



## original papers

Psychiatric Bulletin (2005), 29, 53–55

S. FLEMINGER, E. LEIGH, P. EAMES, L. LANGRELL, R. NAGRAJ AND S. LOGSDAIL

# HoNOS—ABI: a reliable outcome measure of neuropsychiatric sequelae to brain injury?

### AIMS AND METHOD

The Health of the Nation Outcome Scale for Acquired Brain Injury (HoNOS—ABI) is a relatively new outcome measure designed to assess the neuropsychiatric sequelae of brain damage. This study investigated the interrater reliability of this scale. Fifty patients with traumatic brain

injury receiving rehabilitation were each rated twice on the HoNOS—ABI, by two different raters. There were 24 raters in total.

### RESULTS

Weighted kappa values ranged from 0.43 to 0.84 and intraclass correlation coefficients from 0.58 to 0.97 for the ten items assessed. This indicated

that agreement was moderate to substantial for all items.

### CLINICAL IMPLICATIONS

The scales consistently measured the items of interest across different raters. This indicates that HoNOS—ABI is a reliable outcome measure when applied by different raters in routine clinical practice.

Neuropsychiatric sequelae of brain injury are numerous and can include disturbances in cognition, mood and behaviour. They commonly impede individuals' ability to function in their work and family life and are responsible for at least as much disability as the associated physical symptoms (Lishman, 1998).

Treatment includes pharmacological medication and rehabilitation. Rehabilitation is multifaceted (Rao & Lyketsos, 2000) and it is often difficult to determine which specific interventions are responsible for improvements in an individual. It is therefore necessary to have good outcome measures. Fleming & Powell (1999) highlighted that most importantly, outcome measures need to be relevant to the patient and carer; they must also be trustworthy, and ideally used consistently across studies to facilitate comparison of treatment methods.

The Health of the Nation Outcome Scales (HoNOS; Wing *et al*, 1998) were produced to provide an easily administered and reliable measure to be used in general adult mental health. Subsequent versions have been developed for more specialist settings. The Health of the Nation Outcome Scale for Acquired Brain Injury (HoNOS—ABI) for the assessment of individuals who have sustained a brain injury was developed by the UK Brain Injury Psychiatrists Group in conjunction with the Royal College of Psychiatrists' Research Group, and has been available since 1999 (further details available from the author upon request).

Little work has been done to investigate the clinical relevance of HoNOS—ABI. Coetzer & Du Toit (2001) found promising correlations between the scale and three other outcome measures, including post-injury employment status. These findings indicate that the HoNOS—ABI is

valid and pertinent to patients, as it relates outcome to reintegration into the community. However, there is no published assessment of interrater reliability for HoNOS—ABI, an index Portney & Watkins have argued is 'especially important when measuring devices are new' (Portney & Watkins, 1993, p. 60). Our study therefore investigated the interrater reliability of this measure.

## Method

### Raters

The 24 raters consisted of staff from five neuropsychiatric brain injury rehabilitation sites in the UK. They were all healthcare professionals and included psychiatrists, psychologists and charge nurses. None had received formal training in using the HoNOS—ABI but all were familiar with the completion of outcome measures.

### Participants

The 50 in-patients ranged in age from 18 years to 65 years. All had disabling neuropsychiatric sequelae following severe traumatic brain injury requiring in-patient rehabilitation.

### Design and materials

Every patient was assessed independently by two raters. Each pair of raters saw only a sample of the study participants. The assessments were made separately by the two raters in the course of routine clinical practice and were not conducted through an interview process.



original papers

The raters were well acquainted with the patients they assessed, who were all residents of the in-patient units. The HoNOS–ABI consists of 12 items, each reflecting a different domain of symptoms rated on a five-point scale (with 0 indicating no problem). Items 11 and 12 are designed for patients in community settings and were therefore excluded from the analysis. Item 3 relates to problems associated with alcohol and drug use, which can be difficult to assess among in-patients ( $n=36$  for this item). All analyses were performed on a personal computer using Microsoft Excel, the Statistical Package for the Social Sciences version 11.0 and Stata version 8.

### Statistical analysis

The intraclass correlation provides an assessment of interrater reliability by comparing the amount of variation between raters with the amount of variation between individuals. A one-way random analysis of variance was used to calculate the intraclass correlation coefficient (ICC) because each pair of raters had not assessed all the participants. In order to take the closeness of agreement between raters into account, weighted kappa values ( $\kappa_w$ ) were also calculated for each item (Fig. 1). Significance tests can identify whether raters show an agreement above chance or not. Values of  $\kappa_w$  are always lower than their corresponding ICC. To interpret the degree of agreement, the guidelines provided by Landis & Koch (1977) were used: 0.21–0.40 is seen as a fair level of agreement, 0.41–0.60 as moderate, 0.61–0.80 as substantial and  $>0.81$  as almost perfect. Although these ‘divisions are arbitrary, they do provide useful benchmarks’ (Landis & Koch, 1977, p.165).

### Results

Table 1 shows the mean scores for each item, which range from 0.21 to 3.04. The low mean score (0.21) and associated standard deviation (0.01) on the item relating to alcohol or drug misuse may be because the participants were in-patients without ready access to alcohol or drugs. The domains with the highest mean scores were those assessing cognitive problems, problems with activities of daily living, and relationships. The standard deviation for all items is quite low, suggesting that most participants had similar levels of problems.

Table 2 shows the  $\kappa_w$  and ICC values and their confidence intervals for each item. The interrater reliability ranges from 0.43 to 0.84 for  $\kappa_w$  and 0.58 to 0.97 for ICC.

		Rater A				
		0	1	2	3	4
Rater B	0	1	0.5	0	0	0
	1	0.5	1	0.5	0	0
	2	0	0.5	1	0.5	0
	3	0	0	0.5	1	0.5
	4	0	0	0	0.5	1

Fig. 1. Weight matrix used to calculate weighted kappa values.

Table 1. Scores on each HoNOS–ABI item (0–4) and across all ten items (0–40)

Item	Symptoms	Mean (s.d.)
1	Antisocial	2.17 (0.41)
2	Self-injury	0.47 (0.21)
3	Alcohol/drugs <sup>1</sup>	0.21 (0.01)
4	Cognitive	3.04 (0.34)
5	Physical	1.78 (0.34)
6	Psychotic	1.17 (0.47)
7	Depression	1.25 (0.47)
8	Other	1.45 (0.55)
9	Relationships	2.29 (0.41)
10	Activities of daily living	2.80 (0.37)
Total scores		19.09 (1.46)

HoNOS–ABI, Health of the Nation Outcome Scale for Acquired Brain Injury.  
1.  $n=36$ .

Table 2. Weighted kappa and interclass correlation values for all items of the HoNOS–ABI

Item	Symptoms	$\kappa_w$ (95% CI)	ICC (95% CI)
1	Antisocial	0.54 (0.38–0.70)	0.73 (0.57–0.84)
2	Self-injury	0.56 (0.36–0.76)	0.71 (0.55–0.83)
3	Alcohol/drugs <sup>1</sup>	0.84 (0.59–1.09)	0.97 (0.94–0.98)
4	Cognitive	0.53 (0.35–0.71)	0.68 (0.50–0.80)
5	Physical	0.66 (0.50–0.82)	0.81 (0.68–0.88)
6	Psychotic	0.51 (0.32–0.70)	0.64 (0.44–0.78)
7	Depression	0.43 (0.26–0.60)	0.62 (0.41–0.76)
8	Other	0.53 (0.35–0.71)	0.58 (0.36–0.74)
9	Relationships	0.52 (0.36–0.68)	0.73 (0.57–0.84)
10	Activities of daily living	0.54 (0.36–0.73)	0.68 (0.50–0.80)

HoNOS–ABI, Health of the Nation Outcome Scale for Acquired Brain Injury; ICC, intraclass correlation coefficient.  
1.  $n=36$ .

Calculation of test statistics for  $\kappa_w$  ( $z = \kappa_w / \text{standard error}$ ) indicated that, for all items, the level of agreement between the pairs of raters was significantly greater than chance ( $P < 0.001$ ), which was also supported by the finding that the confidence intervals for both  $\kappa_w$  and ICC values did not include zero. The level of agreement was highest for the item relating to drug or alcohol problems ( $\kappa_w=0.82$ , ICC=0.97). The lowest interrater reliability was for the item corresponding to depressive symptoms ( $\kappa_w=0.43$ ), whereas the lowest reliability for the ICC values was for the item relating to other symptoms (0.58). All values showed at least moderate agreement (Landis & Koch, 1977).

### Discussion

#### Reliability across items

The Z test of significance on the  $\kappa_w$  values and the finding that none of the confidence intervals for  $\kappa_w$  or ICCs contained zero indicated that the level of agreement



between the pairs of raters was significantly greater than chance for all items of the HoNOS—ABI assessed. The interrater reliability of the item relating to problems with drugs or alcohol was 0.84, which may be spuriously high owing to the smaller sample size used, but is more likely to be due to the lack of access to drugs in rehabilitation settings and the concrete nature of the question. The lowest  $\kappa_w$  value was for the depressive symptoms item, indicating that it is more difficult to rate this item consistently. The ICC was lowest for the item assessing other mental and behavioural problems, with a wide confidence interval of 0.38. This item is disparate in nature and might be more clinically useful if viewed qualitatively. Interrater reliability values for the other items were all similar and showed at least moderate agreement.

### Rater training

It is possible that differences in raters' interpretation of the items, rather than their assessment of the patients, led to lower reliability. Brooks (2000) examined the efficacy of staff training on the reliability of the generic HoNOS and concluded that although 'reasonable improvements could be gained' (p. 509), staff training could also be of 'no value' (p. 609). The findings are clearly inconclusive, although training could increase the reliability of the scales.

### Neuropsychiatric sequelae

Overall, the participants were rated as having particular problems within the domains of cognitive functioning, relationships and activities in daily life, in line with the findings of Coetzer & Du Toit (2001). In contrast, Orrell et al (1999) investigated scores on the generic HoNOS among a population of psychiatric patients and found that the patients were rated as having more problems with depressed mood, other mental health problems and relationships. In this investigation, participants were rated as having more severe problems on most items (mean scores 0.21–3.04) compared with those assessed in the study by Orrell and colleagues (mean scores 0.24–1.64; Orrell et al, 1999). The findings reported here and by Coetzer & Du Toit (2001) support the proposal that individuals who have sustained a brain injury tend to present with particular difficulties in cognition which then affect their general psychosocial functioning and ability to perform activities of daily living (Ponsford et al, 1995).

### Potential criticisms

The generalisability of the study is limited; all the participants were in-patients on neuropsychiatric brain injury units with moderate to severe cognitive, behavioural and/or neuropsychiatric problems resulting from traumatic brain injury. Reliability will be higher because all raters had the opportunity to observe and discuss the patients in some detail. Raters did not have to rely on taking a history

from the patient or an informant to rate the patient, as they would probably need to do in an out-patient or community setting. On the other hand, we used data from several sites and the ratings were made during the course of routine clinical care without specific training.

### Implications of the study

The findings from our study indicate that interrater reliability was good for most items of the HoNOS—ABI and moderate for one item. Although there is still a need for further evaluation of the validity and reliability (including test–retest) of this scale, our study complements the work of Coetzer & Du Toit (2001) and highlights the potential usefulness of the scale both in a research setting and as part of routine clinical practice. The HoNOS—ABI proved to be a reliable outcome measure of the neuropsychiatric sequelae of brain injury across different raters when used for assessment within a rehabilitation setting, during the course of routine clinical practice.

### Declaration of interest

None.

### References

- BROOKS, R. (2000) The reliability and validity of the Health of the Nation Outcome Scales: validation in relation to patient derived measures. *Australian and New Zealand Journal of Psychiatry*, **34**, 504–511.
- COETZER, R. & DU TOIT, P. L. (2001) HoNOS—ABI; a clinically useful outcome measure? *Psychiatric Bulletin*, **25**, 421–422.
- FLEMINGER, S. & POWELL, J. (1999) Evaluation of outcomes in brain injury rehabilitation. *Neuropsychological Rehabilitation*, **9**, 225–230.
- LANDIS, J. R. & KOCH, G. G. (1977) The measurement of observer agreement for categorical data. *Biometrics*, **33**, 159–174.
- LISHMAN, W. A. (1998) Head injury. In *Organic Psychiatry: The Psychological Consequences of Cerebral Disorder* (3rd edn), pp. 161–217. Oxford: Blackwell.
- ORRELL, M., YARD, P., HANDYSIDES, J., et al (1999) Validity and reliability of the Health of the Nation Outcome Scales in psychiatric patients in the community. *British Journal of Psychiatry*, **174**, 409–412.
- PONSFORD, J., SLOAN, S. & SNOW, P. (1995) *Traumatic Brain Injury: Rehabilitation for Everyday Adaptive Living*. Hove: Erlbaum.
- PORTNEY, L. G. & WATKINS, M. P. (1993) *Foundations of Clinical Research: Applications to Practice*. Stamford, Ct: Appleton & Lange.
- RAO, V. & LYKETSOS, C. (2000) Neuropsychiatric sequelae of traumatic brain injury. *Psychosomatics*, **41**, 95–103.
- WING, J. K., BEEVOR, A. S., CURTIS, R. H., et al (1998) Health of the Nation Outcome Scales (HoNOS): research and development. *British Journal of Psychiatry*, **172**, 11–18.

\***Simon Fleminger** Consultant Neuropsychiatrist, Lishman Brain Injury Unit, Maudsley Hospital, Denmark Hill, London, SE5 8AZ (e-mail: s.fleminger@lop.kcl.ac.uk), **Eleanor Leigh** Research Assistant, Lishman Brain Injury Unit, Maudsley Hospital, London, **Peter Eames** Consultant Neuropsychiatrist, Grafton Manor, Towcester, **Lorna Langrell** Associate Specialist in Psychiatry, Robert Ferguson Unit, Royal Edinburgh Hospital, Edinburgh, **Raj Nagraj** Consultant Neuropsychiatrist, Kemsley Division, St Andrew's Hospital, Northampton, **Steven Logsdail** Consultant Psychiatrist, Campbell Unit, Hospital Campus, Milton Keynes