Early Intervention IN PSYCHIATRY
First Impact Factor released in June 2010 and now listed in MEDLINE!
Early Intervention in Psychiatry 2012; **6:** 423–431 doi:10.1111/j.1751-7893.2011.00329.x

Original Article

Validation of the Health of the Nation Outcome Scales as a routine measure of outcome in early intervention programmes

Antonio Preti,[1,2] Alessia Pisano,[2] Maria Teresa Cascio,[2] Federico Galvan,[2] Emiliano Monzani,[2] Anna Meneghelli[2] and Angelo Cocchi[2]

[1]Centro Medico Genneruxi, Cagliari, and [2]Azienda Ospedaliera Ospedale Niguarda Ca' Granda, Dipartimento di Salute Mentale: Programma2000, Milano, Italy

Corresponding author: Professor Dott. Angelo Cocchi, A.O. Ospedale Niguarda Ca' Granda, Programma 2000 – Via Livigno, 3, 20128 Milan, Italy. Email: angelo.cocchi@ospedaleniguarda.it; programma2000@ospedaleniguarda.it

Received 9 July 2011; accepted 6 November 2011

Abstract

Aim: So far, no study has assessed the validity of the Health of the Nation Outcome Scales (HoNOS) in patients enrolled in early intervention programmes, nor has any study evaluated the validity of the HoNOS in people at ultra high-risk (UHR) of psychosis. This study set out to assess the validity and reliability of the HoNOS as a measure of outcome in the patients enrolled in an early intervention programme.

Methods: The concurrent, discriminant and predictive validity, and the reliability of the HoNOS as a measure of outcome in an early intervention programe were assessed in 87 first-episode psychosis (FEP) patients, and in 81 patients at UHR of psychosis.

Results: Reliability indexes were good in the FEP sample, and less good in the UHR sample. HoNOS total scores differentiated between FEP and UHR patients, and the HoNOS subscales proved able to assess a specific profile of symptoms in the two samples, demonstrating a helpful adjunctive measure of health status without complete overlap with other scales. Sensitivity to change was also very good, again with differences between FEP and UHR patients. HoNOS scores at intake did not predict failure to attain remission in FEP patients. There were too few cases of transition to psychosis ($n = 2$) to assess predictive validity of HoNOS in the UHR sample.

Conclusion: HoNOS possesses satisfactory sensitivity and validity to be used in the routine assessment in early intervention programmes.

Key words: early intervention, Health of the Nation Outcome Scales, psychosis, ultra high-risk.

INTRODUCTION

The Health of the Nation Outcome Scales (HoNOS) was developed for the routine measurement of outcome in people with mental disorders.[1] Stemming from the HoNOS, the HoNOSCA and the HoNOS65+ were developed to tailor the specific needs of children and adolescents, and of people aged 65 years and older, respectively.[2,3]

The original version of the HoNOS was reported to possess good reliability,[1,4,5] factor structure,[1,5,6] sensitivity to change[7,8] and usability in a wide range of settings, including busy psychiatric services.[9] However, some studies found the reliability of the total score only marginally adequate,[10] and several alternative factor structures have been proposed.[11–13]

As for patients diagnosed with schizophrenia, most studies on HoNOS validity dealt with patients with a long course of the disorder. One study assessed the validity of the HoNOS in patients with first-episode psychosis (FEP) and found good correlation with established outcome scales, but poor agreement between trained researchers and

patients' usual keyworkers.[14] No study has assessed the validity of the HoNOS in patients enrolled in early intervention programmes, nor has any study evaluated the validity of the HoNOS in people at ultra high-risk (UHR) of psychosis. This recently defined category includes people with signs of incipient psychosis, and principally involves three clusters of subjects: young people with attenuated positive symptoms, as revealed by dedicated interviews[15]; people with diagnosable transient psychotic symptoms, not stabilized in a syndrome yet[16,17]; and a third category of people with genetic risk (first-degree relatives of subjects with psychosis), or meeting the criteria for Schizotypal Personality Disorder, who are showing symptoms of deterioration.[18] These subjects are generally ill, and they can benefit greatly from appropriate psychiatric/psychological help.[19]

At present, HoNOS is used in early intervention programmes as a measure of treatment outcome.[20,21] In the past concerns were raised about the validity and reliability of the HoNOS in patients with psychosis.[5,7] Therefore, despite the advantage of an easy and quick administration, its use in the routine assessment of FEP or UHR patients enrolled in early intervention programs was still to be tested.

This study set out to assess the validity and reliability of the HoNOS as a measure of outcome in the patients enrolled in the Programma2000, an early intervention program active in Milan (Italy) since 1999.[22,23]

METHODS

The Programma2000 is a centre operating in Milan (Italy) since 1999 and specifically devoted to the early diagnosis of and intervention on people with, or at a high risk of psychosis, under the authority of the Niguarda Ca' Granda General Hospital.[22,23] The Institutional Review Board approved the protocol of the study, which complies with the provisions of the Declaration of Helsinki in 1995 (as revised in Tokyo 2004); all the patients gave their informed consent. The sample included 168 patients from a catchment area catering for approximately 200 000 inhabitants.

Diagnosis and assessment

All the patients referred for evaluation underwent a comprehensive, multidimensional evaluation.[22,23] For the purpose of this study, the following standardized assessment instruments were considered: (i) a socio-demographic form; (ii) the Early Recognition Inventory Retrospective Assessment of Symptoms (ERIraos), a 17-item screening checklist intended to select the subjects in need of a more in-depth assessment[24]; (iii) the 24-item Brief Psychiatric Rating Scale (BPRS), to assess general psychopathology;[25,26] (iv) the Global Assessment of Functioning (GAF), which measures the level of functioning of the patient, with higher scores corresponding to better functioning[27]; and (v) the HoNOS.[1,28]

The HoNOS comprises 12 items that rate various aspects of mental and social health with a severity score ranging from 0 (no problem) to 4 (severe to very severe problem). These items are grouped in four subscales: behavioural problems, impairment problems, symptoms problems and social problems (see Table 4 for details). Clinicians rate HoNOS before and after interventions so that changes attributable to the interventions (outcomes) can be measured.

Duration of untreated illness (DUI) and duration of untreated psychosis (DUP) were based on the interview of an informant (a close relative, and preferably a parent); they were both measured as the time elapsed from the onset of key symptoms (anxiety, depression or social withdrawal for DUI; hallucinations, delusions or bizarre behaviour for DUP) to the start of treatment (pharmacotherapy or psychotherapy) prescribed by a psychiatrist; DUP was measured in days, whereas DUI was measured in months.

Inclusion and exclusion criteria

Patients are enrolled in the Programma2000 upon condition that they are aged 17 to 30, and they are at their first contact with any public mental health service of the catchment area for a first episode of psychosis, that is, with DUP $\leq$ 24 months, or have been referred to the programme out of suspicion of an impending psychosis.

The main criterion for the inclusion of an FEP was a diagnosis of schizophrenia or related syndromes (F20-29 in the *Tenth Revision of the International Classification of Diseases and Related Health Problems* (ICD-10)) according to both ICD-10 and *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* criteria.[29,30]

For the UHR group, patients were diagnosed according to the criteria of the Personal Assessment and Crisis Evaluation Clinic in Melbourne, the so-called Melbourne criteria for the identification of young people at incipient or 'ultra high risk' of developing a psychotic disorder.[17,31] To screen UHR patients, the ERIraos was used rather than the Comprehensive Assessment of At-Risk Mental States.[32] A score $\geq$12 on the 17-item ERIraos was deemed necessary for UHR to be further evaluated. This

threshold proved able to discriminate between patients in need for treatment and those showing symptoms that fluctuated without any serious risk of transition to psychosis.[33]

In both FEP and UHR patients, affective psychosis (bipolar disorder, or unipolar disorder with psychotic features) was an exclusion criterion, as was a co-morbid persistent substance-use dependent disorder, whereas substance use/abuse without dependence was not.

Statistical analyses

Data were analysed with the Statistical Package for Social Science for Windows (SPSS Inc., Chicago, IL, USA), version 13. All tests were two tailed, with threshold set at $P < 0.05$. All data were analysed with non-parametric statistics due to non-normality (Kolmogorov–Smirnov, with Lilliefors significance correction, $P < 0.05$ in all analyses). Categorical data were analysed in inter-group comparisons with χ^2, or Fisher exact test, when appropriate ($n < 5$ in any cell). The Mann–Whitney test for non-normal distributions was used to compare the ordinal variables. Spearman's rho correlation coefficients were used to examine associations between two continuous variables.

Scale reliability was measured by Guttman's lambda-2,[34] now regarded as a better lower bound estimate of reliability than Cronbach's alpha.[35] To compare groups, reliability values of 0.70 are considered satisfactory.[36] Test–retest stability was measured 6 months after the assessment with the intraclass correlation coefficient (ICC), with 95% confidence interval (CI). The ICC is dimensionless statistics that describes the reproducibility of repeated measurements in the same population: ICC values ≥0.60 are considered acceptable for clinical use.[37]

Discriminant validity was measured by comparison between FEP and UHR patients, with FEP patients being expected to be more severe than UHR patients. The effect sizes of the differences between FEP and UHR patients were calculated according to Cliff's delta (with CI), which is appropriate in case of violations of normality. Cliff's delta represents the degree of overlap between two score distributions.[38] It ranges from −1 to +1 (according to the order of overlap between the two groups), and takes the expected value of 0 if the two sample distributions are extracted from the same population.

Concurrent validity was measured by correlating HoNOS with BPRS and GAF, as the four HoNOS subscales are expected to relate to symptoms severity and level of functioning.

Sensitivity to change was measured with the Wilcoxon rank sum test on each HoNOS item from the assessment to the 12-month follow-up.

In FEP patients, the predictive validity of HoNOS scores at assessment was measured on the basis of remission at 12-month follow-up. According to the Remission in Schizophrenia Working Group, a patient is considered in remission when he/she scores 'mild' or better (BPRS ≤ 3) simultaneously on all seven items of BPRS that are considered to represent the core symptoms of psychosis: grandiosity, suspiciousness, unusual thought content, hallucinatory behaviour, conceptual disorganization, mannerism/posturing, and blunted affect.[39,40] A threshold score of 3 is given to low-frequency symptoms, with a minimal burden and without consequences on functioning.

In UHR patients, the predictive validity of HoNOS scores at assessment was measured on the basis of conversion to psychosis at 12-month follow-up; the McGorry *et al.*[41] criteria were used, that is, a score of 3 or higher on the hallucinations subscale, or a score of 4 or higher on the unusual thought content subscale, or a score of 4 or higher on the conceptual disorganization subscale of the BPRS; all for a duration exceeding 1 month.

Predictive validity was assessed with binary logistic regression with stepwise backward elimination (likelihood ratio), by measuring the association between baseline HoNOS total scores (continuous) and the previously described outcomes, taking into account age and gender as potential confounders. Odds ratio with 95% CI was reported.

RESULTS

The study included 87 FEP and 81 UHR patients (Table 1).

Reliability and test–retest stability

Reliability, as measured by Guttmann's lambda-2, was good for all scales (lambda-2 > 0.70), but less good for HoNOS in UHR patients (lambda-2 = 0.60).

As for test–retest of HoNOS, ICC was 0.76 (95% CI = 0.67 to 0.82) in FEP patients, and 0.64 (95% CI = 0.51 to 0.75) in UHR patients. In both cases, ICC values can be considered acceptable.

Descriptive item analysis

A lower than 1 score was found on five items in FEP patients and on eight items in UHR patients (Table 2). Floor effects (threshold > 25%) were

TABLE 1. Baseline characteristics of patients enrolled in Programma2000 (data refer to individuals in treatment until September 2009)

Variables of interest	First-episode psychosis *n* = 87	High-risk subjects *n* = 81
Age at entry	22.6 (3.8)	22.3 (3.6)
Gender (*n*, % of males)	*n* = 70 (80.5%)	*n* = 57 (70.4%)
Age of males	22.3 (3.7)	22.0 (3.7)
Age of females	23.7 (4.3)	23.0 (3.5)
Education		
College graduate or higher	6 (6.9%)	4 (4.9%)
High school diploma	37 (42.5%)	40 (49.4%)
Lower than high school diploma	44 (50.6%)	37 (45.7%)
Family psychiatric history		
Yes	58 (66.7%)	50 (61.7%)
First/second degree relative with psychosis	16 (27.5%)	9 (18.0%)
First/second degree relative with affective disorder	27 (46.5%)	24 (48.0%)
First/second degree relative with substance abuse	6 (10.3%)	8 (16.0%)
First/second degree relative with personality disorder	4 (6.9%)	6 (12.0%)
Unspecified/Unclassified	5 (8.6%)	3 (6.0%)
None	29 (33.3%)	31 (38.3%)
Duration of untreated psychosis (days)	168.8 (217.5)	–
Duration of untreated illness (months)	28.4 (20.4)	29.9 (21.9)
Clinical assessment at intake		
BPRS	53.3 (17.0)	44.8 (11.4)
BPRS nuclear symptoms of psychosis	17.9 (6.8)	12.2 (4.6)
HoNOS	15.3 (6.7)	13.0 (4.9)
GAF	43.7 (9.0)	52.9 (9.6)

All data: no. (%) or mean (SD).
BPRS, Brief Psychiatric Rating Scale; GAF, Global Assessment of Functioning; HoNOS, HoNOS, Health of the Nation Outcome Scales.

observed on eight items in FEP patients, and on nine items in UHR patients. Ceiling effects (threshold > 25%) were found on one item only in UHR patients (item 8: emotional problems and other related symptoms). Skewness was acceptable (<1) for seven items in FEP patients and five items in UHR patients. Kurtosis > 3 was evident for items 2 (non-accidental injury) and 5 (physical illness or disability problems) in both samples, and for item 3 (alcohol and/or drug misuse) in UHR patients.

Discriminant validity

As expected following the classification, at intake FEP patients were more severe than UHR patients (Table 1, at the bottom).

At the initial assessment, FEP patients scored higher than UHR patients on the BPRS ($z = -3.43$, $P < 0.001$), on the BPRS subscale on nuclear symptoms of psychosis ($z = -5.93$, $P < 0.0001$) and on the HoNOS ($z = -2.61$, $P < 0.009$), and lower on the GAF ($z = -6.09$, $P < 0.0001$).

The effect size of the differences between FEP and UHR patients was higher on the BPRS (Cliff's delta = 0.30; 95% CI = 0.13 to 0.46) and the GAF (Cliff's delta = –0.54; 95% CI = –0.67 to –0.39) than on the HoNOS (Cliff's delta = 0.23; 95% CI = 0.05 to 0.39).

Concurrent validity

Scores at intake on the HoNOS, and on its subscales 'Impairment' and 'Symptoms', were positively related to BPRS scores at intake in both FEP and UHR patients (Table 3).

Overall, the correlations between HoNOS total scores and subscales and BPRS scores were stronger in the FEP sample than in the UHR sample.

GAF scores were negatively related specifically to the 'Social problems' subscale of the HoNOS. At 12-month follow-up, the correlations between GAF and HoNOS total scores, and with the 'Symptoms and Social problems' subscale, became stronger and statistically significant in both samples.

Sensitivity to changes

At the 12-month follow-up, data were available for 63 FEP and 61 UHR patients, and missing for 24 FEP and 20 UHR patients. The reasons for missing

TABLE 2. Distribution of HoNOS items in the FEP and UHR samples at intake

	FEP patients (n = 87)					UHR patients (n = 81)				
	Mean (SD)	Floor (%)	Ceiling (%)	Skewness	Kurtosis	Mean (SD)	Floor (%)	Ceiling (%)	Skewness	Kurtosis
Behavioural problems										
1. Aggressive and disruptive behaviour	1.02 (1.11)	43.9	1.2	0.72	–0.63	0.76 (0.98)	55.8	0	0.91	–0.43
2. Non-accidental injury	0.45 (0.95)	75.6	3.7	2.43	5.67	0.37 (0.88)	80.5	2.6	2.63	6.81
3. Alcohol and (or) drug misuse	0.74 (1.15)	64.6	2.4	1.31	0.45	0.39 (0.81)	75.3	1.3	2.47	6.38
Impairment problems										
4. Cognitive and (or) scholastic problems	1.12 (1.13)	40.2	3.7	0.68	–0.36	0.57 (0.80)	61.0	0	1.09	–0.01
5. Physical illness or disability problems	0.38 (0.81)	78.0	4.9	2.18	3.83	0.36 (0.85)	80.5	1.3	2.56	6.13
Symptoms problems										
6. Hallucinations and (or) delusions	1.63 (1.24)	13.4	24.4	–0.72	–0.47	0.79 (1.08)	54.5	2.6	1.33	0.93
7. Depressed mood	1.63 (1.24)	24.4	6.1	0.14	–1.07	2.40 (1.10)	6.5	16.9	–0.38	–0.34
8. Emotional problems and other related symptoms	1.97 (1.21)	18.3	9.8	–0.24	–0.73	2.81 (0.92)	1.3	27.3	–0.33	–0.27
Social problems										
9. Problems with relationships	2.53 (1.13)	7.3	18.3	–0.71	–0.15	2.48 (1.24)	9.1	23.4	–0.52	–0.65
10. Activities of daily living and (or) self-care	1.29 (1.42)	46.3	9.8	0.61	–1.03	1.09 (1.35)	50.6	7.8	0.92	–0.49
11. Living conditions and (or) family life	0.73 (1.10)	61.0	3.7	1.46	1.33	0.50 (0.95)	71.4	1.3	1.98	3.16
12. Occupation and (or) school attendance	0.64 (1.10)	67.1	4.9	1.76	2.35	0.54 (1.04)	72.7	2.6	1.93	2.76

Floor effects occur with answers '0 = no problems' in more than 25% of the sample.
Ceiling effects occur with answers '4 = severe problems' in more than 25% of the sample.
FEP, first-episode psychosis; UHR, ultra high-risk.

TABLE 3. Matrix of correlation between the four subscales of HoNOS and BPRS and GAF scores – Spearman's rho correlation coefficient

	HoNOS behaviour	HoNOS impairment	HoNOS symptoms	HoNOS social problems	HoNOS total
FEP patients					
At intake (n = 87)					
BPRS	0.46*	0.45*	0.35*	0.57*	0.78*
BPRS core symptoms psychosis	0.40*	0.34*	0.33*	0.45*	0.67*
GAF	0.02	–0.15	–0.15	–0.26*	–0.23
At 12-month follow-up (n =63)					
BPRS	0.22	0.52*	0.60*	0.69*	0.42*
BPRS core symptoms psychosis	0.22	0.40*	0.54*	0.62*	0.58*
GAF	–0.06	–0.25	–0.42*	–0.33*	–0.38*
UHR patients					
At intake (n = 81)					
BPRS	0.18	0.33*	0.42*	0.19	0.44*
BPRS core symptoms psychosis	0.11	0.25	0.29*	0.02	0.19
GAF	.02	–0.16	–0.03	–0.36*	–0.35*
At 12-month follow-up (n =61)					
BPRS	0.38*	0.08	0.63*	0.46*	0.64*
BPRS core symptoms psychosis	0.22	–0.01	0.48*	0.30	0.47*
GAF	–0.24	–0.21	–0.36*	–0.60*	–0.57*

*$P < 0.01$.
BPRS, Brief Psychiatric Rating Scale; FEP, first-episode psychosis; GAF, Global Assessment of Functioning; HoNOS, HoNOS, Health of the Nation Outcome Scales; UHR, ultra high-risk.

data at the 12-month follow-up were the first year of treatment being not completed yet; or raters' not reporting the rating in the report (there were no dropouts in the first year of treatment). Patients with available data at 12-month follow-up did not differ from patients whose data were missing on the BPRS, HoNOS or GAF as measured at initial assessment (Mann–Whitney U-test $P > 0.05$ in all analyses in both the FEP and the UHR samples).

The FEP sample showed statistically significant changes on total HoNOS scores and on 10 out of its 12 items (Table 4).

In the UHR sample, improvement in the dimensions assessed by the HoNOS was less evident: there was a statistically significant change on the total HoNOS scores, but the changes concerned 7 items out of 12.

The items related to symptoms responsive to pharmacotherapy were more likely to change (e.g. hallucinations and/or delusions, depressed mood, emotional problems, and other related symptoms). Behavioural problems, such as non-accidental self-injury and alcohol or drug misuse, were less likely to change, particularly in the UHR sample.

Predictive validity

Predictive validity was assessed on 63 FEP and 61 UHR patients, for whom we had collected a complete dataset at the initial assessment and at the 12-month follow-up. In the FEP group, 39 patients reached remission criteria at 12-month follow-up. In the UHR group, two patients were positive for the criteria of transition to psychosis at 12-month follow-up.

HoNOS total scores at intake did not predict the FEP group's failure to reach remission (Wald test = 1.51, d.f. = 1, $P = 0.218$; OR = 1.05; 95% CI = 0.96 to 1.15). There were too few cases of transition to psychosis to assess predictive validity of HoNOS in the UHR sample.

DISCUSSION

The use of the HoNOS in an early intervention programme proved satisfactory for routine assessment. Reliability indexes were good in the FEP sample, and less good in the UHR sample (i.e. lower than the suggested threshold of 0.70),[36] although reliability values close to ours were considered acceptable in past studies. For example, in a study on FEP patients the ICC agreement between researchers and key workers on the HoNOS total scores was 0.59 (95% CI = 0.36 to 0.75).[14]

Item analysis showed that problems with alcohol or drug misuse, physical illness or disability, and non-accidental injury were rare in both FEP and UHR patients, a reflection of the composition of the

TABLE 4. Changes in total scores and individual HoNOS 12 items between baseline and 12-month assessment

	FEP patients (n = 63)		UHR patients (n = 61)	
	Mean change in score*	Wilcoxon rank test	Mean change in score*	Wilcoxon rank test
Behavioural problems				
1. Aggressive and disruptive behaviour	0.70	$z = -4.20$, $P < 0.0001$	0.20	$z = -2.12$, $P < 0.05$
2. Non-accidental injury	0.22	$z = -2.27$, $P < 0.05$	–3.32	$z = -0.92$, NS
3. Alcohol and (or) drug misuse	0.09	$z = -1.33$, NS	–0.16	$z = -1.80$, NS
Impairment problems				
4. Cognitive and (or) scholastic problems	0.28	$z = -2.44$, $P < 0.05$	0.33	$z = -3.38$, $P < 0.001$
5. Physical illness or disability problems	0.19	$z = -1.88$, NS	0.13	$z = -1.72$, NS
Symptoms problems				
6. Hallucinations and (or) delusions	1.41	$z = -5.38$, $P < 0.0001$	0.49	$z = -3.65$, $P < 0.0001$
7. Depressed mood	0.60	$z = -3.45$, $P < 0.001$	1.18	$z = -5.29$, $P < 0.0001$
8. Emotional problems and other related symptoms	0.76	$z = -3.73$, $P < 0.0001$	1.32	$z = -5.44$, $P < 0.0001$
Social problems				
9. Problems with relationships	0.66	$z = -3.88$, $P < 0.0001$	0.93	$z = -4.19$, $P < 0.0001$
10. Activities of daily living and (or) self-care	0.44	$z = -2.79$, $P < 0.01$	0.63	$z = -3.45$, $P < 0.001$
11. Living conditions and (or) family life	0.35	$z = -2.78$, $P < 0.01$	0.16	$z = -1.38$, NS
12. Occupation and (or) school attendance	0.25	$z = -2.24$, $P < 0.05$	0.25	$z = -1.89$, NS
HoNOS total	6.14	$z = -5.57$, $P < 0.0001$	5.80	$z = -5.40$, $P < 0.0001$

*Positive value means improvement.
FEP, first-episode psychosis; HoNOS, HoNOS, Health of the Nation Outcome Scales; UHR, ultra high-risk

sample that included young people who are unlikely to be affected by severe physical illness or disability, and specifically excluded those with alcohol or drug dependence, who have a higher risk of self-harm and somatic complications.[42] Indeed, this is an area of potential further research, as individuals with co-morbid drug and alcohol dependence form a challenging population, and the HoNOS is tailored to assess alcohol and drug misuse and its related problems.[43]

Overall, the HoNOS subscales proved able to assess a specific profile of symptoms in the two samples of patients, proving a helpful adjunctive measure of health status without complete overlap with other scales. Discriminant and concurrent validity were good. HoNOS differentiated between FEP and UHR patients, who are expected to show a different degree of severity concerning symptoms of psychosis; however, the effect size of the differences was smaller on the HoNOS than on the BPRS or the GAF, as a reflection of HoNOS measuring a more extended profile of psychopathology and impairment, which is more common to both conditions than severity as measured on the BPRS or social functioning as measured on the GAF. Nevertheless, HoNOS total scores and subscales correlated with other measures of psychopathology and functioning, and the correlations between HoNOS values and BPRS or GAF at 12-month follow-up were stronger than at the initial assessment, as expected for a measure of 'outcome'.

Sensitivity to change was also very good, again with differences between FEP and UHR patients that are attributable to the different profile of symptoms in the two samples.

As for predictive validity on a very specific aspect of early intervention programmes, that is, attainment of remission in FEP patients or transition to psychosis in UHR patients, the investigated samples were small, and this might have affected poor predictive validity of the HoNOS in FEP patients, and precluded analysis of transition to psychosis in the UHR sample. Moreover, we could not analyse the predictive validity of HoNOS on other indicators such as the number of admissions to psychiatric units or bed-days in psychiatric units, because the patients enrolled in the Programma2000 were rarely hospitalized.[22,23]

Overall, HoNOS possesses satisfactory sensitivity and validity to be used in routine assessment in early intervention programmes. The HoNOS is easy to use and comprehensive, but for an optimal application it requires the continuous engagement of the clinicians who complete the outcome measure. This is necessary to produce reliable and reproducible values across the samples.

ACKNOWLEDGEMENTS

The authors thank Dr Davide Sisti (Institute of Biomathematics, University of Urbino, Italy) for help in calculating the Cliff's delta.

Programma2000 is entirely financed by a grant from the Lombardy Regional Health Authority (Italy).

The Programma2000 Team includes: Prof Angelo Cocchi, MD (team leader); Anna Meneghelli, PsyD (clinical and scientific coordinator); Andrea Alpi, PsyD, Laura Bislenghi, PsyD, Tiziana Bolis, PsyD, Noelia Moreno Granados, PsyD, Giovanni Patelli, PsyD, Nicoletta Pafumi, PsyD, Maria T. Cascio, PsyD (all clinical psychologists); Federico Galvan, MA (educator); Alessia Pisano, PsyD (psychologist); Emiliano Monzani, MD; Simona Barbera, MD; Stefano Romanò MD; Daniele von Morghen, MD (all psychiatrists). The team includes a very active group of psychologists in training, too. Antonio Preti, MD, is a scientific advisor and consultant to the Programma2000 project.

Maria T. Cascio is the recipient of a grant of the Regione Sardegna (Grant n. T2-MAB-A2008-138). The Regione Sardegna had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication. No other forms of financial support were received for this study.

REFERENCES

1. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). research and development. *Br J Psychiatry* 1998; **172**: 11–8.
2. Gowers SG, Harrington RC, Whitton A *et al*. Brief scale for measuring the outcomes of emotional and behavioural disorders in children. Health of the Nation Outcome Scales for children and Adolescents (HoNOSCA). *Br J Psychiatry* 1999; **174**: 413–6.
3. Burns A, Beevor A, Lelliott P *et al*. Health of the Nation Outcome Scales for elderly people (HoNOS 65+). Glossary for HoNOS 65+ score sheet. *Br J Psychiatry* 1999; **174**: 435–8.
4. Shergill SS, Shankar KK, Seneviratna K, Orrell MW. The validity and reliability of the Health of the Nation Outcome Scales (HoNOS) in the elderly. *J Ment Health* 1999; **8**: 511–21.
5. Orrell M, Yard P, Handysides J, Schapira R. Validity and reliability of the Health of the Nation Outcome Scales in psychiatric patients in the community. *Br J Psychiatry* 1999; **174**: 409–12.
6. Preston NJ. The Health of the Nation Outcome Scales: validating factorial structure and invariance across two health services. *Aust N Z J Psychiatry* 2000; **34**: 512–9.
7. McClelland R, Trimble P, Fox ML, Stevenson MR, Bell B. Validation of an outcome scale for use in adult psychiatric practice. *Qual Health Care* 2000; **9**: 98–105.
8. Parabiaghi A, Barbato A, D'Avanzo B, Erlicher A, Lora A. Assessing reliable and clinically significant change on Health of the Nation Outcome Scales: method for displaying longitudinal data. *Aust N Z J Psychiatry* 2005; **39**: 719–25.
9. Monzani E, Erlicher A, Lora A, Lovaglio P, Vittadini G. Does community care work? A model to evaluate the effectiveness of mental health services. *Int J Ment Health Syst* 2008; **2**: 10.
10. Brooks R. The reliability and validity of the Health of the Nation Outcome Scales: validation in relation to patient derived measures. *Aust N Z J Psychiatry* 2000; **34**: 504–11.
11. Trauer T. The subscale structure of the Health of the Nation Outcome Scales (HoNOS). *J Ment Health* 1999; **8**: 499–509.
12. Newnham E, Harwood K, Page A. The subscale structure and clinical utility of the Health of the Nation Outcome Scale. *J Ment Health* 2009; **18**: 326–34.
13. Andreas S, Harfst T, Rabung S *et al*. The validity of the German version of the Health of the Nation Outcome Scales (HoNOS-D): a clinician-rating for the differential assessment of the severity of mental disorders. *Int J Methods Psychiatr Res* 2010; **19**: 50–62.
14. Amin S, Singh SP, Croudace T, Medley I, Harrison G. Evaluating the Health of the Nation Outcome Scales. Reliability and validity in a three-year follow-up of first-onset psychosis. *Br J Psychiatry* 1999; **174**: 399–403.
15. Olsen KA, Rosenbaum B. Prospective investigations of the prodromal state of schizophrenia: review of studies. *Acta Psychiatr Scand* 2006; **113**: 247–72.
16. Simon AE, Dvorsky DN, Boesch J *et al*. Defining subjects at risk for psychosis: a comparison of two approaches. *Schizophr Res* 2006; **81**: 83–90.
17. Phillips LJ, McGorry PD, Yuen HP *et al*. Medium term follow–up of a randomized controlled trial of interventions for young people at ultra high risk of psychosis. *Schizophr Res* 2007; **96**: 25–33.
18. Cornblatt B, Lencz T, Smith CW, Correll CU, Auther A, Nakayama E. The schizophrenia prodrome revisited: a neurodevelopmental perspective. *Schizophr Bull* 2003; **29**: 633–51.
19. Preti A, Cella M. Randomized controlled trials in people at ultra high risk of psychosis: a review of treatment effectiveness. *Schizophr Res* 2010; **123**: 30–6.
20. Nash L, Gorrell J, Cornish A, Rosen A, Miller V, Tennant C. Clinical outcome of an early psychosis intervention program: evaluation in a real-world context. *Aust N Z J Psychiatry* 2004; **38**: 694–701.
21. Chee KY. Determinants of subjective quality of life in first-episode schizophrenia: perspective from Malaysia. *Early Interv Psychiatry* 2010; **4**: 111–8.
22. Cocchi A, Meneghelli A, Preti A. 'Programma 2000': celebrating ten years of activity of an Italian pilot program on early intervention in psychosis. *Aust N Z J Psychiatry* 2008; **42**: 1003–12.
23. Meneghelli A, Cocchi A, Preti A. 'Programma2000': a multimodal pilot program on early intervention in psychosis underway in Italy since 1999. *Early Interv Psychiatry* 2010; **4**: 97–103.
24. Häfner H, Riecher-Rossler A, Hambrecht M. IRAOS: an instrument for the assessment of onset and early course of schizophrenia. *Schizophr Res* 1992; **6**: 209–23.
25. Overall JE, Gorham DE. The Brief Psychiatric Rating Scale. *Psychol Rep* 1962; **10**: 799–812.
26. Roncone R, Ventura J, Impallomeni M *et al*. Reliability of an Italian standardized and expanded Brief Psychiatric Rating Scale (BPRS 4.0) in raters with high versus low clinical experience. *Acta Psychiatr Scand* 1999; **100**: 229–36.
27. Moos RH, McCoy L, Moos BS. Global assessment of functioning (GAF) ratings: determinants and roles as predictors of one-year treatment outcomes. *J Clin Psychol* 2000; **56**: 449–61.

28. Lora A, Bai G, Bianchi S *et al*. La versione italiana della HoNOS (Health of the Nation Outcome Scales), una scala per la valutazione della gravità e dell'esito nei servizi di salute mentale [The Italian version of the HoNOS (Health of the Nation Outcome Scales), a scale to assess the severity and the outcome in mental health services] [Italian]. *Epidemiol Psichiatr Soc* 2001; **10**: 198–212.
29. American Psychiatric Association. *DSM-IV. Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. Washington, DC: Author, 1994.
30. World Health Organization. *Tenth Revision of the International Classification of Diseases and Related Health Problems (ICD-10)*. Geneva: Author, 1992.
31. Yung AR, McGorry PD, McFarlane CA, Jackson HJ, Patton GC, Rakkar A. Monitoring and care of young people at incipient risk of psychosis. *Schizophr Bull* 1996; **22**: 283–303.
32. Yung AR, Yuen HP, McGorry PD *et al*. Mapping the onset of psychosis: the comprehensive assessment of at-risk mental states. *Aust N Z J Psychiatry* 2005; **39**: 964–71.
33. Maurer K, Hörrmann F, Trendler G, Schmidt M, Häfner H. Früherkennung des Psychoserisikos mit dem Early Recognition Inventory (ERIraos) Beschreibung des Verfahrens und erste Ergebnisse zur Reliabilität und Validität der Checkliste [Identification of psychosis risk by the Early Recognition Inventory (ERIraos) – description of the schedules and preliminary results on reliability and validity of the checklist] [German]. *Nervenheilkunde* 2006; **25**: 11–6.
34. Guttman L. A basis for analyzing test-retest reliability. *Psychometrika* 1945; **10**: 255–82.
35. Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika* 2009; **74**: 107–20.
36. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *Br Med J* 1997; **314**: 572.
37. Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *Br Med J* 1992; **304**: 1491–4.
38. Cliff N. Dominance statistics: ordinal analyses to answer ordinal questions. *Psychol Bull* 1993; **114**: 494–509.
39. Andreasen N, Carpenter W, Kane J, Lasser R, Marder S, Weinberger D. Remission in schizophrenia: proposed criteria and rationale for consensus. *Am J Psychiatry* 2005; **162**: 441–9.
40. Lasser RA, Nasrallah H, Helldin L *et al*. Remission in schizophrenia: applying recent consensus criteria to refine the concept. *Schizophr Res* 2007; **96**: 223–31.
41. McGorry PD, Yung AR, Phillips LJ *et al*. Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. *Arch Gen Psychiatry* 2002; **59**: 921–8.
42. Verdoux H, Liraud F, Gonzales B, Assens F, Abalan F, van Os J. Suicidality and substance misuse in first-admitted subjects with psychotic disorder. *Acta Psychiatr Scand* 1999; **100**: 389–95.
43. Carr JA, Norman RG, Manchanda R. Sociodemographic and clinical characteristics of patients presenting with first-episode psychosis and concurrent substance misuse. *Early Interv Psychiatry* 2009; **3**: 75–9.