

# Cognitive sequelae of severe head injury in relation to the Glasgow Outcome Scale

DN BROOKS, J HOSIE,\* MR BOND, B JENNETT, M AUGHTON

From the Department of Psychological Medicine, University of Glasgow; Department of Clinical Psychology, Royal Dundee Liff Hospital; Department of Psychological Medicine, Southern General Hospital, Glasgow, UK

**SUMMARY** Aspects of reliability and validity of the Glasgow Outcome Scale were studied by examining outcome scale categories and cognitive test scores in 57 patients with severe closed head injury seen within 2 years of injury. Inter-rater reliability of a three and six point scale was high. The three point scale, unlike the six point, showed a significant statistical association with cognitive test scores. The relationship between test score and outcome category was accurate within the first 3 months of injury, but considerably less so thereafter.

With improved surgical and nursing care many seriously damaged head injured patients survive with a variety of physical, cognitive and behavioural disabilities. The need for a set of clearly defined criteria of outcome is obvious not only in enabling researchers to communicate unambiguously, but also to allow clear predictions about the effects of treatment and management on outcome after injury, and to analyse the nature of disabilities. Such a scale has been devised by two of the present authors,<sup>1</sup> and its practical utility in assessing outcome in a general sense has been demonstrated in a variety of contexts,<sup>2-7</sup> although see reference 8 for a contrary view. The reliability and validity of the scale have been demonstrated obliquely, by showing that the distribution of cases in different outcome grades is very similar in three separate neurosurgical centres,<sup>2</sup> but the present study aims explicitly to examine reliability. Also, one aspect of validity (that relating to cognitive outcome), by comparing outcome scores of two experienced raters on a large group of head injured patients, and by examining the distribution of cognitive disability in the different outcome grades of the same group of patients. Levin *et al.*<sup>9</sup> have reported such a relationship for IQ data, and the present study examines the relationship for a variety of other neuropsychological data.

## Method

### *Patients studied (table 1)*

Fifty seven severely head injured patients admitted to a neurosurgical unit following head injury were examined for the purpose of this study. To fulfil the definition of "severe", patients must have had a coma duration assessed by the Glasgow Coma Scale,<sup>10</sup> of at least 8 hours, or a post traumatic amnesia (PTA, the interval between the injury and regaining continuous day to day memory), assessed retrospectively by careful clinical questioning, of at least 48 hours.

### *Procedure*

Patients had been examined cognitively within 2 years of injury (many within 6 months) on a variety of neuropsychological tests. The neurosurgical case sheets were given blind to two of the authors, both experienced users of the scale (MB and BJ), who were asked to rate each case of the three and six point outcome scales at a time corresponding to the time after injury at which psychological testing had been carried out. If the rater felt that insufficient information was available for reliable grading, that patient was excluded from the analysis, resulting in 51 cases rated in common on the three point scale, and 48 on the six point.

The Glasgow Outcome Scale has been described in detail elsewhere.<sup>17</sup> (Appendix 1). Four grades of survivor are described as follows: Persistent vegetative state (PVS), severe disability (SD), moderate disability (MD), and good recovery (GR). The current paper focuses only on survivors who became better than PVS. The scale assesses aspects of functional dependence, with patients in the SD category being conscious but disabled enough to be dependent upon others for daily support. Those in the MD category are disabled but independent, in that they can live alone and could undertake some kind of work. Patients in the GR category resume a normal life, although they may still have minor deficits.

Address for reprint requests: Professor DN Brooks, University Department of Psychological Medicine, 6 Whittinghame Gardens, Great Western Road, Glasgow G120AA, UK.

Received 16 July 1985 and in revised form 18 September 1985.  
Accepted 21 September 1985

The raters were asked to rate each patient on the conventional three point scale as indicated above, and where sufficient information was available, raters also used a six point scale by rating the patient as "high" or "low" on whichever of the three point scale is appropriate (that is a very severe SD would be rated as "low").

The patients were examined on a variety of cognitive tests of intelligence, learning and memory, and language. Precise descriptions of the tests may be obtained in reference 11. Very briefly, intelligence was assessed using Raven's Progressive Matrices and Mill Hill Vocabulary Scale.<sup>12,13</sup> Learning and memory were assessed by immediate and delayed recall of the logical memory story in the Wechsler Memory Scale,<sup>14</sup> and by noting errors to learn three pairs of words both initially and at relearning.<sup>15</sup> Language was assessed by a word fluency measure<sup>16</sup> in which patients had 60 seconds to give as many category exemplars as possible (for example occupations), and then two 60 second periods in which to produce words beginning with an "easy" or "hard" letter.

## Results

### Inter-Rater Reliability

Measures of inter-rater reliability were made by comparing the ratings of the two experienced assessors (MB and BJ). In this situation, it is the number of cases differently classified that matter, rather than a level of statistical association, and statistics were not, therefore, computed. There was very high agreement on the three point scale (table 2). In 36 cases (70%) the agreement was perfect, and in all cases the agreement was within one outcome category. The six point scale (table 3) once again showed a high agreement between raters, with only four cases rated differently by more than one scale category, and none by more than this discrepancy. It is in the middle category of

moderate disability that disagreement occurs, although even here, the disagreements are not large.

### Validity

In view of the high inter-rater reliability, the ratings of one rater (MB) were used giving a total of 57 cases. The validity of the scale in terms of distinguishing between the three groups on cognitive test performance was assessed by comparing the cognitive scores of patients in the three sub groups via a one way analysis of variance (table 4). The F ratios were significant at at least  $p = 0.05$  in eight of the nine tests, and in all tests except the "hard" word fluency measure, there was a clear reduction in score with increasing severity of outcome category. Nevertheless, as the standard deviations indicate, the

Table 2 Distribution of three outcome categories by two raters

Rater (MB)		Rater (BJ)		
		SD	MD	GR
	SD	9	6	0
	MD	1	10	3
	GR	0	5	17

Table 3 Distribution of six outcome categories by two raters

Rater (MB)		SD	Rater (BJ)					
			SD		MD		GR	
			-	+	-	+	-	+
	SD	-	2	8	0	0	0	0
		+	0	7	4	0	0	0
	MD	-	0	0	6	1	1	0
		+	0	1	0	3	1	0
	GR	-	0	0	2	4	2	0
		+	0	0	0	0	4	2

Table 1 Details of head injured cases

(a)								
Coma (days)	0	1	2	3	4	5	6	Unknown
No. of cases	1	15	11	7	6	0	3	14
(b)								
PTA (days)	1-7	8-14	Over 14	Unknown				
No. of cases	13	10	25	9				
(c)								
Time after injury (months)	Up to 1 month	2-3	4-6	7-12	13-24			
No. of cases	18	14	12	10	3			
(d)								
Age (yr)	15-30	31-45	46-60	Over 60				
No. of cases	25	13	17	2				

groups overlapped considerably in test scores, with particular overlap between the MD group and each of the other two groups. When Scheffé tests were performed, the significant F ratio was found to be due either to the difference between the GR and SD groups, or between the SD and the remaining two groups combined. The crucial significance therefore appears to be due to the very poor scores of the SD cases who performed conspicuously below patients in the other two groups. On the one measure failing to show a significant F ratio (immediate associated learning), all three head injury groups performed very poorly, and there is therefore probably a "floor" effect here.

An alternative examination of validity may be made by comparing three outcome grades in terms of an external measure of severity of brain damage such

as the duration of post traumatic amnesia. As the results in table 4 show, the three severity groups did indeed differ significantly in terms of PTA duration, and the significance was seen on Scheffé testing to arise from the difference between the GR and the other two groups combined.

The validity of the six point scale was assessed in a similar manner by one way analysis of variance across the six categories, although the larger number of groups resulted in a small number of cases in each group. Table 5 shows that seven of the nine F ratios were significant at at least 0.05, but inspection of the data shows that the expected gradation of scores did not always appear, and the clear correspondence between outcome and cognitive performance found with the three point scale was not obvious on the six point scale. However, on Associate Learning, a test

Table 4 Cognitive test mean scores in three outcome groups: summary of analyses of variance

Cognitive test	Outcome Grade						F
	(SD)		(MD)		(GR)		
	Mean (N = 15)	SD	Mean (N = 16)	SD	Mean (N = 26)	SD	
Progressive Matrices	81.2	13.2	89.7	16.5	94.0	13.2	3.83*
Mill Hill Vocabulary	81.8	10.9	92.6	14.6	93.2	7.2	5.91†
<i>Logical memory</i>							
Immediate	4.0	2.4	8.4	4.4	11.3	4.4	16.21†
Delayed	2.7	2.0	5.4	3.7	9.0	5.0	12.28†
<i>Associate learning</i>							
Immediate	13.6	12.0	13.4	9.7	8.0	6.3	2.37
Delayed	14.4	12.2	8.7	11.7	4.3	7.1	4.59†
<i>Word fluency</i>							
Category	7.7	5.2	10.6	3.6	14.1	3.4	11.90†
Easy letter	4.0	4.9	8.0	5.3	9.1	4.3	5.56†
Hard letter	1.9	2.5	4.4	3.5	3.8	2.3	3.48*
PTA (Days)	24.2	7.6	20.5	8.2	12.2	10.1	7.86†

\*p ≤ 0.05; †p ≤ 0.01.

Table 5 Cognitive test mean scores in six outcome groups: summary of analyses of variance

Cognitive test	Outcome Grade						F
	(SD)		(MD)		(GR)		
	- (6)	+ (8)	- (6)	+ (8)	- (11)	+ (10)	
Progressive Matrices	78.8	84.0	92.0	89.0	100.5	90.9	2.27
Mill Hill	77.3	87.4	95.5	94.5	92.1	94.6	2.69*
<i>Logical memory</i>							
Immediate	3.0	5.1	7.6	10.0	11.1	12.0	5.78†
Delayed	2.0	3.2	4.4	6.4	9.0	10.0	4.20†
<i>Associate learning</i>							
Immediate	21.3	6.9	16.6	8.1	6.8	9.6	3.40†
Delayed	25.2	7.2	13.2	5.6	4.8	4.6	4.53†
<i>Word fluency</i>							
Category	6.0	9.2	10.4	10.5	13.0	15.0	4.05†
Easy letter	1.8	6.1	7.2	9.3	10.1	8.8	2.45*
Hard letter	1.2	2.6	4.0	5.3	3.4	4.4	1.60
PTA	28.0	21.4	15.4	22.3	11.3	10.9	4.10

\* p ≤ 0.05; † p ≤ 0.01.

which head injured patients find conspicuously difficult, patients in both categories of severe disability showed no reduction in error between the immediate and delayed testing showing the extremely poor verbal retention in these patients.

#### *Validity as a function at time of testing*

As the patients in this study were seen at any time from 1 to 24 months after injury, this afforded the opportunity to examine the correspondence between outcome grade and cognitive performance at different times after injury. In order to have sufficient cases for data analysis, only the three point outcome grading was used, and the 32 "early" patients seen within 3 months after injury were analysed separately as were the 25 "late" patients seen after 3 months. The choice of 3 months was to some extent arbitrary, being dictated partly by the need to have adequate numbers of patients for analysis, but also by clinical experience which shows that in many patients cognitive recovery on testing is slowing considerably by 3 months. This does not, of course, deny the reality of continuing improvement in many patients. Table 6 gives a summary of the individual analyses of variance, although interpretation is somewhat problematic due to the small shows of cases in some of the categories (only 5 late SD and MD cases). The results are nevertheless clear in that the F ratios are significant for eight of the nine tests in the early group, but only two in the late group. Scheffé testing showed that the significance arose due to the particularly poor cognitive performance of the SD group.

#### Discussion

This study concerned both reliability and validity of the outcome scale, and reliability was found to be satisfactory when examined by comparing two raters.

Table 6 *Significant F ratios comparing three outcome grades in "early" and "late" patients*

Cognitive test	Early		Late	
	F	Sig	F	Sig
Progressive Matrices	11.28	≤0.01	0.01	NS
Mill Hill	8.22	≤0.01	0.32	NS
<i>Logical Memory</i>				
Immediate	12.29	≤0.01	3.82	≤0.05
Delayed	8.11	≤0.01	4.15	≤0.05
<i>Associate Learning</i>				
Immediate	2.77	NS	1.60	NS
Delayed	5.25	≤0.01	1.68	NS
<i>Word Fluency</i>				
Category	11.06	≤0.01	2.00	NS
"Easy"	8.80	≤0.01	0.36	NS
"Hard"	4.25	≤0.05	0.30	NS
PTA	6.84	≤0.01	1.81	NS

Given adequate reliability, an analysis of validity in terms of cognitive performance showed highly significant and consistent differences between the outcome subgroups, with a clear gradation in score between GR and SD groups. This gradation was not, however, so evident when dividing the scale into six rather than three points, and this reduction in validity cannot be due simply to lowered reliability as this was still acceptably high on the six point scale.

Scheffé comparisons across the three outcome groups showed that in many cases the significant difference in cognitive performance arose from the difference between the two extremes, with the moderate disability patients overlapping with both the good recovery and severe disability cases.

The scale is a relatively crude one (as it was intended to be), assessing very general aspects of outcome related more to functional disability than to cognitive deficit. Despite this, there is a correspondence with cognitive outcome on the three point scale, but it is important to note that this correspondence is dependent upon the time after injury at which a patient was assessed. Comparisons across the three outcome groups gave many significant F ratios within 3 months of injury, but not later, and it may be that during this early period both scales (outcome scale and cognitive performance) are driven by a third factor: the severity of injury. This is supported by the significant F ratio for PTA early, but not late after injury. As time passes, both psychological test performance and outcome scale categorisation appear to become more multi-determined, with factors such as severity of injury or persisting focal neurological deficit; the patient's affective and behavioural status, possibilities of return to work, and opportunities to live in a psychologically stimulating environment, all interacting to give the final outcome picture.

Outcome scales are very useful for epidemiologically based research, whereas psychological tests are useful for neuropsychologically based research and for individual patient rehabilitation. It would be unreasonable to expect one to predict the other in anything other than a broad sense, and it is important to be aware that patients may achieve considerable clinical improvement yet still remain within the same broad outcome category. It is not adequate to use the Outcome Scale to categorise outcome in an individual case, particularly so when behavioural and psychological disturbances are evident. To do justice to individual cases requires a range of measures which relate clearly to the individual disability of an individual patient.

Dr J Hosie and Ms ME Aughton were supported by a Mental Health Foundation Research Grant.

## Appendix

## THE GLASGOW OUTCOME SCALE

*The three categories of non-vegetative survival*

**Severe disability** The patient is conscious but needs assistance from another person for some activities of daily living every day. The state may range from total dependency (for washing and feeding, etc.) to the need for help with only one activity such as dressing, or mobility. Very often dependency is due to a combination of physical and mental disability, and it is not uncommon for a severely head injured patient to have little residual physical deficit, but to have such an inability to organise his/her life that he is functionally severely disabled. Patients in this category may be described as "conscious but dependent".

**Moderate disability** These patients may look after themselves at home, and are independently mobile. Nevertheless, some previous activities (either vocational or social) are no longer possible due to either mental or physical disabilities or both. It is not impossible for patients in this category to be able to return to certain kinds of work, but nevertheless they would be described as "independent but disabled".

**Good recovery** Patients in this category may resume normal vocational and social activities, although often in the context of persisting minor physical or mental deficits. Nevertheless, for a variety of reasons (often mental rather than physical) the patient may not have resumed all previous social activities, and may not be working.

## References

- <sup>1</sup> Jennett B, Bond M. Assessment of outcome after severe brain damage. A practical scale. *Lancet* 1975;1:480-4.
- <sup>2</sup> Jennett B, Teasdale G, Galbraith S, *et al.* Severe head injuries in three countries. *J Neurol Neurosurg Psychiatry* 1977;40:291-8.
- <sup>3</sup> Jennett B, Snoek J, Bond MR, Brooks N. Disability after severe head injury: observations on the use of the Glasgow Outcome Scale. *J Neurol Neurosurg Psychiatry* 1981;44:285-93.
- <sup>4</sup> Miller JD, Butterworth JF, Gudeman SK, *et al.* Further experience in the management of severe head injury. *J Neurosurg* 1981;54:289-99.
- <sup>5</sup> Gennarelli TA, Spielman GM, Langfitt TW, *et al.* Influence of the type of intracranial lesion on outcome from severe head injury. *J Neurosurg* 1982;56:26-32.
- <sup>6</sup> Teasdale E, Cardoso E, Galbraith S, Teasdale G. CT scan in severe diffuse head injury: physiological and clinical correlations. *J Neurol Neurosurg Psychiatry* 1984;47:600-3.
- <sup>7</sup> Jennett B. The Measurement of Outcome. In: Brooks N, ed. *Closed Head Injury: Psychological, Social, and Family Consequences*. Oxford: Oxford University Press, 1984:37-43.
- <sup>8</sup> Hall K, Cope N, Rappaport M. Glasgow Outcome Scale and Disability Rating Scale: comparative usefulness in following recovery in traumatic head injury. *Arch Phys Med Rehab* 1985;66:35-7.
- <sup>9</sup> Levin HS, Grossman RG, Rose JE, Teasdale G. Long-term neuropsychological outcome of closed head injury. *J Neurosurg* 1979;50:417-22.
- <sup>10</sup> Teasdale G, Jennett B. Assessment of coma and impaired consciousness. *Lancet* 1974;ii:81-4.
- <sup>11</sup> Brooks N, Aughton ME. Psychological consequences of blunt head injury. *Int Rehab Med* 1979;1:160-5.
- <sup>12</sup> Raven JC. *Guide to the Standard Progressive Matrices*. 1960. London: HK Lewis.
- <sup>13</sup> Raven JC. *Extended guide to the Mill Hill Vocabulary Scales*; 1962. London: HK Lewis.
- <sup>14</sup> Wechsler D. A standardised memory scale for clinical use. *J Psychol* 1945;19:87-95.
- <sup>15</sup> Inglis J. A paired-associate learning test for use with elderly psychiatric patients. *J Ment Sci* 1959;105:440-3.
- <sup>16</sup> Borkowski JG, Benton AL, Spreen O. Word fluency and brain damage. *Neuropsychologia* 1967;5:135-40.