CHAPTER VI

DEVELOPMENT OF MENTAL HEALTH MEASUREMENT

This chapter is to present the development of mental health measurement, the Health of the Nation Outcome Scale (HoNOS), as a tool to collect clinical data to construct a classification model. Many steps in the development process were implemented to test the new measurement. The chapter covers selected criteria for mental health measurement for allocation, materials and methods, results of measurement psychometric properties test, discussions, and conclusions.

1. Targeting mental health measurement for budget allocation

An outcome measurement is one group of measurements that play an important role in mental health care and research. It enhances quality of psychiatric service, increases service efficiency, controls service costs, introduces innovation, and demonstrates accountability (Rush et al., 2000). Especially, outcome measurement can be used as a tool for budget allocation.

Generally, there are three ways to find a measurement: *(1) Self development:* by this method, the measurement will be developed to meet developer's objectives but it will need more budget, tremendous time, and, very importantly, high caliber developer's skills. *(2) Adopt foreign standard measurement without adjustment:* this method needs lower budget compared to *self development* but translation validity and difference in context and culture should be aware of. *(3) Adopt foreign standard measurement with adjustment:* after selecting and translating the measurement, the new tool should follow through standard procedures and finally test for psychiatric qualities. Similar to the second alternative, this method is also comparatively cheap (Guillemin, Bombardier & Beaton, 1993); (Greco, Walop & Eastridge, 1987, pp. 817-845); (Greco, Walop & McCarthy, 1987, pp. 699-701).

The results from the chapter V show that outcome measurements such as Global Assessment of Functioning (GAF), Clinical Global Impression scale CGI), and the Health of the Nation Outcome Scale (HoNOS) can be used as a tool for budget allocation. This study selected to adopt the Health of the Nation Outcome Scale (HoNOS) on the basis of objectivity of outcome measurement, content, rater, time used and suitability to, adult psychiatric. HoNOS was developed into Thai version to measure health outcomes and to provide data on clinical progress. HoNOS has been used for clinical audit in the Care Program Approach (CPA), and proved to reflect patients' needs, matched with practitioners' skills in casemix and caseload (Wing et al, 1998, pp. 11-18), and passed satisfactory level in both clinical and administrative requirements. HoNOS provides more comprehensive picture of mental health outcomes than any other brief measurement (Orrell et al., 1998, pp. 409-412). HoNOS was first translated into Thai, and tested its psychometric properties.

2. Materials and Methods

2.1 The Health of the Nation Outcome Scales (HoNOS)

The HoNOS (Wing, Curtis & Beevor, 1990) provides more comprehensive mental health outcomes than any other brief measurement (Orrell et al., 1998, pp. 409-412). Clinical research also benefit (Wing et al, 1996).

The HoNOS were designed for a brief general assessment of functioning. It was intentionally designed for routine work to be handled by mental health staff to fill in easily; and to be useful for everyday clinical work (Orrell et al, 1998, pp. 409-412). This measurement, consisting of 12 rating scales, covers mental and physical health and social functioning. Each scale has a score from 0 (no problem) up to 4 (very severe problem). Score can be completed in a few minutes by clinicians with routine assessments (Wing et al, 1998, pp. 11-18). The 12 scales cover four areas, including behavioral problems, impairment, symptomatic problems, and social problems. It is designed to be used with secondary mental healthcare services. It covers clinical and social areas relevant to adult mental illness, is appropriate for routine practical use in any.

setting, provides a brief numerical record of the clinical assessment, and has a variety of uses for clinicians, administrators and researchers. The HoNOS has to be employed at the beginning of admission and the end of care (discharge). In an acute patient setting, the domestic situation is not often known to the staff, so items 11 and 12 are not rated. Additionally, the patient should be rated for every 3 to 6 months but not longer than annually. It is suitable for routine use by nurses and psychiatrists (Rush et al., 2000); (Bebbington, et al, 1999, pp. 389-94); (Amin, et al, 1998. pp. 399-403); (Wing et al, 1998, pp. 11-18); (Wing et al, 1996). When compared with the Brief Psychiatric Rating Scale (BPRS) (Overall & Gorham, 1962, pp. 799–812), the HoNOS is simpler and relatively easier to use. Up to now, HoNOS has identified five outstanding characteristics as follows: short enough for routine use, coverage of common clinical problems and social functioning, sensitive to change, reliable, and highly correlated with established scales (Sharma, Wilkinson & Fear, 1999, pp. 395-98). The HoNOS scores show a strong association with service utilization, and therefore is likely to play a major role in casemix systems (Trauer, Callaly, Hantz, Little, Shields & Smith, 1999, pp. 380-88); (Orrell et al., 1998, p. 409-12). In Australia, HoNOS was used in developing a casemix classification for mental health service costs (MH-CASC) (Buckingham et al, 1998).

2.2 Translation Process

After granted an official permission to translate HoNOS into Thai by the Royal College of Psychiatrists' Research Unit, London, translation procedures as recommended by Greco and colleagues were carefully applied (Grecoet et al., 1987, pp. 699-701); (Greco et al.,1987, pp. 817-845). Two bilingual mental health experts and Thai native speakers living in the UK and the US for more than five years translated the original HoNOS into Thai. Then the other two bilingual non-mental health specialists back-translated the Thai HoNOS into English. One back-translator was an English native speaker who was fluent in Thai and had lived in Thailand for more than 5 years. The other back-translator was a Thai native speaker who was fluent in English and had lived in the UK for more than 5 years. The first draft of Thai HoNOS was pre-tested by six psychiatric nurses to identify the statements difficult to understand.

Then, the English version that was back-translated from Thai HoNOS draft, and the opinions of field testing nurses were presented to an expert panel comprising researchers, translators, back-translators, psychiatric experts, language experts, and psychiatric nurses. The expert panel commented on all aspects of the Thai HoNOS draft. The comments were reasonably positive. Their suggestions and concerns were similar to those presented by Orrell et al. (Orrell et al., 1998, pp. 409-412).

Items 11 (accommodation) and 12 (occupational problems) of the HoNOS were also translated, but were excluded from reliability and validity tests and factor analysis studies, (Trauer, et al, 1999, pp. 380-8); (Wing et al, 1996;, became these items cannot be used to evaluate hospitalized patients.

2.3 Subjects for testing psychometric properties

The subjects for testing psychometric properties were selected from inpatient at Suanprung psychiatric hospital. Group 1 comprised 23 acute patients, and group 2 was 23 sub-acute patients who were waiting for transportation arrangement after being discharged. The inclusion criteria were patients aged 18-80 years with diagnosis of psychiatric and/or substance use disorders.

Human Research Ethics Committee of Naresuan University approved this study. Written consent was received from all participants after the full study details were explained. The patients were assured that the study was anonymous and confidential. This study was carried out in August till October 2004.

2.4 Analysis plan

The validity tests included criterion and discrimination validity. The Cronbach's alpha and inter-rater reliability were tested for its reliability. Factor analysis was performed to examine its structure components.

2.5 Other measurements

To test the criterion validity of the Thai HoNOS, rater 1 used 3 other mental health measurement Psychiatric Rating Scale (BPRS), Global Assessment of Functioning (GAF), and Clinical Global Impression (CGI) scale to rate acute patients in group 1 at the same evaluation.

2.6 Raters and rating

Two raters were psychiatric nurses at Suan Prung psychiatric hospital who had over 10 year experiences with psychiatric patients. Prior to the field-testing, two raters carefully studied the HoNOS trainers guide, introduced by the Royal College of Psychiatrists, (Wing, Curtis & Beevor, 1990) and made themselves familiar with the scoring system. Rating was made according to the HoNOS trainers guide.

Group 1 and Group 2 subjects were assessed and scored under the Thai HoNOS item 1 to item 10. The patients in group 1 were also rated with other three measurements (GAF, CGI, and BPRS). The ratings were much relied upon direct observation and interview with patients and information obtained from their medical records. Before assessing, both raters independently reviewed the subjects' medical records, relevant information including diagnostic category, mental health status, date of birth, and marital and employment status, etc. After that, they directly observed and performed joint interviews but independently rated scores. Thai HoNOS were scored by using the information received from the medical records, followed by GAF, CGI, and BPRS, respectively.

The total mean scores of the Thai HoNOS (H1-10) was significantly correlated with other popular outcome measures. The mean scores of the Thai HoNOS were highly correlated with BPRS ($r = 0.915$, $p < 0.000$), followed by GAF ($r = -0.896$, $p < 0.000$), and CGI ($r = 0.880$, $p < 0.000$), respectively.

2.7 Reliability tests

The internal consistency of the Thai HoNOS was assessed by using scores from the Subjects' assessed by rater 1. The magnitude of the consistency was examined by calculating the Cronbach's alpha coefficient. The scale was internally consistent and reliable if a Cronbach's alpha was equal to 0.70 or higher (John et al, 1990 p. 1657-1661).

The correlation was assessed by using Spearman rank correlation coefficient (r). Inter-rater reliability was assessed by means of intra-class correlation coefficients (ICCs) of Thai HoNOS scores assessed by the two raters. The ICCs was examined by the use of two-way ANOVA.

2.8 Validity tests

Concurrent validity was tested by comparing among GAF, CGI, and BPRS. Criterion validity was assessed by comparing total Thai HoNOS scores with GAF, CGI, and BPRS mean scores. The difference between Thai HoNOS scores of Group 1 and those of Group 2 by rater 1 was examined by evaluating the discrimination validity. The 2-tailed significant difference of the scores was examined by using Wilcoxon Rank Sum Test. A p-value being less than 0.05 indicated the significant correlation or difference (John et al, 1990. pp. 1657-61).

The Cronbach alpha coefficient of the Thai HoNOS was 0.68. In respect of inter-rater reliability, most items had high intra-class correlation coefficients (r's > 0.70). A high correlation (r's > 0.80) with the Clinical Global Impression (CGI) and the Brief Psychiatric Rating Scale (BPRS) suggested high concurrent validity. In addition, the discriminating power of the overall clinical outcomes between acute and sub-acute psychiatric inpatients was satisfactory (p < 0.05).

2.9 Factor analysis study

Thai HoNOS scores of Group 1 and Group 2 assessed by rater 1 were included in the factor study. All 10 items were subject to a principal component analysis to identify the distinct factors. Eigen value-one test was applied to keep or discard factors. Finally, a varimax rotation was performed to elicit the factor components. An item with the loading of 0.4 or higher was considered as a significant loading factor.

2.10 Data analysis

Data included in the analysis were summarized in Table 28.

Criterion validity: The Thai HoNOS scores by rater 1 (Group 1 'set A' and Group 2 'set B') show a correlation with GAF, CGI, and BPRS

Discrimiation validity: The mean of Thai HoNOS scores by rater 1 for Group 1 'set A' was compared with the mean score for Group 2 'set B' patients

Cronbach alpha: internal consistency of Thai-HoNOS scores was analyzed for rater 1 (Group 1 'set A' and Group 2 'set B')

Interrater reliability: ICCs of Thai HoNOS scores were assessed by scores of rater 1 (Group 1 'set A' and Group 2 'set B') and score of rater 2 (Group 1 'set C' and Group 2 'set D')

Factor analysis: The Thai-HoNOS scores by rater 1 (Group 1 'set A' and Group 2 'set B')

Table 28 Measures rated by both raters

	Group 1[a] (n=23)	Group 2[b] (n=23)
Rater 1	Thai HoNOS 'set A', GAF, CGI, BPRS	Thai HoNOS 'set B', GAF, CGI, BPRS
Rater 2	Thai HoNOS 'set C'	Thai HoNOS 'set D'

Thai HoNOS = Thai Health of the Nation Outcome Scales;   GAF = Global Assessment of Functioning Scales;

CGI = Clinical Global Impression Scales;   BPRS = Brief Psychiatric Rating Scale

Group 1[a] = acute patient: inpatients hospitalized for less than 48 hours;

Group 2[b] = non-acute patient: inpatients who were discharged but waiting for transportation arrangement

3. Subjects of the study

The inpatients participating in this study were admitted to the hospital during August to October 2004, 29 males and 17 females, with a mean age of 37.7 years (s.d. 13.6) (range from 18-60 years). Subjects had the illness for the average of 7.9 years (s.d.7.2), and stayed in the hospital for the average of 37.1 days (s.d.28.8).

Subjects were diagnosed and code the International Statistical Classification of Diseases version 10 (ICD 10). Four diagnostic categories were included in the study. The predominant diagnosis was schizophrenia, paranoid, and acute psychotic disorders (F20.0-F29.9) (28; 60.9%); 14 patients (30.4%) had alcohol-related disorders (F10.0-F10.9); 3 patients (6.5%) had mood disorders (F30.0-F39.9); and only one (2.2%) had mental retardation (F70.0-89.9).

On average, the time spent for rating HoNOS was 10.7 minutes (s.d. 5.8). Actual time consumed depended on complexity of inpatient's problems.

Table 29 demonstrates the profile (mean scores) of the individual Thai HoNOS item for the two patient groups (acute and non-acute). The maximum Thai HoNOS score was 22 with mean score of 6.07. The highest scores were seen in item 6 (hallucinations/delusions), followed by item 8 (other symptoms), and item 1 (aggression), respectively. The lowest scores were seen in item 5 (physical illness/handicap), followed by item 10 (activities of daily living), and item 3 (alcohol & drug use), respectively. In group I (acute inpatient), only mean scores of items 1, 6, and 8 were higher than 1, especially item 6 was higher than 2 indicating an existence of hallucination and delusion problem. In group 2 (sub-acute inpatient), only mean scores of items 6 and 8 were higher than 1.0.

Table 29 Number and mean of patients and Thai HoNOS scale

HoNOS items and total	Number of cases	% scores	Group 1 (acute)	Goup 2 (sub-acute)	Total Group 1 & 2
H1 Aggression	15 (10%)	14	1.52 (1.50)	0.00 (0.00)	0.76 (1.30)
H2 Self-harm	5 (3%)	5	0.78 (1.54)	0.04 (0.21)	0.41 (1.15)
H3 Alcohol & drug abuse	10 (7%)	8	0.61 (1.12)	0.04 (0.21)	0.33 (0.85)
H4 Cognitive problems	11 (7%)	6	0.52 (0.95)	0.26 (0.75)	0.39 (0.86)
H5 Physical illness/ handicap	9 (6%)	4	0.26 (0.69)	0.13 (0.46)	0.20 (0.58)
H6 Hallucinations/delusions	24 (16%)	22	2.43 (1.47)	0.39 (0.90)	1.41 (1.59)
H7 Depression	12 (8%)	6	0.83 (1.19)	0.30 (0.64)	0.57 (0.98)
H8 Other symptom	32 (21%)	18	1.74 (1.10)	0.70 (0.70)	1.22 (1.05)
H9 Relationship problems	20 (13%)	10	0.83 (0.83)	0.22 (0.52)	0.52 (0.75)
H10 Activities daily living	13 (9%)	8	0.52 (0.85)	.000 (0.00)	0.26 (0.65)
H1-10 HoNOS total (range 0-22)			10.04 (4.15)	2.09 (2.17)	6.07 (5.19)

4. Reliability

Table 30 indicates that the ICCs of Thai HoNOS was excellence at total scale score 0.96. The interrater reliability by item was satisfactory to excellence, between 0.75 and 0.98. For the item 8 (H8 Other symptom) the reliability was satisfactory at 0.75, while the rest had good to very good level of ICCs at 0.88 or better.

Cronbach's alpha of the Thai HoNOS was 0.68, indicating that the scale of Thai HoNOS had unsatisfactry internal consistancy. The Cronbach's alpha, lower than 0.70, was unsatisfactory of internally consistent reliable.

Table 30 Comparing intra-class reliability coefficients of Thai HoNOS with the original and Australia studies

Thai HoNOS items and total	Wing et al. 96*		Victorian field trials*	Thai HoNOS
	Nottingham	Manchester	Geelong	
	N=100	N=100	N=50	N=54
H1 Aggression	0.97	0.80	0.73	0.93***
H2 Self-harm	0.88	0.92	0.83	0.96***
H3 Alcohol & drug abuse	0.99	0.61	0.86	0.92***
H4 Cognitive problems	0.81	0.92	0.41	0.94***
H5 Physical illness/ handicap	0.88	0.89	0.62	0.91***
H6 Hallucinations/delusions	0.87	0.92	0.83	0.98***
H7 Depression	0.84	0.89	0.79	0.92***
H8 Other symptom	0.95	0.52	0.61	0.75***
H9 Relationship problems	0.74	0.78	0.60	0.88***
H10 Activities daily living	0.71	0.90	0.68	0.88***
H1-3 Behaviour total	0.89	0.74		0.94***
H4-5 Impairment total	0.87	0.95		0.94***
H6-8 Symptom total	0.88	0.81		0.95***
H9-10 Social total	0.82	0.68		0.95***
H1-10 HoNOS total	0.86	0.77	0.71	0.96***

* Wing et al. 1996

** Trauer et al. 1999, p. 16.

***p<0.001 [a]Nottingham and [b]Manchester

Table 31 indicates changes of the Cronbach alpha if Items were deleted, Alpha would reach the highest value (0.71) if item on cognitive problems was deleted. While alpha would become worst if item on hallucinations and delusions was deleted, followed by items on aggression and alcohol and drug abuse.

Table 31 Item-total statistics for Cronbach's alpha

Scale	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
H1 Aggression	5.30	22.84	0.19	0.33	0.70
H2 Self-harm	5.65	21.83	0.35	0.50	0.66
H3 Alcohol & drug abuse	5.74	25.44	0.09	0.42	0.70
H4 Cognitive problems	5.67	25.83	0.04	0.26	0.71
H5 Physical illness	5.87	24.43	0.37	0.45	0.66
H6 Hallucinations	4.65	15.65	0.70	0.66	0.56
H7 Depression	5.50	22.92	0.32	0.60	0.66
H8 Other symptom	4.85	20.80	0.52	0.43	0.62
H9 Relationship problems	5.54	22.34	0.56	0.49	0.63
H10 Activities daily living	5.80	23.54	0.47	0.43	0.65

5. Validity

Total average score of Thai HoNOS (H 1-10) was 6.07 (see table 29). The mean score for acute patients (group 1) was higher than the mean score for sub-acute patients (group 2). The discrimination analysis found that mean scores of group 1 and group 2 were different (Wilcoxon test, p<0.01). It indicates that the acute patients' Thai HoNOS mean score (group 1) had significantly higher than sub-acute patients' (group 2).

The total mean scores of Thai HoNOS (H1-10) was significantly correlated with other popular outcome measures. The scores of the Thai HoNOS were highly correlated with BPRS (r = 0.915, p<0.01), followed by GAF (r = -0.896, p<0.01), and CGI (r = 0.880, p<0.01), respectively.

6. Factor analysis

Table 32 shows the principle component analysis of the four-factor model of Thai HoNOS. The first factor (21.60% of the variance) comprised item 7 depression, item 2 self-harm, and item 5 physical illness/handicap. The second factor (21.09% of the vairance) consisted of item 10 activities of daily living, item 4 cognitive problems, item 9 relationship problems, and item 8 other symptoms. The third factor (16.08% of the variance) comprised item 3 alcohol/drug use and item 6 hallucination/ delusions. The fourth factor (13.96% of the variance) consisted of only item 1 aggression. The Cronbach's alphas of the first, second, and third factors were 0.70, 0.64, and 0.49, respectively.

Table 32 Factor analysis of Thai HoNOS (weight > 0.40 shown)

	Factor 1	Factor 2	Factor 3	Factor 4
H1 Aggression	0.89			
H2 Self-harm	0.83			
H3 Alcohol & drug abuse	0.59			
H4 Cognitive problems		0.77		
H5 Physical illness/ handicap		0.76		
H6 Hallucinations/delusions		0.57		
H7 Depression		0.55		
H8 Other symptom			0.89	
H9 Relationship problems			0.67	
H10 Activities daily living				0.93
Eigen value	2.16	2.11	1.61	1.40
Percentage of variance[a]	21.60	21.09	16.08	13.96

[a] Total = 72.724.

7. Discussion and conclusions

The results indicated that Thai HoNOS generally fulfills the requirements of a clinically acceptable outcome scale for routine use in mental health services. The psychometric properties of Thai HoNOS were similar to those of the original one.

The time to complete Thai HoNOS (mean 10.72 minutes, s.d. 5.76) was comparable to McClelland's study, which ranged between 5-15 minutes, depending on rater's experience and the complexity of the patient's problem (McClelland et al, 2000, pp. 98-105).

In comparison to general measures, the Cronbach's Alpha coefficient of the Thai HoNOS (r = 0.68) is acceptable a little better than the study of Orrell et al. (Cronbach's alpha = 0.65) (Orrell et al., 1998, pp. 409-12). Though, it is satisfactory for a measure to have small number of items (Feinstein, 1987). But no items were deleted.

The ICCs of all items were good to very good (r > 0.88), except item the other symptoms (r = 0.75), which was still a satisfactory level. This suggests that Thai HoNOS is a measure that possesses satisfactory to excellent inter-rater reliability (Brooks, 2000, pp. 504-11); (Trauer et al, 1999, pp. 380-8); (Wing et al, 1998, pp. 11-8); (Green & Gracely 1997). The item and total ICCs, obtained from this study, alongside with those reported from the British reliability studies (Trauer et al, 1999, pp. 380-8); (Wing et al, 1998, pp. 11-8). The second and the ICC coefficients are good to very good for all items but two items (moderately good and acceptable). This study shows that reliability levels (0.96) were generally higher than those found in the original work (Trauer et al, 1999, pp. 380-8); (Wing et al., 1996).

The validity in discriminating those having good and bad overall clinical outcomes was in line with the Victorian field trial's study (Orrell et al., 1998, pp. 409-12). The difference in the mean total Thai HoNOS score between acute patients and sub-acute patients was significant.

Similar to several previous studies, this study examined the validity of HoNOS and reported a good correlation with other standard measures. Examination of the validity showed that the Thai HoNOS had very high correlation with GAF, CGI, and BPRS.

Thai HoNOS/ GAF and Thai HoNOS/ BPRS correlations were higher than the original HoNOS. McClelland et al (McClelland et al., 2000), compared HoNOS with two well-established measures, GAF and BPRS in specialist mental health services, the correlations were high for BPRS ($r = 0.72$, $p < .000$, $n = 93$ and for GAF $r = 0.71$, $p < .000$, $n = 149$).

While this study found a 4-factor model of Thai HoNOS, the item of hallucination/ delusions considerably loaded on two factors. The unclear loading of this item may be caused by small samples.

However, there were a number of methodological limitations in this study. First, all subjects were limited to psychiatric inpatients. Further study should include other groups of psychiatric patients e.g. outpatients, patients in community. Second, the raters of this study were nurses only. It is interesting to test whether nurse is a good representative for other mental health professionals, e.g., psychiatrists, psychologists. Finally, the sample size for factor analysis study was relatively small comparing with number suggested by experts (more than 5 subjects per variable or more than 100 subjects per study) (Norman & Streiner, 1994).

For an inpatient setting, the Thai HoNOS was brief and easy to use. Reliability and validity tests generally provided good to excellence results, in line with other previous studies that later used as routine care of psychiatric inpatients (Brooks, 2000, pp. 504-11). Trauer et al. (Trauer et al, 1999, pp. 380-8) suggested that HoNOS scores showed strong association with utilization, and may therefore have a role to play in casemix systems. As a measure of the overall clinical outcome of a psychiatric patient, this measure can be used in psychiatric casemix study in Thailand.

After reliability and validity established, the next logical steps in the development of Thai HoNOS scale are how compatible of the Thai HoNOS with other psychiatric population (e.g., outpatients); what result of the HoNOS if used by other mental health professionals. In particular, follow-up studies need to be done to show how good Thai HoNOS be used as a measure to predict outcome rather than limited interpretation from a cross-sectional study.