-whiplash patients (ASTM score >85). Then the scores on conventional memory and attention tests of these subgroups were compared with those of patients with closed head injury and normal controls. Raw scores on the Stroop and trail making tests were log transformed before statistical testing to normalise the distributions. Overall group comparisons were done by univariate analyses of variance (ANOVA) with the exception of the AVLT recognition score, which was assessed by Kruskal-Wallis test. To control for multiple comparisons, significance was accepted at a Bonferroni corrected level of 0.003. When a variable showed significant overall group differences, post hoc analyses were performed with Scheffé test at a significance level of 0.05. All reported p values are two tailed.

Results

Forty three post-whiplash patients scored below the cut off value on the ASTM malingering test. The prevalence of underperformance in the litigation patients was 0.61 (22 of 36; 95% confidence interval (95% CI) 0.45–0.77). The prevalence in the non-litigation patients was 0.29 (21 of 72; 95% CI 0.18–0.40). The difference was significant (Yates corrected $\chi^2$=8.93, df=1, p=0.003). The prevalence among the outpatients who were involved in a damage or workmen’s compensation claim was 0.33 (20 of 60; 95% CI 0.21–0.45), whereas it was only 0.08 (1 of 12; 95% CI 0.0–0.24) in those who had no such claims. Although suggestive, this difference was not significant (Yates corrected $\chi^2$=1.94, df=1, p>0.20).

Table 1 shows the demographic characteristics of the post-whiplash subgroups, patients with closed head injury, and control subjects.

---

**Table 1**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Post-whiplash</th>
<th>Closed Head Injury</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>45.2 ± 15.8</td>
<td>50.1 ± 12.3</td>
<td>49.7 ± 10.6</td>
</tr>
<tr>
<td>Gender (M/F)</td>
<td>27/16</td>
<td>12/28</td>
<td>20/30</td>
</tr>
<tr>
<td>Injury Type</td>
<td>whiplash</td>
<td>closed head injury</td>
<td>normal</td>
</tr>
<tr>
<td>Disability</td>
<td>5.8 ± 3.2</td>
<td>4.5 ± 2.1</td>
<td>3.2 ± 1.8</td>
</tr>
<tr>
<td>Employment</td>
<td>0.7 ± 0.5</td>
<td>0.6 ± 0.4</td>
<td>0.5 ± 0.3</td>
</tr>
<tr>
<td>Education</td>
<td>15.1 ± 2.5</td>
<td>15.2 ± 3.0</td>
<td>14.7 ± 2.8</td>
</tr>
</tbody>
</table>

---

The difference was significant (Yates corrected $\chi^2$=8.93, df=1, p=0.003). The prevalence among the outpatients who were involved in a damage or workmen’s compensation claim was 0.33 (20 of 60; 95% CI 0.21–0.45), whereas it was only 0.08 (1 of 12; 95% CI 0.0–0.24) in those who had no such claims. Although suggestive, this difference was not significant (Yates corrected $\chi^2$=1.94, df=1, p>0.20).

Table 1 shows the demographic characteristics of the post-whiplash subgroups, patients with closed head injury, and control subjects.
The groups were not significantly different from each other for sex, age, educational level, or premorbid intelligence as measured by DART-IQ.

Table 2 shows the results of the memory and concentration tests. All these tests showed significant differences across the four groups. Scheffé tests showed that the malingering group performed as poorly as patients with closed head injury in most instances, whereas the non-malingering post-whiplash patients scored better than these two groups, but worse than the normal controls. The recognition score of the word learning task (AVLT) was significantly lower in the malingering group than in the other two patient groups.

**Discussion**

The main result of our study is that a significant proportion of post-whiplash patients seemed to be performing below their actual capacities. In the clinical sample this proportion is about one out of three or four patients; in the context of litigation the prevalence of underperformance is twice as high. Furthermore, both malingering and non-malingering patients scored below normal controls on memory and concentration tests. This result replicates earlier findings of compromised memory and attention in late post-whiplash patients. However, the malingering post-whiplash patients performed as poorly as the patients with closed head injury. The patients with closed head injury were purposely selected because of incapacitating cognitive disorders as a consequence of documented severe brain damage. Our results suggest that the extremely poor performance of some post-whiplash patients is not caused by organic brain disorder but can be explained by underperformance. This alternative explanation does not only depend on the ASTM malingering test, but is also supported by the results of the word learning test (AVLT). The malingers had low recognition scores compared with the patients with closed head injury, whereas the recall scores of these two groups were not significantly different. This dissociation is generally regarded as another sign of malingering.

The finding that underperformance was twice as frequent in litigation cases as in clinical patients suggests that financial claims may strongly influence test behaviour. This is consistent with a recent meta-analysis of research on closed head injury, which concluded that “patients with less severe injuries, as measured by post-traumatic neurological data, are more likely to seek monetary compensation”. It underscores the importance of applying formal tests of motivation and effort in the examination of patients who present with questionable syndromes, or who have financial claims.

The non-malingering post-whiplash patients scored about 1 SD below normal controls on the memory and concentration tasks. This is a clinically significant finding which cannot be explained as the result of malingering. Its order of magnitude is similar to that of the mild cognitive disorders found in patients with other types of chronic pain, chronic fatigue, and non-psychotic depression. Thus it seems plausible that the reduced cognitive function in non-malingering post-whiplash patients is a consequence of (a combination of) these factors.

Some remarks on methodology are in order. Firstly, our sample is not representative for post-whiplash patients in general. We examined our patients two to three years after the accident, and all of them had cognitive complaints. By contrast, most whiplash cases that come to medical attention do not develop a chronic condition, and less than half of them report memory or concentration problems. Secondly, we defined malingering by ASTM scores less than 86 points, based on the validation study. In view of the score range of the patients with closed head injury this was perhaps too conservative. With a sharper cut off (≤87) the prevalence figures would have been higher. However, we preferred to remain on the safe side, as the ASTM test is a new test with as yet relatively few validity data. Thirdly, it could be argued that low scores on the ASTM test are perhaps due to the effect of fatigue. Although we did not expressly examine this possibility in an experimental way, we think that it is unlikely. In one of the
participating centres the tests were always administered in the same order. The data from this centre (not reported), did not show a trend indicating a build up of fatigue.

Finally, we stress that the concept of malingering was used in this paper not only to mean deliberate fabrication of bad test results, but also in the sense of a possibly unconscious tendency to perform below the actual level of competency. Such a tendency might be induced by factors such as assumption of patient role, the need to get recognition for complaints in the face of medical scepticism, or perhaps by a strategy of self protection against exhaustion. There might also be an element of self deception, in that the patients' beliefs about their complaints change in a direction of greater consonance with their illness behaviour. It is impossible to distinguish between these alternatives with the ASTM or similar malingering tests.

We do not conclude that the problems of patients with late post-whiplash complaints are mere products of their imagination. On the contrary, we think that the complaints should be taken seriously. Whenever it has to be assumed that an underperforming post-whiplash patient is acting in good faith, it is relevant to thoroughly assess the emotional aspects and the behavioural consequences of his situation. This is important in view of patient management, rehabilitation, and prevention of medical shopping. Our findings only indicate that the neuropsychological test results of groups of post-whiplash patients are strongly influenced by a subgroup of patients who perform way below their actual level. Explanations of the poor results of this subgroup in terms of brain damage are not warranted.

We are grateful to M. Fiedeldij Dop, Dr B.P. Radanov, and Professor J. Staam for their helpful comments.