Research Article   **Open Access**

Presentation of the HoNOS-Secure French Version, an Outcome Measure for Users of Secure and Forensic Mental Health Services

Eytan A[*1], Golay D[2], Della Polla S[2], Hofer F[2], Sillitti F[2] and Gex-Fabry M[3]

[1]Division of General Psychiatry, Geneva University Hospitals, Geneva, Switzerland
[2]Division of Penitentiary Psychiatry, Geneva University Hospitals, Geneva, Switzerland
[3]Division of Psychiatric Specialties, Geneva University Hospitals, Geneva, Switzerland

[*]Corresponding author: Eytan A, Division of General Psychiatry, Geneva University Hospitals, Geneva, Switzerland, Tel: +41 22 3054755, E-mail: ariel.eytan@hcuge.ch

Abstract

Background: Despite a number of instruments designed for the assessment of dangerousness, standardized outcome measures for forensic psychiatric patients are lacking, especially in other languages than English.

Aim: To validate the French version of HoNOS-secure, the Health of the Nation Outcome Scales for users of forensic services.

Method: The psychometric properties of HoNOS-secure-French were evaluated in 66 patients rated independently by two clinicians on two occasions, 6 to 8 weeks apart. There were 7 raters in total, both psychiatrists and psychologists.

Results: Internal consistency was acceptable for all 4 assessments of the HoNOS security scale and 3 of 4 assessments of the clinical scale (Cronbach's alpha ≥ 0.70). Inter-rater reliability of single ratings was moderate for the security scale (ICC 0.48-0.56) and fair to moderate for the clinical scale (ICC 0.37-0.60). Test-retest reliability was substantial for both the security (ICC 0.69) and clinical scales (ICC 0.73) when patients were rated by their attending therapists. It was substantial for the security scale (ICC 0.80), but only fair for the clinical scale (ICC 0.28) when raters were less familiar with the patient's condition. Increased needs for secure measures (security scale) and for care (clinical scale) were significantly associated with higher severity of psychopathology (Clinical Global Impression) and lower self-perceived physical health (Medical Outcomes Study 36-Item Short Form Health Survey).

Conclusions: The value of the HoNOS-secure-F for routine outcome monitoring deserves closer examination. Some of its psychometric properties might be less than optimal. Therefore, we cannot recommend relying solely on its scores for medico-legal decisions.

Keywords: HoNOS-secure-F; Mental disorders; Medico-legal decisions

Introduction

Many individuals with mental disorders are detained either in prisons or in secure psychiatric facilities throughout the world. Others are free to come and go but are submitted to compulsory psychiatric treatments by order of justice. Typically, these forensic patients have been found guilty of a serious offence or crime because of an irresponsible behavior caused by a mental disorder. These patients often present complex psychopathologies, with co-morbid substance misuse and/or personality disorder [1].

A number of studies conducted in the last four decades have shown a higher prevalence of mental disorders among prisoners, compared with the general population. In recent years, there was a significant increase of the number of medium and high secure psychiatric beds in most European countries [2]. Consequently to security driven policies, this is occurring also in French speaking countries such as Switzerland and France [3].

Regarding standardized assessment in the field of forensic psychiatry, there is an abundant multinational literature on the issue of dangerousness and several risk assessment tools exist [4]. Most of these instruments address long-term dangerousness, but short-term risk assessment measures aiming at the improvement of clinical decision-making are lacking [5]. There are few routine instruments for assessing clinical outcome of secure services' users and, to our knowledge, none was designed or translated in French. For general psychiatry, outcome measurement was implemented in Switzerland at a national level in 2012. The Health of the Nation Outcome Scales (HoNOS) were chosen for this purpose [6]. HoNOS is a diagnosis-independent scale for mental health and social functioning. It was first developed and validated in the UK for adult psychiatric patients [7] and became widely used in Europe [8]. Adaptations of HoNOS were designed for specific clinical populations such as children and adolescents [9] or elderly people [10]. HoNOS-secure is the declination of the instrument for forensic patients [11]. The translation and psychometric evaluation of a new French version of HoNOS-secure is described in the present paper.

Method

Setting

The study took place in three facilities located in Geneva, in the French speaking region of Switzerland: two medical units that serve a remand detention centre and a jail, respectively, and a secure mental health unit for detainees presenting with acute or sub-acute psychiatric disorders. All participants were psychiatric patients in one of these three facilities. Participation in the study was proposed to patients who had a sufficient knowledge of the French language. Selection was opportunistic, based on the therapeutic alliance between caregivers and patients. The study included 66 participants over a 12-month period.

The study protocol was approved by the Ethics Committee of Geneva University Hospitals. Written informed consent was obtained from all participants.

HoNOS-secure French version

HoNOS-secure consists of a seven-item security scale measuring the need for secure measures, plus the 12 original HoNOS items measuring the need for care, modified to account for secure settings. While HoNOS items are rated retrospectively for the observed problem behavior, the security scale is rated prospectively for the period "in the near future, including if living unsupported in the community" [11]. All items are rated from 0 to 4, with total scores in the ranges 0-28 for the security scale and 0-48 for the clinical scale.

A validated French translation was already available for the original HoNOS [12]. We kept the 13th item proposed in the French version, which measures problems with drug management and medication adherence [13]. In the English version of HoNOS-secure, Dickens and colleagues modified some items of the original HoNOS in order to account for secure settings (e.g., in scale 12, point 4 by adding an example of enforced inactivity: being in a prison cell) [11]. We modified similarly the original French HoNOS. The modified clinical items and the secure items were translated independently by three of the authors (SDP, AE, DG). Discrepancies were resolved in a second step by consensus, to obtain the HoNOS-secure-F used in the present validation study (available upon request to the corresponding author). One of the authors (AE) translated the glossary and training material.

Instruments

Psychiatric diagnosis was assessed with the Mini International Neuropsychiatric Interview (MINI). The MINI is a standardized, structured diagnostic interview, which has been tested against the Structured Clinical Interview for DSM-III-R (SCID) and the Composite International Diagnostic Interview for ICD-10 (CIDI) and found to be reliable and valid [14]. The MINI includes suicide risk assessment (rated as absent, low, moderate or high).

Subjective physical and mental health was assessed using the Medical Outcomes Study 36-Item Short Form Health Survey (SF-36), which considers a 4-week time frame for most questions [15]. The SF-36 was translated into several languages, including French [16]. It allows calculating a physical component summary score (PCS) and a mental component summary score (MCS), which are norm-based scores, i.e. values 10 points below or above 50 represent differences of one standard-deviation from average values in a US reference population.

Severity of psychopathology was assessed with the Clinical Global Impression Severity score (CGI-S), rated on a seven-point scale (from 1= normal, not at all ill to 7= among the most extremely ill patients). The CGI was introduced as a brief, clinician-rated summary measure to be used before and after treatment (Guy 1976). Because it is easy and quick to administer, it has been widely used in different settings and psychiatric conditions.

The acceptability, feasibility and face validity of the HoNOS-secure-F were addressed through 4 questions to raters: the time needed to fill out the scales, the perceived difficulty with the instrument (very easy, rather easy, rather difficult, very difficult) and the degree to which they felt the security and clinical scales reflected the actual severity of the patient's condition (largely underestimated, truly reflected, largely overestimated).

Procedure

Each participant was assessed with the HoNOS-secure-F independently by two mental health professionals (raters 1 and 2) on two occasions (T1 and T2), within a time interval of 6 to 8 weeks. Both clinicians rated the CGI-S and answered the acceptability questions on the two occasions. Participants rated their self-perceived mental and physical health with the SF-36 at T1. Between T1 and T2, a research psychologist (SDP, FH) conducted the MINI diagnostic interviews and recorded socio-demographic characteristics of participants, as well as type of offence.

Raters

Seven mental health professionals participated to the assessments, including 4 psychiatrists and 3 psychologists. The first rater (rater 1) was generally the attending therapist (98.5% of ratings), a psychiatrist (56.1%) or psychologist (43.9%) who knew the patient since at least one month (95.5%). The second rater (rater 2) was another professional working in the facility, generally a psychologist (97.0%) who knew the patient since less than one month (90.9%).

Two of the authors (AE and DG) participated in the "training for the trainers" for HoNOS-secure in the UK. In Switzerland, they trained the other professionals involved in the project.

Statistical analysis

Internal consistency of the HoNOS-secure-F was assessed with the Cronbach's alpha coefficient, which is generally considered as acceptable when equal or larger than 0.70 [17]. Interrater reliability was examined with the intra-class correlation coefficient (ICC). Because each patient was rated by a different subset of clinicians, a one-way random model was used. We considered both single measure agreement, which generalizes to other possible individual raters, and average measure agreement, which refers to the mean rating of two raters and generalizes to other possible sets of two raters [18]. These two aspects complement each other, because the instrument might be used either by a single rater as part of routine clinical practice or by two raters when a clinical decision is based on the mean rating of two observers in a multidisciplinary team. Landis and Koch [19] provided guidelines for interpreting agreement, with 0 to 0.20 indicating slight agreement, 0.21 to 0.40 indicating fair agreement, 0.41 to 0.60 indicating moderate agreement, 0.61 to 0.80 indicating substantial agreement, and 0.81 to 1 indicating almost perfect agreement. Test-retest reliability was assessed with the ICC, considering a two-way random model and single measure agreement. Associations between the HoNOS-secure-F, SF-36 and CGI scores were evaluated with Spearman's rank correlation coefficients. Differences between groups were examined with the Mann-Whitney U test or the Kruskal-Wallis test. All statistical analyses were conducted using SPSS version 20 (IBM Corporation, Armonk, NY, USA). All tests were two-tailed, with significance level at 0.05.

Results

Patient characteristics

Characteristics of patients are presented in Table 1. Most of them were males (95.5%) born outside Switzerland (71.2%), who had committed violent offences (68.2%) and had been sentenced (56.1%). Mood disorders were the most frequent psychiatric diagnoses (71.2%), with 28.8% of participants considered at moderate to high suicide risk.

Characteristic		frequency	%
Gender	Male	63	95.5
	Female	3	4.5
Age	18 – 29	26	39.4
	30 – 39	21	31.8
	40 – 60	19	28.8
Education (n=65)	Primary school or less	9	13.8
	Secondary school	24	36.9
	Vocational school	27	41.5
	University or equivalent	5	7.7
Place of birth	Switzerland	19	28.8
	Other European country	21	31.8
	North Africa	13	19.7
	Sub-Saharan Africa	7	10.6
	Latin America or Caribbean	6	9.1
Court judgment	Yes	37	56.1
	No	29	43.9
Violent offence	Yes	45	68.2
	No	21	31.8
Sexual offence	Yes	16	24.2
	No	50	75.8
Psychiatric diagnosis[a]	Mood disorders[b] F30 – F39	47	71.2
	Anxiety disorders[b] F40 – F43	29	43.9
	Psychotic disorders[b] F20 – F29	12	18.2
	Substance use disorders[c] F10 – F19	29	43.9
	Current suicide risk (moderate to high)	19	28.8

[a]According to the Mini International Neuropsychiatric Interview (MINI), with corresponding ICD codes
[b]Present or past
[c]Present
Table 1: Patient characteristics (n=66)

Acceptability, feasibility and face validity

When asked about their perception of the HoNOS-secure-F, a majority of clinicians rated the instrument as easy or very easy to use. Approximately 10% of all ratings at T1 were perceived as rather difficult (12.1% for rater 1, 10.6% for rater 2). More than 80% of ratings were considered to actually reflect the severity of the patient's condition, both for security and clinical scales. The time needed to fill the HoNOS-secure-F ranged from 2 to 30 minutes, with median values of 15 min and 6 min for raters 1 and 2, respectively.

Item and total score distributions

For the security scale (rater 1 at T1) item scores ranged from 0 to 4 for 6 of 7 items, with mean values between 0.7 and 1.4. Mean total score was 8.3 (standard deviation 6.0, range 0-24). For the clinical scale, item scores were in the range 0-4 for 5 of 13 items, with mean values between 0.2 and 1.5. Mean total score was 8.4 (standard deviation 5.5, range 1-26). No floor or ceiling effect was observed for the security and clinical total scores, with less than 10% of observations at the lowest or highest possible scores.

Internal consistency

The internal consistency of the HoNOS security scale was considered acceptable for all 4 assessments (2 raters, 2 time points), with Cronbach's alpha values between 0.71 and 0.80. Alpha values were in the range 0.72 to 0.75 for 3 out of 4 assessments of the clinical scale, but only 0.50 for the assessment performed by rater 2 at T2.

Inter-rater reliability

Inter-rater reliability is summarized in Table 2. It was fair to moderate for a majority of items in the security and clinical scales. It was poor (≤0.20) on at least one occasion for one item in the security scale (need for building security to prevent escape) and 5

HoNOS-secure items	Inter-rater reliability ICC		Test-retest reliability ICC	
	T1 (n=66)	T2 (n=57)	Rater 1[a] (n=56)	Rater 2[b] (n=56)
Security scale				
Risk of harm to adults or children	0.60	0.43	0.68	0.76
Risk of self-harm (deliberate or accidental)	0.41	0.24	0.60	0.75
Need for building security to prevent escape	0.14	0.24	0.66	0.51
Need for a safely-staffed living environment	0.41	0.38	0.77	0.63
Need for escort on leave (beyond secure perimeter)	0.53	0.61	0.65	0.32
Risk to individual from others	0.40	0.31	0.61	0.59
Need for risk management procedures	0.38	0.48	0.72	0.76
Total score (single rating)	0.48	0.56	0.69	0.80
Total score (average of two ratings)	0.65	0.72		
Clinical scale				
Overactive, aggressive, disruptive or agitated behaviour	0.65	0.44	0.60	0.32
Non-accidental self-injury	0.58	0.20	0.50	0.09
Problem drinking or drug taking	0.26	0.39	0.77	0.35
Cognitive problems	0.13	0.67	0.59	0.55
Physical illness or disability problems	0.21	0.08	0.62	0.33
Problems associated with hallucinations and delusions	0.69	0.33	0.84	0.38
Problems with depressed mood	0.53	0.46	0.37	0.15
Other mental and behavioural problems	0.46	0.30	0.47	0.22
Problems with relationships	0.15	-0.06	0.31	0.17
Problems with activities of daily living	0.26	0.39	0.53	0.23
Problems with living conditions	0.13	0.14	0.50	0.29
Problems with occupation and activities	0.25[c]	-0.09	0.75[c]	0.34
Problems with drug management	0.25	0.44[c]	0.51[c]	0.30
Total score (single rating)	0.60	0.37	0.73	0.28
Total score (average of two ratings)	0.75	0.54		

[a]Psychiatrist or psychologist who was patient's attending therapist
[b]Psychologist who was not patient's attending therapist
[c]n=1 missing value
ICC : intra-class correlation coefficient
Table 2: Inter-rater reliability and test-retest reliability of the HoNOS-secure-F

items of the clinical scale (cognitive problems, physical illness or disability problems, problems with relationships, problems with living conditions and problems with occupation). When considering interchangeability of raters (single measure agreement), inter-rater reliability was moderate for the security total score and fair to moderate for the clinical total score. When the mean rating of the two observers was considered, reliability was substantial for the security scale and moderate to substantial for the clinical scale.

Test-retest reliability

For attending therapists, test-retest reliability over a 6-8 weeks period was substantial for all items in the security scale and moderate to substantial for all but 2 items in the clinical scale (Table 2). As a result, test-retest reliability was substantial for both the security total score (ICC 0.69) and clinical total score (ICC 0.73). When considering clinicians who were less familiar with the patient's condition, test-retest reliability remained substantial for the security total score (ICC 0.80), but was only fair for the clinical total score (ICC 0.28).

Associations with SF-36 and CGI scores

Associations between the HoNOS-secure-F and the SF-36 and CGI scores are presented in Table 3. Increased need for secure measures, as estimated with the security scale by both raters, was significantly associated with poorer self-perceived SF-36 physical health and higher CGI severity score, whereas no significant association was observed with the SF-36 mental health dimension. In keeping with expectations, a higher need for care, as measured with the HoNOS clinical scale, was significantly associated with lower SF-36 physical and mental health scores and higher severity on the CGI scale.

	n	HoNOS-secure security scale		HoNOS-secure clinical scale	
		Spearman correlation coefficient	P-value	Spearman correlation coefficient	P-value
Rater 1 (at T1)					
SF-36 Physical component summary	65	-0.31	0.012	-0.45	<0.001
SF-36 Mental component summary	65	-0.16	0.217	-0.30	0.014
CGI – Severity scale	65	0.71	<0.001	0.54	<0.001
Rater 2 (at T1)					
SF-36 Physical component summary	65	-0.39	0.001	-0.40	0.001
SF-36 Mental component summary	65	-0.05	0.707	-0.25	0.046
CGI – Severity scale	66	0.68	<0.001	0.52	<0.001

SF-36: Medical Outcomes Study 36-Item Short Form Health Survey
CGI: Clinical Global Impression
Table 3: Associations between the HoNOS-secure-F and the SF-36 and CGI scores

Differences between groups

HoNOS-secure-F security and clinical scores were investigated for possible differences according to age, history of violent offence, duration of detention and psychiatric diagnosis (Table 4). Need for secure measures, as estimated by rater 1 at T1, was significantly higher for patients who met diagnostic criteria for mood disorders or disorders due to substance use. These effects were not independent from each other, because comorbidity was frequent, with 75.9% of patients with substance-related disorders presenting mood disorders as well. Security score was not associated with violent offence and duration of detention. Estimated need for care was also significantly higher for patients with mood disorders or disorders due to substances, whereas it was significantly lower for participants whose detention had been longer.

Discussion and Conclusion

The current study aimed to test and validate the HoNOS-secure locally translated into French. Given the absence of a gold standard for measuring outcome among French speaking forensic patients, validity of the HoNOS-secure-F was assessed through inter-rater reliability, test-retest reliability and comparison of scores with participants' self perceived health on the SF-36 and clinicians' global clinical impression. In a study conducted in a private psychiatric clinic in Australia, the HoNOS and the SF-36 were concordant in providing reliable and valid measures of aspects of patient function [20].

Although the selection of subjects was based on their willingness to participate to the study, their characteristics were similar to the profile of the population treated for mental health problems at our regional detention facilities: young, male, non-Swiss, low level of education and affective disorder being the most frequent diagnosis [21]. We did not record the specific characteristics of non-participants in the present study.

		n	HoNOS-secure security scale[a]			HoNOS-secure clinical scale[a]		
			median	range	P-value[b]	median	range	P-value[b]
Age	18 – 29	26	9	0 - 20	0.29	9	1 - 26	0.25
	30 – 39	21	7	0 - 18		7	1 - 25	
	40 – 60	19	5	0 - 24		6	2 - 23	
Violent offence	Yes	45	10	0 - 24	0.23	7	1 - 26	0.11
	No	21	6	0 - 18		8	2 - 25	
Length of detention at the time of the study	< 12 months	35	8	0 - 24	0.43	8	1 - 26	0.035
	≥ 12 months	31	5	0 - 20		6	1 - 17	
Mood disorders (present or past)[c]	Yes	47	11	0 - 24	0.001	8	1 - 26	0.003
	No	19	4	0 - 13		4	1 - 11	
Disorders due to substance use (present)[c]	Yes	29	11	0 - 20	0.003	10	1 - 26	0.004
	No	37	5	0 - 24		6	1 - 23	
Current suicide risk[c]	Moderate to high	19	7	1 - 24	0.38	9	4 - 26	0.058
	Absent to low	47	7	0 - 20		7	1 - 25	

[a]Assessed by rater 1 at T1
[b]Mann-Whitney U test or Kruskal-Wallis test
[c]According to the Mini International Neuropsychiatric Interview (MINI)
Table 4: Differences between groups on the HoNOS-secure-F

The score distributions, the absence of floor or ceiling effects and the internal consistency results point toward a satisfying intrinsic structure of the HoNOS-secure-F. Indeed, internal consistency was acceptable for all 4 assessments of the security scale and 3 of 4 assessments of the clinical scale.

Inter-rater reliability was non-homogeneous across items of both the security and clinical scales. In the validation study of the original HoNOS-secure, inter-rater ICC values were similarly comprised between 0.39 and 0.88 for security items and between 0.29 and 0.96 for clinical items [11]. Reliability of the security total score was nevertheless substantial in the present study, when focus was placed on the mean rating of two observers with different background and experience. This result indicates that the HoNOS-secure-F could be used by several raters in order to reach a consensus for a given patient, as this is often an objective in a multidisciplinary team.

Regarding the total scores on security and clinical scales, 3 out of 4 test-retest reliability coefficients were in accordance with the recommended standards. Indeed, ICC values comprised between 0.6 and 0.8 are often used as the minimum standards for reliability [22]. Only the clinical total score rated by a clinician less familiar with the patient was below 0.6 (ICC 0.28).

As expected, high scores on the clinical scale of the HoNOS-secure-F were significantly associated with poor perceived physical and mental health, and with high severity on the CGI. The association between high scores on the security scale, poor self-perceived physical health and increased clinician-rated severity was more surprising. A possible explanation might be the frequent comorbidity of mood disorders, suicide risk and substance abuse, which might influence both the clinical global impressions and the perceived needs for security measures. Another issue might be the difficulty for therapists to rate the needs for secure measures independently from the needs for care.

Rating differences between the attending therapist and the second rater, who was less familiar with the patient's condition, need to be pointed out. The instrument performances were globally satisfying when used by the clinician in charge of the patient, but problematic when used by a second rater with more superficial information about the patient. This was also apparent in the validation study of the HoNOS-65+F, in which the type of clinical setting and length of the patient-caregiver relationship were the main determinants of inter-rater reliability for the instrument [23]. This is not surprising, since Bebbington, *et al.*, in the initial validation study of HoNOS, already showed that there were serious problems in using the instrument as a routine measure of clinical status in busy psychiatric services [24]. In this original study, the performance of HoNOS appeared to be closely related to the training and experience of key workers.

The present study has several limitations. The number of participants was small and the design did not allow obtaining a representative sample of the forensic population treated at our psychiatric facilities. Sensitivity to change was not assessed, in contrast with other studies addressing the utility and validity of HoNOS-secure in male [25] and female [26] detainees. Because of the setting and its constraints, it was practically too complicated to assess systematically subjects before and after treatment in our study. This parameter should be monitored in further studies. Convergent validity with other instruments designed to assess either risk of aggression, such as the Dynamic Appraisal of Situational Aggression (DASA) [27] or security needs, such as the Security Needs Assessment Profile (SNAP) [28], was not measured. The DASA is supposed to be used with psychiatric inpatients, while the

SNAP takes into account specific procedural security items which are usually under the responsibility of the prison administration. Therefore, none is completely convergent with HoNOS. We did not assess personality with structured or semi-structured instruments, knowing from previous studies that the prevalence of personality disorders is high in this population [21].

The HoNOS family of measures was developed in order to monitor clinical outcome of patients with severe mental illness in the UK. These tools were part of a national strategy for improving mental health care at a national level. Therefore managerial preoccupations were also present [29]. HoNOS addresses a wide range of problems [30] and appears to be globally appropriate for routinely monitoring outcome, despite uneven psychometric properties of items [31]. Since the original validation study, HoNOS assessments made by individual key workers were shown to be of limited value and input from several involved members of the mental health team was recommended [24].

As underlined by Long, *et al*, most outcome studies conducted in forensic services usually address long term issues such as readmission or recidivism [26]. Research has confirmed that long-term risk assessment and management is a complex task and standardized instruments cannot be used as sole determinants of detention, sentencing or release from a secured environment [4]. Compared with such instruments, the HoNOS-secure is supposed to fulfill a somewhat simpler mission. Indeed, it was designed with the objective of facilitating ongoing clinical assessment and short term risk monitoring of forensic patients. In this perspective, the seven security items (A-G scales) were added to the 1-13 clinical scales of the general HoNOS [26].

We observed in the present study that the performances of the HoNOS-secure-F are only fair, despite some interesting psychometric properties. Therefore, we cannot recommend its use for justifying medico-legal decisions. We believe that the HoNOS-secure-F is relevant for routine clinical assessment of forensic patients, insofar as several members of the team rate the scales together, confront their ratings and consider the tool as a checklist and a support for clinical multidisciplinary discussions.

Acknowledgment

References

1. Doyle M, Coid J, Archer-Power L, Dewa L, Hunter-Didrichsen A, et al. (2014) Discharges to prison from medium secure psychiatric units in England and Wales. Br J Psychiatry 205: 177-82.

2. Gordon H, Lindqvist P (2007) Forensic psychiatry in Europe. Psychiatr Bull 31: 421-4.

3. Velpry L, Eyraud B (2014) Confinement and psychiatric care: a comparison between high-security units for prisoners and for difficult patients in France. Cult Med Psychiatry 38: 550-77.

4. Fazel S, Singh JP, Doll H, Grann M (2012) Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: systematic review and meta-analysis. BMJ 345: e4692.

5. van de Sande R, Nijman HL, Noorthoorn EO, Wierdsma AI, Hellendoorn E, et al. (2011) Aggression and seclusion on acute psychiatric wards: effect of short-term risk assessment. Br J Psychiatry 199: 473-8.

6. Koch N, Wasserfallen Jean-B, Camus D, Gebel S, Bonsack C (2013) Implementing validated routine outcome assessment in psychiatry: A field study. Serveur académique lausannois 164: 134-44.

7. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, et al. (1998) Health of the Nation Outcome Scales (HoNOS). Research and development. Br J Psychiatry 172: 11-8.

8. Stephen K, Leslie AC, Jennifer C, Martha C, John C, et al. (2010) Do the Health of the Nation Outcome Scales measure outcome? Can J Psychiatry 55: 431-9.

9. Gowers SG, Harrington RC, Whitton A, Lelliott P, Beevor A, et al. (1999) 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 174: 413-6.

10. Burns A, Beevor A, Lelliott P, Wing J, Blakey A, et al. (1999) Health of the Nation Outcome Scales for elderly people (HoNOS 65+). Br J Psychiatry 174: 424-7.

11. Dickens G, Sugarman P, Walker L (2007) HoNOS-secure: A reliable outcome measure for users of secure and forensic mental health services. J Forensic Psychiatry Psychol 18: 507-14.

12. Lauzon S, Corbière M, Bonin JP, Bonsack C, Lesage AD, et al. (2001) Validation of the French version of the Health of the Nation Outcome Scales (HoNOS-F). Can J Psychiatry 46: 841-6.

13. Bonsack C, Borgeat F, Lesage A (2002) Mesurer la sévérité des problèmes des patients etleurévolutiondans un secteurpsychiatrique: Uneétudesur le terrain du Health of Nation Outcome Scales en français: (HoNOS-F). = Measuring patients' problems severity and outcomes in a psychiatric sector: A field study with the French version of the Health of Nation Outcome Scales (HoNOS-F). Annales Médico-Psychologiques 160: 483-8.

14. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, et al. (1998) The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 20: 22-33, quiz 34-57.

15. Ware JE, Sherbourne CD (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 30: 473-83.

16. Leplège A (2001) Le questionnaire MOS SF-36, manuel de l'utilisateur et guide d'intrprétation des scores. Paris: ESTEM éditions.

17. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, et al. (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 60: 34-42.

18. Shrout PE (1979) Fleiss, Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86: 420-8.

19. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33: 159-74.

20. Page AC, Hooke GR, Rutherford EM (2001) Measuring mental health outcomes in a private psychiatric clinic: Health of the Nation Outcome Scales and Medical Outcomes Short Form SF-36. Aust N Z J Psychiatry 35: 377-81.

21. Eytan A, Haller DM, Wolff H, Cerutti B, Sebo P, et al. (2011) Psychiatric symptoms, psychological distress and somatic comorbidity among remand prisoners in Switzerland. Int J Law Psychiatry 34: 13-9.

22. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, et al. (2011) Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 64: 96-106.

23. Canuto A, Weber K, Gold G, Notaridis G, Michon A, et al. (2007) Structured assessment of mental health status in psychogeriatrics: validity of the French HoNOS 65+. Can J Psychiatry 52: 37-45.

24. Bebbington P, Brugha T, Hill T, Marsden L, Window S (1999) Validation of the Health of the Nation Outcome Scales. Br J Psychiatry 174: 389-94.

25. Dickens G, Sugarman P, Picchioni M, Long C, et al. (2010) HoNOS-Secure: Tracking risk and recovery for men in secure care. The British J Forensic Practice 12: 36-46.

26. Longa CG, Dickensa G, Sugarmana P, Lorraine (2010) Tracking risk profiles and outcome in a medium secure service for women: Use of the HoNOS-Secure. The Int J Forensic Mental Health 9: 215-25.

27. Ogloff JR, Daffern M (2006) The dynamic appraisal of situational aggression: an instrument to assess risk for imminent aggression in psychiatric inpatients. Behav Sci Law 24: 799-813.

28. Collins M, Davies S, Ashwell C (2003) Meeting patients' needs in secure forensic psychiatric units. Nurs Stand 17: 33-4.

29. Stein GS (1999) Usefulness of the Health of the Nation Outcome Scales. Br J Psychiatry 174: 375-7.

30. Orrell M, Yard P, Handysides J, Schapira R (1999) Validity and reliability of the Health of the Nation Outcome Scales in psychiatric patients in the community. Br J Psychiatry 174: 409-12.

31. Pirkis JE, Burgess PM, Kirk PK, Dodson S, Coombs TJ, et al. (2005) A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures. Health Qual Life Outcomes 3: 76.