Routledge
Taylor & Francis Group

Validation and factor structure of the Thai version of the EURO-D scale for depression among older psychiatric patients

Tawanchai Jirapramukpitak[a]*, Niphon Darawuttimaprakorn[b], Sureeporn Punpuing[b] and Melanie Abas[c]

[a]Department of Psychiatry, Faculty of Medicine, Thammasat University, Klong Luang, Pathumthani, Thailand; [b]Institute for Population and Social Research, Mahidol Univeristy, Salaya, Nakhonpathom, Thailand; [c]Institute of Psychiatry, King's College London, Denmark Hill, London, UK

Objectives: To assess the concurrent and the construct validity of the Euro-D in older Thai persons.
Method: Eight local psychiatrists used the major depressive episode section of the Mini International Neuropsychiatric Interview to interview 150 consecutive psychiatric clinic attendees. A trained interviewer administered the Euro-D. We used receiver operating characteristic (ROC) analysis to assess the overall discriminability of the Euro-D scale and principal components factor analysis to assess its construct validity.
Results: The area under the ROC curve for the Euro-D with respect to major depressive episode was 0.78 [95% confidence interval (CI) 0.70–0.90] indicating moderately good discriminability. At a cut-point of 5/6 the sensitivity for major depressive episodes is 84.3%, specificity 58.6%, and kappa 0.37 (95% CI 0.22–0.52) indicating fair concordance. However, at the 3/4 cut-point recommended from European studies there is high sensitivity (94%) but poor specificity (34%). The principal components analysis suggested four factors. The first two factors conformed to affective suffering (depression, suicidality and tearfulness) and motivation (interest, concentration and enjoyment). Sleep and appetite constituted a separate factor, whereas pessimism loaded on its own factor.
Conclusion: Among Thai psychiatric clinic attendees Euro-D is moderately valid for major depression. A much higher cut-point may be required than that which is usually advocated. The Thai version also shares two common factors as reported from most of previous studies.

Keywords: aged; depressive disorders; Thailand; validation studies

Introduction

Depression in older persons is reported to be associated with substantially reduced quality of life and increased mortality (Penninx et al., 1999; Penninx, Leveille, Ferrucci, van Eijk, & Guralnik, 1999) and increased use of all health and social service resources (Watts et al., 2002). In spite of its public health significance, levels of recognition and treatment of depression by physician are reported to be rather low (Crawford, Prince, Menezes, & Mann, 1998; Dearman, Waheed, Nathoo, & Baldwin, 2006; Koenig, 2007). There has been little research from Thailand on depression in older people. Given the importance of detection of depression in older people, there is a need to validate a Thai version of a depression screening instrument.

Euro-D was an instrument developed by 14 European countries for screening depression in the elderly (Prince et al., 1999). It has also been used in low income and middle income countries in other regions (Castro-Costa et al., 2007; Prince et al., 2004). Its concurrent and criterion validity across 14 different settings in Europe was satisfactory (Prince et al., 1999). Most of the previous Euro-D studies suggest two common factors; affective suffering (including depression, tearfulness and wishing to death) and motivation

(including loss of interest, poor concentration and lack of enjoyment). Evidence for its cross-cultural validity in developing countries including China, India, Latin America and Africa suggest a similar factor structure across these settings (Prince et al., 2004). There is as yet no evidence for the validity of Euro-D in Thailand.

In the current study we attempted to test Euro-D's internal reliability, its concurrent validity against the diagnosis of major depressive episodes provided by the mini international neuropsychiatric interview (MINI) (Sheehan et al., 1998) and its construct using principal component analysis (PCA) on a sample of psychiatric clinic elderly attendees.

Method

Participants

We sampled consecutive patients attending a busy outpatient practice in a psychiatric hospital in the suburb of Eastern Bangkok. Patients aged 60 or above were approached in the waiting room when they attended the clinic to see their doctor. There were eight psychiatrists participating in the study. A research worker, who was a postgraduate student in population and social research, approached the patients and asked for informed consent. Study was

---

*Corresponding author. Email: tawanchaij@gmail.com

Procedures

After reading a study information sheet, those who agreed to take part were interviewed either before or after seeing their doctor in a private room on the premises. Those with evidence of severe cognitive impairment, dangerous behaviours, very frail, or with limitations of comprehension were excluded. We continued recruitment until our target of 150 completed interviews was met. The majority of the patients were presenting with common conditions such as anxiety, depression, dementia, and substance-related disorders.

Interviews with Euro-D were carried out before or after the participants saw their doctors. Each of the eight psychiatrists administered MINI-Thai version, the major depressive episode section (Kittiratpaiboon & Wongkampin, 2004), while the trained research worker administered Euro-D. We blinded the MINI interviewers to the Euro-D responses obtained by the research worker and vice versa.

Measures

The Euro-D is a structured common depressive symptoms scale derived from the geriatric mental state-AGECAT (GMS-AGECAT) interview (Copeland, Dewey, & Griffiths-Jones, 1986), SHORT-CARE (Gurland, Golden, Teresi, & Challop, 1984) and other measures including the Centre for Epidemiological Studies Depression Scale (CES-D) (Radloff, 1977), Zung Self-Rating Depression Scale (ZSDS) (Zung, 1965), and the Comprehensive Psychological Rating Scale (CPRS) (Asberg & Schalling, 1979). It was designed to be administered by trained lay interviewers and only consists of 12 items dealing with: depression, pessimism, wishing death, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment and tearfulness. A cut-point of 3/4 was identified using receiver operating characteristic (ROC) analysis in studies carried out in 14 European countries and produced a range of sensitivity (63%–83%) and specificity (49%–95%). Principal components analysis generated two factors (affective suffering and motivation) that were common to nearly every participating European country (Prince et al., 1999) and for Indian, Latin-American and Caribbean centres (Prince et al., 2004). Internal consistency was universally satisfactory, ranging from 0.83–0.93 (Prince et al., 2004).

In Thailand, the Euro-D items appeared to cover symptoms recognised locally as common in psychological disorders in older adults. The original version was carefully translated and back-translated into English. First, a team of bilingual mental health professionals and social scientists developed the first translation, paying particular attention to conceptual and semantic equivalence. Two English speaking old age psychiatrist with training and extensive experience with the GMS assisted to answer any questions about the original version. The first translation was piloted on an urban community sample of 12 elderlies to test its clarity and comprehensibility. Some questions (e.g., sleep, irritability, appetite, weight loss) simply required translation. Others were adapted to ensure local equivalence or to aid comprehension (i.e. depression, concentration and pessimism). 'Can you concentrate on entertainment or reading' was replaced with 'Can you concentrate on daily activities that you like, such as listening to the monks' teachings, the radio or watching television'. 'Have you been feeling depressed?' was replaced with 'have you been feeling sad, gloomy or in despair'. It was most difficult to translate the question on pessimism. The original question asks 'tell me your hopes for the future' and rates pessimism if the person cannot describe at least one hope. In the Thai context we anticipated that older people, most of whom were Buddhist, would not mention any hopes because they were expected at their age to view life with contentment and to take each day as it comes. Asking about 'future hopes' might therefore elicit feelings of unfamiliarity and discomfort. This could lead to misinterpreting true pessimism. We changed the wording to 'do you have any hope that in the near future good things might happen to you?'. We hoped that the modified closed question would enable to valid response.

The MINI was used as gold standard criterion measure. MINI is a short structured diagnostic interview, developed for DSM-IV and ICD-10 psychiatric disorders. The section on major depressive episode was used. The Thai version of MINI was validated against diagnoses made by local psychiatrists in a clinical setting (Kittiratpaiboon & Wongkampin, 2004) and used as a gold standard criterion against a newly developed screening test in a previous study (Arunpongpisal et al., 2006). It was also used as a main diagnostic instrument in a recent national mental health survey (Department of Mental Health, 2003).

Data analysis

The output data from the MINI and Euro-D were read into STATA and the following parameters were assessed.

(1) The prevalence of major depressive disorder according to each method of assessment.
(2) The absolute numbers of participants in each cell of the cross tabulation by test and criterion status, that is true positive, false negative, false positive and true negative.
(3) The kappa with 95% confidence intervals (CI) as the primary measure of agreement.
(4) Coefficient of agreement (the proportion of participants classified as positive by either

assessment that are classified as positive by both.

(5) To assess concurrent validity, the scores of the Euro-D were compared with the diagnosis of major depressive disorder provided by the MINI. A positive case was defined as an individual who was diagnosed as having a major depressive episode. The area under the rROC, with the correspondent 95% CI, was estimated for major depressive disorder criterion. The optimal cut-off point was identified and sensitivity, specificity, positive, and negative predictive values were calculated.

(6) The inter-item correlations and Cronbach's alpha coefficients were calculated as measures of the internal consistency of the Euro-D.

(7) A PCA of the Euro-D was performed, in an exploratory approach.

Results

Demographic and clinical characteristics

Over a 2-month period we approached 167 patients of whom 150 completed the interview (response rate 89.8%). The demographic characteristics of the sample are shown in Table 1. The principal diagnoses were mood disorders (29.3%), anxiety disorders (23.3%), schizophrenia and delusional disorders (16.7%), organic mental disorders (16%), headache (4%), substance abuse and dependence (2.7%) and other disorders (7.3%).

Scale distribution

Table 2 shows the prevalence of each Euro-D symptom across the whole sample and the prevalence of each Euro-D symptom comparing patients who were diagnosed and not diagnosed with depression.

As shown in Table 2, the mean number of items rated as present during the past month per subject was 4.9 (95% CI 4.4–5.3). The frequencies of items ranged from 16.7% (concentration) to 58.7% (irritability). Only one item (pessimism) did not significantly discriminate patients depressed from those who

Table 1. Sociodemographic characteristics (*n* = 150).

Variables		Mean 69.4 (SD = 7.0, range 59–88)	*N*	%
Age				
Gender	Male		60	40
	Female		90	60
Marital status	Single		11	7.33
	Married		85	56.7
	Widowed		40	26.7
	Divorced/separated		14	9.3
Education	<4 years		31	20.7
	4 years (primary school)		81	54.0
	>4 years		38	25.3

were not. Over 75% of depressed patients reported that they had been depressed, had sleep trouble, had been irritable, and felt fatigue.

Concurrent validity

The concordance between Euro-D and MINI for major depressive episode was fair. The level of agreement for the sample (kappa) was 0.37 (95% CI 0.22–0.52). Euro-D tended to overdiagnosis with respect to MINI with a prevalence of 39.3% compared with 34% for major depressive disorder according to MINI. The area under ROC curve for the overall discriminability of the Euro-D scale against the criterion of MINI major depressive disorder was 0.78 (95% CI 0.70–0.85) (Figure 1). Inspection of psychometric indices at different cut-points suggests a cut-point of 5/6 with sensitivity of 84.3% and specificity of 58.6%. The positive and negative predictive values were 55.9 and 80.2, respectively (Table 3). A higher cut-point of 6/7 will yield lower sensitivity (64.7%) and higher specificity (73.7%) (Table 3).

Reliability

The Cronbach's alpha for the total scale was sufficiently high at 0.72, although the internal consistency could be improved upon by the omission of EURO4 (pessimism), which marginally increased the standardised alpha to 0.74.

Construct validity

For factor analysis, the adequacy of the data for factor analysis was assessed by inspecting the correlation matrix, which showed that 10 coefficients in 66 were >0.3. Moreover, the Kaiser–Mayer–Oklin value was 0.77 and the Barlett's test of sphericity achieved a high level of statistical significance ($p < 0.001$), supporting the factorability of the correlation matrix.

Principal component analysis revealed the presence of at least three components with eigenvalues exceeding 1, following inspection of the scree plot (Table 4). Factor one conformed to affective suffering (depression, suicidality and tearfulness) with smaller contribution (0.4–0.5) from guilt, irritability and fatigue, and factor two indicated motivation (interest, concentration and enjoyment) but negligible loadings from other items. The affective suffering factor account for 26% of the variance in the items, and the motivation factors for 12%. The third factor, with eigenvalue marginally over 1, was largely loaded on by sleep and appetite.

Discussion

Depression is an important condition affecting older people's health and quality of life which needs a screening instrument to identify individuals at risk. Euro-D is a useful instrument which serves

Table 2. Euro-D item prevalences and scale scores ($N = 150$).

Symptom	N	Total %	Depressive patients diagnosed by MINI		p
			Yes ($n = 51$) %	No ($n = 99$) %	
Depression	76	50.7	79.7	31.9	0.000
Pessimism	85	56.7	64.4	51.7	0.123
Suicidality	40	26.7	50.9	11.0	0.000
Guilt	44	29.3	52.5	14.3	0.000
Sleep	84	56.0	76.3	42.9	0.000
Interest	47	31.3	54.2	16.5	0.000
Irritability	88	58.7	89.8	38.5	0.000
Appetite	64	42.7	71.2	24.2	0.000
Fatigue	86	57.3	93.2	34.1	0.000
Concentration	25	16.7	33.9	5.5	0.000
Enjoyment	40	26.7	44.1	15.4	0.000
Tearfulness	52	34.7	59.3	18.7	0.000

this purpose. To our knowledge, this is the first study to test the reliability and validity of Thai version of Euro-D. Attempts were made to ensure that content and semantic equivalence were achieved by strict translation and back-translation procedures with minor modifications to make the items more comprehensible and relevant to the Thai context. It was found to be feasible and easily administrable by trained interviewers in this clinical setting.

Our results show that the frequency of individual items rated as present during the previous month in this study was somewhat higher than those reported in a previous Euro-D studies in European countries (Prince et al., 1999). The prevalence of the pessimism item was particularly high and it did not discriminate well between the depressed and the nondepressed groups. Although we altered the original version of the item it was still the case that few older people volunteered any hopes for the future. The Buddhist doctrines of contentment and acceptance of the current life situation are likely to explain this view. The hope of better living in the future, for instance in terms of gaining material good, contradicts the traditional path toward merit in the afterlife (Pfanner & Ingersoll, 1962). Guilt was also highly prevalent. This may be due to people misunderstanding the guilt item or the tendency to volunteer feeling a burden on children and this being rated as a symptom. It is widely perceived among Thai older people that they should rely on their children when they get old and dependent (Knodel, Chayovan, & Siriboon, 1992), but many expressed guilt about this. Fatigue was also high, possibly because there were extra untreated physical illnesses or because many older Thai people are still working hard to make ends meet even after their retirement. Further studies using item response theory analysis to explore for differential item functioning (DIF) may be helpful to investigate whether there is item bias in the Thai version and to explain the high levels of these symptoms. The area under the ROC curve for the overall discriminability of the Euro-D scale against the criterion of any MINI depressive
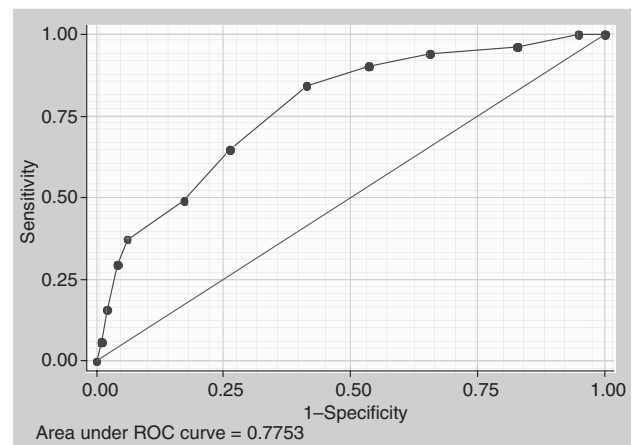

Figure 1. Area under ROC curve.

disorders was 0.78 (0.70–0.85). This rather impressive degree of overall prediction, coupled with the high sensitivity (94%) and poor specificity (34%) of the Euro-D cut-point of 3/4 suggested that this cut-point is too low for the Thai clinical sample. A cut-off score of 5/6 had lower sensitivity (84.3%) but higher specificity (58.6%) to screen major depressive episode compared with the usual cut-off reported previously. This higher cut-off might reflect more depressive symptoms among psychiatric patients. The majority of cases of depressive disorders in clinical settings obviously tend to be more severe compared with cases in general population. In community settings, cases may tend to have lower screening score. Further validation studies in the community setting will be needed. Taking account of the Thai context and improving the clarity of the wordings may help to improve the overall sensitivity and specificity.

The Cronbach alpha coefficient shows a high internal consistency for the whole Euro-D questionnaire, although the item on pessimism would appear to improve the overall internal consistency if deleted. As discussed earlier, the concept of pessimism may be different among Thai elderly. Finding a better way to

Table 3. Performance of the Thai Euro-D in screening major depressive disorder in older patients.

Cut-off score	Major depression			
	Sensitivity	Specificity	Positive predictive value	Negative predictive value
3/4	94.1	34.3	46.5	90.2
4/5	90.2	46.5	51.2	87.9
5/6	84.3	58.6	55.9	80.2
6/7	64.7	73.7	59.5	75.9
7/8	49.0	82.8	76.0	74.4

Table 4. Principal component analysis of Euro-D.

	Factor 1	Factor 2	Factor 3	Factor 4
Eigenvalue	3.1	1.5	1.1	1.0
Variance (%)	26	12	9	9
Depression	**0.65**	0.03	0.34	−0.18
Pessimism	−0.03	0.08	−0.02	**0.92**
Suicidality	**0.64**	0.26	−0.02	0.04
Guilt	0.56	0.19	0.02	−0.22
Sleep	−0.03	0.06	**0.79**	−0.21
Interest	0.09	**0.60**	0.37	0.17
Irritability	0.42	0.23	0.43	-0.17
Appetite	0.24	0.01	**0.69**	0.28
Fatigue	0.55	0.20	0.34	0.17
Concentration	0.17	**0.68**	0.00	0.10
Enjoyment	−0.00	**0.83**	0.01	0.01
Tearfulness	**0.73**	−0.12	−0.00	0.08

Bold figures indicate items loading at >0.60.

address the issue of pessimism may help to improve the overall psychometric properties.

One of the questions raised in a validation study is about what is the real construct behind the instrument to be validated. The Euro-D was mainly derived from GMS/AGECAT and designed to be used by trained lay interviewers, whereas MINI was based on ICD-10 and DSM-IV criteria and used by mental health professionals as a diagnostic tool for major depressive episode. These discrepancies may provide an explanation for the fair kappa for the agreement between Euro-D and MINI cases.

It is striking that the internal structure of the Euro-D as suggested by principal components analysis is consistent with most of the previous studies (Castro-Costa et al., 2007; Prince et al., 2004). Our version of the Euro-D also shares two common constructs, the first one being affective suffering, including depression, tearfulness and wishing to die, and second one being motivation including loss of interest, poor concentration and lack of enjoyment. They were mapped well onto the two common main factors reported previously, although there were slight differences in the individual items included in each factor (Castro-Costa et al., 2007; Prince et al., 1999; Prince et al., 2004). The two extra factors were identified in our study, one

including sleep and appetite and the other one including pessimism. The three items had eigenvalues of about 1. This may not be so surprising as three- or four-factor structure was also reported in some of the European settings. For example, sleep and appetite were found to be loading on another factor in Sweden, Netherlands and Iceland and pessimism loading on a separate factor in Germany (Prince et al., 1999). Although the four-factor structure is somewhat different, it remains difficult to establish that it does vary from the European studies. Confirmatory factor analysis may be a useful technique to test whether the four-factor model has better overall goodness-of-fit than the two-factor solution.

There were some limitations in this study. First, the study was conducted in a psychiatric hospital on a sample of patients with a variety of mental conditions. The study sample was obviously different from those based in communities or general practice. Second, the sample size available for this validation study was relatively small. More studies with a larger sample size in a variety of clinical settings are warranted. In addition, item bias cannot be excluded as an explanation for the relatively high prevalences of many Euro-D items in our study. Further studies in cross-cultural settings using item response theory analysis should provide more clues.

In conclusion, this is the first study to test the reliability and validity of Thai version of Euro-D. This validation study reports the use of Euro-D in a clinical setting. The instrument has demonstrated rather different properties to those shown in previous studies. The different cut-point found in the current study suggests the need for caution in interpretation of Euro-D scores in Thailand. Further studies are still needed to validate Euro-D in the Thai community.

References

Arunpongpisal, S., Assanangkornchai, S., Nimnuan, C., Kongsuk, T., Kittiratpaiboon, P., Taemepradit, A., et al. (2006). Reliability and validity study for Thai mental Illness screening test in community population. *Journal of the Psychiatric Association of Thailand, 51*(2), 22–38.

Asberg, M., & Schalling, D. (1979). Construction of a new psychiatric rating instrument, the Comprehensive Psychopathological Rating Scale (CPRS). *Progress in Neuro-psychopharmacology, 3*(4), 405–412.

Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Huppert, F., Mendonca-Lima, C., et al. (2007). Prevalence of depressive symptoms and syndromes in later life in ten

European countries: The SHARE study. *British Journal of Psychiatry, 191*, 393–401.

Copeland, J.R., Dewey, M.E., & Griffiths-Jones, H.M. (1986). A computerized psychiatric diagnostic system and case nomenclature for elderly subjects: GMS and AGECAT. *Psychological Medicine, 16*(1), 89–99.

Crawford, M.J., Prince, M., Menezes, P., & Mann, A.H. (1998). The recognition and treatment of depression in older people in primary care. *International Journal of Geriatric Psychiatry, 13*(3), 172–176.

Dearman, S.P., Waheed, W., Nathoo, V., & Baldwin, R.C. (2006). Management strategies in geriatric depression by primary care physicians and factors associated with the use of psychiatric services: A naturalistic study. *Aging & Mental Health, 10*(5), 521–524.

Department of Mental Health. (2003). The 2003 National Epidemiology Survey on Mental Health in Thailand. Bangkok: Department of Mental health.

Gurland, B., Golden, R.R., Teresi, J.A., & Challop, J. (1984). The SHORT-CARE: An efficient instrument for the assessment of depression, dementia and disability. *Journal of Gerontology, 39*(2), 166–169.

Kittiratpaiboon, P., & Wongkampin, M. (2004). The Validity of the Mini International Neuropsychiatric Interview – Thai version. Nonthaburi: Department of Mental health.

Knodel, J., Chayovan, N., & Siriboon, S. (1992). The familial support system of Thai elderly: An overview. *Asia-Pacific Population Journal, 7*(3), 105–126.

Koenig, H.G. (2007). Physician attitudes toward treatment of depression in older medical inpatients. *Aging & Mental Health, 11*(2), 197–204.

Penninx, B.W., Geerlings, S.W., Deeg, D.J., van Eijk, J.T., van Tilburg, W., & Beekman, A.T. (1999). Minor and major depression and the risk of death in older persons. *Archives of General Psychiatry, 56*(10), 889–895.

Penninx, B.W., Leveille, S., Ferrucci, L., van Eijk, J.T., & Guralnik, J.M. (1999). Exploring the effect of depression on physical disability: Longitudinal evidence from the established populations for epidemiologic studies of the elderly. *American Journal of Public Health, 89*(9), 1346–1352.

Pfanner, D.E., & Ingersoll, J. (1962). Theravada Buddhism and village economic behaviour: A Burmese and Thai comparison. *Journal of Asian Studies, 21*, 341–361.

Prince, M.J., Acosta, D., Chiu, H., Copeland, J., Dewey, M., Scazufca, M., et al. (2004). Effects of education and culture on the validity of the geriatric mental state and its AGECAT algorithm. *British Journal of Psychiatry, 185*, 429–436.

Prince, M.J., Reischies, F., Beekman, A.T., Fuhrer, R., Jonker, C., Kivela, S.L., et al. (1999). Development of the EURO-D scale – A European Union initiative to compare symptoms of depression in 14 European centres. *British Journal of Psychiatry, 174*, 330–338.

Radloff, L.S. (1977). The CES-D scale: A self report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., 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. *Journal of Cinical Psychiatry, 59*, (Suppl 20), 22–33.

Watts, S.C., Bhutani, G.E., Stout, I.H., Ducker, G.M., Cleator, P.J., McGarry, J., et al. (2002). Mental health in older adult recipients of primary care services: Is depression the key issue? Identification, treatment and the general practitioner. *International Journal of Geriatric Psychiatry, 17*(5), 427–437.

Zung, W.W. (1965). A Self-Rating Depression Scale. *Archives of General Psychiatry, 12*, 63–70.