Validity and Reliability of the Thai Mental Health Questionnaire in Psychiatric Patients in the Community†

Sucheera Phattharayuttawat, Ph.D.*
Thienchai Ngamthipwattana, M.D.*

Abstract : **The Thai Mental Health Questionnaire (TMHQ) has been developed for assessing the effectiveness of mental health services. The purpose of this study was to investigate the validity and reliability of the TMHQ in patients in contact with mental services. Subjects (age 12-74 years) came from out-patient and acute in-patient units in general practice. We obtained the opinions of experienced professionals, advocacy group and patient group to evaluate consensual and content validity. The results of this study showed that 1,205 patients were assessed using the TMHQ scale. Test-retest, Odd-even, and Cronbach's Alpha reliability were good for some items and poor for others. The TMHQ had good criterion validity; acute in-patients had higher scores than out-patients. TMHQ also had good concurrent validity, correlating well with other scales. Comments suggested that the TMHQ was a useful and suitable scale for this population although social factors were not sufficiently covered.**

เรื่องย่อ : **ค่าความตรงและความเที่ยงของแบบวัดสุขภาพจิตไทยสำหรับผู้ป่วยจิตเวชในชุมชน**

สุชีรา ภัทรายุตวรรตน์ ปร.ด.*, เธียรชัย งามทิพย์วัฒนา พ.บ.*
**ภาควิชาจิตเวชศาสตร์, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร 10700.*

สารศิริราช 2545; 54: 540-546.

แบบวัดสุขภาพจิตคนไทยได้ถูกพัฒนาขึ้นเพื่อประเมินคุณภาพด้านสุขภาพจิตของคนในชุมชน คณะผู้วิจัยได้ทำการศึกษาถึงประสิทธิภาพของเครื่องมือนี้ในแง่ของความตรงและความเที่ยง โดยทดสอบกับผู้ที่มารับบริการที่คลินิกสุขภาพจิต กลุ่มตัวอย่างมีจำนวนทั้งสิ้น 1,205 ราย (ช่วงอายุระหว่าง 12 ถึง 74 ปี) มีทั้งผู้ป่วยนอกและผู้ป่วยใน ผลการศึกษาพบว่าค่าความเที่ยงจากการทดสอบด้วยวิธีการทดสอบซ้ำ การแบ่งข้อคู่ข้อคี่ และการหาความสอดคล้องภายในแบบครอนบัคแอลฟา ได้ค่าที่สูงในหลายสเกล แต่บางสเกลก็มีค่าที่ต่ำ สำหรับค่าความตรง พบว่าแบบทดสอบชุดนี้มีความตรงตามเกณฑ์ในระดับสูง โดยกลุ่มผู้ป่วยในมีระดับคะแนนสูงกว่ากลุ่มผู้ป่วยนอก รวมทั้งมีความตรงเชิงเกณฑ์สัมพันธ์ที่สูงด้วย กล่าวคือพบว่ามีความตรงเชิงเกณฑ์สัมพันธ์กับแบบวัดอื่น ข้อคิดเห็นจากผู้ใช้เครื่องมือนี้สรุปได้ว่าแบบทดสอบชุดนี้มีความเหมาะสมที่จะนำไปใช้กับการตรวจคัดกรองในประชากรกลุ่มนี้ ถึงแม้ส่วนที่วัดด้านสังคมยังอาจครอบคลุมไม่เพียงพอ

†This research was supported by Chalermprakiet Fund, Faculty of Medicine Siriraj Hospital, Mahidol University. **Department of Psychiatry, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.*

INTRODUCTION

The Thai Mental Health Questionnaire (TMHQ) is a 70-item self-report screening scale that was originally developed to indicate the possible presence of psychopathology in the setting of a community or of a medical out- and in-patient clinic.

The TMHQ was designed to assess a domain of functioning. It is factor-analyzed to identify separable dimensions, representing theoretical constructs, within the domain[1]. Psychometricians strongly recommend that test developers should begin by factor analyzing the items[2-5]. The specific information was obtained through the use of factor analytically divided subscale scores. Then these factors were obtained logically and empirically by using exploratory approaches :principal components, extracting factors, varimax rotating factors procedures[5-11]. The factor-analytically derived dimensions then serve as subscales[12]. The factor analysis was used to deduce the 94 items on the TMHQ by eliminating items that fail to load on any factor. After this procedure, the items were reduced to 70 items which loaded into five factors. The five factors are somatic, depression, anxiety, psychotic, and social functions. The factors were estimated to explain the covariances among the items[6]. The five factors account for 55.8% of the total variance with confirm to Streiner's view point[11] that factors should explain at least 50% of the total variance. Therefore, the TMHQ is a multifactor scale to assess psychopathology[1].

Jansen and Haynes[13] suggested that many questionnaires actually measured several factors and the use of a single index with a multifactor scale was inappropriate. In the same way, Fisher and Corcoran[14] suggested that using items grouped together empirically on the basis of factor analysis would ensure all subdimensions of a multidimensional measure had good reliability and validity, and the multidimensional instrument could be as useful as several unidimensional measures, and perhaps more efficient.

The reliability coefficients for the Alpha range from 0.80 to 0.92, and the reliability coefficients for the Split-half range from 0.80 to 0.90[1]. This shows that the reliability coefficients are in the middle to high values[15]. The acceptable internal consistency of a psychological instrument should be 0.7[16]. Alpha and split-half reliability coefficients are an index of the homogeneity of the measuring instrument and also exhibit acceptable internal consistency[16]. In addition, when an instrument has an adequate reliability it means that the items are tapping a similar domain, and hence, that the instrument is internally consistent[13].

Thailand is faced with the problems of rapid social and cultural changes that are responsible for increasing mental health problems. Meanwhile, the proportion of mental health professionals to the total population is inadequate. So, a future mental-health policy must be well thought out. Because mental-health treatment in the hospitals is neither sufficient nor effective, health promotion and education in the community must be considered in order to find ways to prevent mental illness before professional treatment is required. Moreover, psychiatric instruments which can be easily used in the community or clinical setting must also be considered. The use of rating scales to detect mental disease is useful. The standard in psychiatric practice is usually a class of persons, such as psychotic, neurotic, normal or other reference groups. Many clinicians have found that a schedule of items covering a variety of observable symptoms and relevant questions concerning patients' attitudes, feelings, and behaviors helps to assure a more thorough and complete psychiatric interview or examination. There is less likelihood that possibly significant phenomena may be omitted or overlooked. This is useful, whether or not the information so derived is later reduced to numerical form for another purpose. Also, ratings are more likely to be interpreted within the same semantic frame of reference. Clinicians with different backgrounds and holding different theoretical views find it valuable to have on record information gathered in a systematic and common format[17,18]. In addition, rating scales are very easy to use and their flexibility and face validity recommend them highly[14]. The Thai Mental Health Questionnaire (TMHQ) is now available for use in general psychiatry and a very widely used brief measure for screening for psychiatric cases in the community and in clinical practice. This study aimed to evaluate the validity and reliability of the TMHQ for psychiatric patients in the Thai community.

MATERIALS AND METHODS

The samples were drawn from two main sources of psychiatric contact: acute in-patients and out-patients. The information collected included a battery of scales (described below), age, gender, marital status, employment status, diagnosis according to DSM-IV or ICD-10 and educational level. The patients were assured that the study was anonymous and confidential. Data analysis was carried out using SPSS for Window.

The Thai Mental Health Questionnaire (TMHQ)[1]

The Thai Mental Health Questionnaire (TMHQ) was developed from designs to record the presence of any symptom or group of symptoms of psychopathology. It was based upon a system analytic model of the DSM-IV, and in addition to identify problems across 5 broad mental health-related areas. The Thai Mental Health Questionnaire (TMHQ) is a 70-item self-report screening scale which is loaded into five factors. The five factors are somatic, depression, anxiety, psychotic, and social functions. Each item is directed to indicate the extent to which the sample has been bothered or distressed by the problems or complaints represented by each of the 70 items, over a specified time interval, usually a one month period. The degree of distress ranges from "Not at all" to "Extremely". The reliability coefficients for the Alpha are in a range from 0.80 to 0.92, and the reliability coefficients for the Split-half range from 0.80 to 0.90[2].

The TMHQ bases construction and manual will be determined as follows:

1. The symptoms which the TMHQ measures based on DSM-IV are:

Somatization: This dimension reflects distress arising from perceptions of bodily dysfunction. Complaints focus on cardiovascular, gastrointestinal, respiratory, and other symptoms with strong autonomic mediation. Headaches, pains, and discomfort localized in the gross musculature are also components, as are other somatic equivalents of anxiety.

Depression: Symptoms of dysphoric affect and mood are represented, as are signs of withdrawal of interest in life events, lack of motivation and loss of vital energy. The dimension mirrors feelings or hopelessness, worthlessness, meaninglessness, pessimism, loneliness, downheartedness, or discouragement. Several items are included concerning thoughts of death and suicidal ideation.

Anxiety: General indications such as restlessness, nervousness, and tension are included as are additional somatic signs. Items measuring free-floating anxiety and panic attacks are an integral aspect of this dimension.

Psychotic: Florid, acute symptomatology, as well as behaviors typically viewed as more oblique, less definitive indicators of psychotic process are represented. In addition, secondary signs of psychotic behavior and indications of a schizoid life style are also represented.

Social function: General indications in interpersonal relationships and contact with other people in the society.

2. The TMHQ Items. The principal components analysis was carried out on 70 items. Therefore, the final TMHQ is composed of 70 items[2].

All TMHQ items were found to be significantly different at $p \leq 0.001$. That means the TMHQ has sufficient power to discriminate between those with mental disorders and normal people.

3. Administration. The standard instructions for the person at the top of the scale, and the items, are in very basic Thai, so literacy level is rarely a problem.

4. Scoring and interpretation. Special answer sheets have been developed that can be read and scored. Each of the 5 scoring categories consists of from 10 to 20 items. By transferring the item responses to a separate scoring sheet the responses that make up a scoring category are added, and then divided by the number of items in that category. As a result, the score on any category can have a value ranging from 0, indicating no problem at all, to a maximum of 4, when there is an extreme problem (this is for the negative items, but for the positive items, the score is reversed). And, since all categories are interpreted the same way, screening of even a large number of examinees' records can be quickly accomplished just by scanning the raw scores.

5. Scoring summary. A summary score is a quantitative index of the degree to which a particular problem area is relevant to a client. The TMHQ's scoring summary is as follows:

Somatization scale: Total score devided by 10.

Depression scale: Total score devided by 20.

Anxiety scale: Total score devided by 15.

Psychotic scale: Total score devided by 10.

Social function scale: Total score devided by 15.

6. The TMHQ profiles. To characterize each item in the symptom scale by means of quantitative index, the five scales are:

"0" NOT AT ALL: No stress reported.

"1" A LITTLE BIT: Some stress but infrequent and of low intensity.

"2" MODERATELY: Somewhat regular stress of mild or moderate intensity.

"3" QUITE A BIT: Regular stress of moderate to high intensity.

"4" EXTREMELY: Examinees experience extreme stress associated with these symptoms due to frequency, intensity or both.

For the positive questions, the score must be reversed.

The TMHQ's profile was created in an easily construct profile. The user can visualize it because its numerical values is translated into a plotted profile.

The normal range is between the T-score 40 to 65.

Brief Psychiatric Rating Scale (BPRS)[19]

This is an interviewer-rated psychiatric scale covering a range of mental state phenomena. The scale is often used to measure the overall psychotic state, and the total score indicates its severity. Factor analysis has, however, identified the following five specific factors: thought disorders, emotional withdrawal, anxiety-depression, aggressiveness, and agitation.

The original version included 16 items, but the scale was soon enlarged to include 18 items. The scale is one of the most widely used scales internationally and should be included in studies dealing with schizophrenia and other psychotic conditions.

General Health Questionnaire (GHQ)[20]

This is a very widely used brief measure for screening for psychiatric cases in the community. In UK general practice, the best threshold for case identification for the GHQ-12 is 3/4, with subjects scoring 4 or more being identified as psychiatric cases.

The GHQ was designed for use in general population surveys, in primary medical care settings, or among general medical outpatients[4]. It was meant to be a first-stage screening instrument for psychiatric illness that could then be verified and diagnosed. The questions ask whether the respondent has recently experienced a particular symptom (like abnormal feelings or thoughts) or type of behavior. Emphasis is on changes in condition, not on the absolute level of the problem, so items compare the present state to the person's normal situation with responses ranging from "less than usual" to "much more than usual". The questionnaire begins with relatively neutral questions and leads to the more overtly psychiatric items toward the end[5].

RESULTS

1,205 patients were assessed using the overall package of rating scale. Some patients were not able to complete every scale. The package took approximately 25 minutes per person to complete. Of the total, 399 came from the acute admission wards and 806 from the out-patients. The gender ratio was 503 males: 702 females, with a mean age was 32.54 years (S.D. 9.88, range 12-74 years). 514 patients were from Ubonrajchatani, 312 patients were from Nakornsawan, and 379 patients were from Songkla. 477 were married or co-habiting, 621 were single, separated or divorced; and 107 people lived alone. Subjects had a mean of 6.4 years of education (S.D. 11.7, range 0-20 years); 802 (or 2/3) patients were currently unemployed. The predominant diagnosis was depression (151); 94 patients had schizophrenia, 122 had anxiety disorders, 81 had personality disorders, 91 were alcohol-dependent, and there were a variety of other diagnoses.

Table 1 shows that the maximum TMHQ T-score was 82 and the mean was 70.40. In each of five scales, fewer than 10 subjects had a T-score at the "problem" level and the means were also low, suggesting that there was not an excess of score of 1. Conversely, four of the remaining scored at problem level in more than T-score 65 (somatization, anxiety, depression and psychotic) yielded a score at the problem level for over a half of the subjects.

Table 1. Number of patients with clinical problems in each item on the TMHQ scale (T-score above 65).

TMHQ scale	N	%
1. Somatization	304	25.23
2. Anxiety	398	33.03
3. Depression	227	18.84
4. Psychotic	202	16.76
5. Social function	447	37.09

Remark : Each person had more than one symptom.

Reliability

Reliability was determined using Test-retest, Odd-even and Cronbach's Alpha. Table 2 indicates that despite the low number of scores at the "problem" level for 3/5 scales, the Odd-even and Cronbach's Alpha reliability were reasonably good, at between 0.81 to 0.93. This indicated that the scale had good internal consistency but was not suitable for Test-retest technique.

Concurrent validity

Table 3 shows that the total on the TMHQ correlated well with that on the other standard scales which have potential values as outcome measures.

Table 2. Test-retest, Odd-even, and Cronbach's Alpha reliability coefficients.

TMHQ scale	Test-retest reliability	Odd-even reliability	Cronbach's Alpha reliability
1. Somatization	0.71	0.88	0.91
2. Anxiety	0.34	0.81	0.82
3. Depression	0.55	0.92	0.89
4. Psychotic	0.39	0.90	0.93
5. Social function	0.45	0.93	0.88

Table 3. Concurrent validity of the TMHQ: correlation with other scales.

Scale	TMHQ	BPRS	GHQ
TMHQ	-	0.87*	0.84*
BPRS		-	0.34*
GHQ			-

*p<.001

Criterion validity

Table 4 shows the T-score of TMHQ for each of the two psychiatric groups (out-patients and in-patients) and this shows that out-patients had significantly lower TMHQ T-score than in-patients, in line with expectations. In general, in-patients also had higher TMHQ T-score than out-patients. Criterion validity was also good for the BPRS ($t = 10.4$, $P = 0.000$) and GHQ ($t = 6.62$, $P = 0.000$), each of which predicted use of in-patient psychiatric services.

Table 4. Criterion validity: independent t-test using TMHQ by current psychiatric placement.

Placement	N	Mean	S.D.	P-value
In-patients	399	79.03	3.44	0.000
Out-patients	806	66.33	6.09	

Consensual validity

Many staff commented positively about the TMHQ, saying, for example, that they were "generally well designed", "thorough", and "suitable" for the intended purpose. There was some concern about the "scales" ability to take into account other important factors such as post-traumatic stress disorder or the various types of abuse could be recorded. It was suggested that the scales should also be able to consider the patients' view of their level of symptoms.

Content validity

Comments from the advocacy group and patient group were reasonably positive, suggesting that the TMHQ "seems generally appropriate". However, it was criticised for not giving an accurate picture of performance, psychotic symptoms or the experiences of some patients with schizophrenia, nor was the issue of poverty addressed. Lastly, there was concern about the definitions of abnormality used in TMHQ; for example the meaning of "bizarre".

DISCUSSION

A number of brief scales have been widely used in mental health settings but, unlike TMHQ, most do not have an overall comprehensive scale or a series of items which flag problems across a wide range of areas. The Brief Psychiatric Rating Scale (BPRS) covers only psychiatric symptoms but not social outcome, while the General Health Questionnaire (GHQ) is a useful brief screening scale for identifying psychiatric cases in the community but is of limited value in hospital populations because it fails to cover psychotic symptoms or social functioning.

The TMHQ was relatively easy to use and fairly quick to complete (15 minutes or less) when the full clinical data were available. This suggests that it might be incorporated into ordinary clinical practice with much less inconvenience than some of the longer instruments. The sample of subjects used in this study was primarily drawn from community sources in Northern (Nakornsawan), North Eastern (Ubonrajchathani), and Southern (Songkla) Thailand. They were less socially disadvantageous than an inner-city population. Nevertheless, nearly two-thirds were unemployed, reflecting both the current job market and the disabling effects of mental illness in general, rather than schizophrenia alone. 50% of people had problems in the rather nebulous category of "Other mental health problems", suggesting that this category was too broad.

The Odd-even and Cronbach's Alpha reliabilities were reasonably good, at between 0.81 to 0.93. This indicated that the scale had good internal consistency[14]. However, the test-retest reliability was lower than internal reliability, even though the researcher carrying out the ratings had several years' experience in psychiatry, but this might be due to rapid changes in the patients' state[21].

The concurrent validity of the TMHQ was good and had linear correlation with the BPRS and GHQ, both of which are probably quick and easy to complete.

For a community population with mental health problems, the TMHQ was a brief easy-to-use scale. Reliability and validity were generally good, although on the basis of our results, TMHQ performed no better than other outcome measures. The main advantage of the TMHQ appears to be that the 70 items included provide a more comprehensive picture of mental health outcomes than most of the other brief measures.

Clinical implications:

1. Brief and easy to use
2. Wide range of problems addressed
3. Correlates well with other scales and high scores associated with in-patient status

Limitations:

1. Low test-retest reliability for some items.
2. Limited coverage of psychotic symptoms.

ACKNOWLEDGEMENT

The researchers would like to thank Mr. Maurice Broughton, Department of Foreign Language, Faculty of Sciences, Mahidol University for editing this paper.

REFERENCES

1. Phattharayuttawat S, Ngamthipwattana T, Sukhatungka S. The development of psychometric test "The Thai Mental Health Questionnaire". Siriraj Hosp Gaz 1999; **51:** 244-49.
2. Stewart A, John E. Measuring functioning and well-being: The medical outcomes study approach. London: Duke University Press, 1992: 11-16.
3. Briggs SR, Cheek JM. The role of factor analysis in the development and evaluation of personality scales. J Pers 1986; **54:** 106-48.
4. Comrey AL. Factor analysis methods of scale development in personality and clinical psychology. J Consult Clin Psychol 1988; **56:** 754-61.
5. Cortina JM. What is coefficient alpha? An introduction of theory and applications. J Appl Psychol 1993; **78:** 98-104.
6. Floyd FJ, Widaman KF. Factor analysis in the development and refinement of clinical assessment instruments. Psychological Assessment 1995; **7:** 286-99.
7. Gorsuch RL. Factor analysis (2nd ed). Hilsdale, New Jersey: Erbaum, 1983: 22-28.
8. Cole DA. Utility of confirmatory factor analysis in test validation research. J Consult Clin Psychol 1987; **55:** 584-94.
9. Hayduk LA. Structural equation modeling with LISREL: Essentials and advances. Baltimore: Johns Hopkins University Press, 1987: 44-49.
10. Bollen KA. Structural equations with latent variances. New York: Wiley, 1989: 87-90.
11. Kenny DA, Kashy DA. Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. Psychol Bull 1992; **112:** 165-72.
12. Streiner DL. Figuring out factors: The uses and misuses of factor analysis. Can J Psychiatry 1994; **39:** 135-40.
13. Jansen B, Haynes S. Self-report questionnaires and inventories. Handbook of Behavioral Assessment (3rd ed). New York: John Wiley & Sons, 1989: 34-39.
14. Fischer F, Corcoran K. Measures for clinical practice: A Source Book, Volumn 1 (2nd ed). New York: The Free Press, 1994: 24-28.
15. Aiken LR. Psychological testing and assessment (2nd ed). Boston: Allyn and Bacon, 1994: 34-54.
16. Bech P, Malt UF, Dencker SJ. Scales for assessment of diagnosis and severity of mental disorders. Acta Psychiatr Scand 1993; **16:** 372.
17. Staff of Research and Education Association. Handbook of Psychiatric Rating Scales. New York: Research and Education Press, 1986; **25:** 6-33.
18. MacKinnon RA, Yudofsky SC. Principles of the psychiatric evaluation. New York: Lippincott Company, 1991: 21-26.
19. Overall JE, Gorham DR. The brief psychiatric rating scale. Psychological Reports 1962; **10:** 799-812.
20. Goldberg DP, Williams PA. User's guide to the general health questionnaire. London: Nfer-Nelson, 1988: 50-65.
21. McDowell I, Newell C. Measuring health: A guide to rating scales and questionnaires (2nd ed). New York: Oxford University Press, 1996: 28-33.