Inter- and Intra-Rater Reliability of the Thai Version of SCAN: Use of Alcohol and Use of Tobacco Section

Suwanna Arunpongpaisal MD*, Thawatchai Krisanaprakornkit MD*,
Suchat Paholpak MD*, Jiraporn Keiwyoo PhD**

** Departments of Psychiatry, Faculty of Medicine, Khon Kaen University, Khon Kaen*
*** Department of Biostatistics and Demography,Faculty of Public Health, Khon Kaen University, Khon Kaen*

Objective: *To assess inter and intra rater reliability of the Thai version of Use of Alcohol and Use of Tobacco Section of the WHO Schedules for Clinical Assessment in Neuropsychiatry (SCAN).*
Material and Method: *Fifteen alcohol and/or tobacco dependence patients and fifteen controls with ages of ≥ 18 years were recruited from October 2003 to August 2004 at Srinagarind Hospital. One psychiatrist interviewed and first rated under video-recording, then re-rated two weeks later. Another psychiatrist rated independently by looking at the videotapes for inter-rater reliability testing.*
Results: *The intra-rater kappa was 'excellent' for both sections. The inter-rater kappa for all items of "tobacco use" were 'excellent'(mean κ =0.84), and 'good'(mean κ =0.66) for the "alcohol use". Items of dependence had 'good' kappa (mean κ =0.73-0.95), except items of 'activities limitation (interest neglect)' and 'use despite knowledge of psychological/physical problems (continued use)' was 'fair' (mean κ 062-0.49). The poor kappa (mean κ =0.17-0.38) was found in the item of 'physical and mental health problems due to drinking'.*
Conclusion: *Thai SCAN provided reliable inter-rater diagnostic information for alcohol and tobacco dependence, but fair reliability for alcohol abuse. Improve understanding of the item concepts, better validity in questioning; giving examples of the focus symptoms, and frequent discussion about the respondent's answers, might improve overall validity and reliability.*

Keywords: *Inter rater reliability, Intra-rater reliability, SCAN, Schedules for clinical assessment in neuropsychiatry,Alcohol use disorder, Tobacco use disorder,Validity*

J Med Assoc Thai 2006; 89 (12): 2129-37
Full text. e-Journal: *http://www.medassocthai.org/journal*

From a 2001 survey, it was estimated that 33 million Thais, had at some time in their lives, consumed alcohol (*viz.* 73% of the population between 12 and 65 years of age). Three-fifths (60%) of this population had drunk within the last year and 42% (18.6 million people) within 30 days of the interview. Among the 18.6 million, the most popular alcoholic beverages were beer (78%), Vodka-like products (46%), whiskeys/ hard liquors (34%), and wine/wine coolers (15%), medicinal herbal spirits (13%). Some people consumed more than one kind of alcoholic beverage. The wine coolers, a newly introduced product in Thailand, have out-paced traditional Chinese liquors in popularity (Sirivonge na Ayudhya & Sai-ngarn, 2004) because they are cheaper and taste sweet. This is important because wine coolers are perceived as having low alcohol content, thus people tend to drink them more frequently and to excess.

In a 2003 survey, it was estimated that 26.5 million Thais consumed alcohol (*i.e.* 58.5% of the population between 12 and 65 years of age), representing a reduction in all three of the 2001 survey-indicators: a) regularly consumed alcohol; b) consumed alcohol this year; and, c) consumed alcohol sometimes in the last 30 days. The number of people in 2003 who drank 30 days before the interview had decreased from

Correspondence to : Arunpongpaisal S, Department of Psychiatry, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand. Phone & Fax: 043-348-384, Mobile: 0-1708-0165, E-mail: suwaru@kku.ac.th

18.6 million (in 2001) to 15.8 million. However, by comparison, the number of regular drinkers (those who drank in 20 of the last 30 days) increased from 8.4 to 11.9 percent between 2001 and 2003, respectively.

Dual alcohol and tobacco use/dependence flags potentially serious behavioral and emotional problems since the association with physical health ailments is well documented. However, neither the 2001 nor the 2003 survey checked for any link between tobacco use and alcohol consumption. The estimated percentage of the population between 12 and 65 who smoked and drank in the past year was 20% (Sirivonge na Ayudhya & Sai-ngarn, 2004).

In the field of mental health, the importance of screening the clinical and general population for alcohol and substance use disorders is important. Detection instruments are chosen principally on the basis of their utility in the clinical setting (*e.g.* ease of assessment and scoring), the reliability and validity, and their relevance to clinical decision-making and the principal function of any scale is to alert clinicians to the need for more extensive assessment of substance-use patterns and substance-related symptoms.

The screening for alcohol use disorders and nicotine dependence includes the Michigan Alcoholism Screening Test (MAST), the Alcohol Use Disorders Identification Test (AUDIT), the CAGE Questionnaire, the TWEAK test and the Fagerstrom Test for Nicotine Dependence (FTND) (Schottenfeld & Pantalon MV, 1999). However, these scales are limited because of the high frequency of false positives and negatives reported by subjects.

A reliable and valid diagnostic process is an essential prerequisite to almost all research and clinical practice in the area of alcohol, substance use, and mental disorders. Since mental disorders are defined by their manifestations, rather than direct disease markers, reliability and validity assessment is a challenge. The Schedule for Clinical Assessment in Neuropsychiatry (SCAN: WHO, 1993) is the most prominent instrument developed for cross-cultural diagnosis and classification of mental disorders, alcohol- and substance-related problems. This instrument has been extensively field-tested in a variety of cultural settings for mental disorders including alcohol and substance use disorders (Ustun et al, 1997) and yielded useful information on the reliability and validity of SCAN for diagnosis, symptomatologic criteria, and difficulty of the question level.

Previous studies of the English version of SCAN demonstrated its overall diagnostic concordance coefficients (κ_s) were 'very good' for dependence disorders (range, 0.7 to 0.9), but somewhat lower for the abuse and harmful use categories.

The authors' one-week, pilot test-retest comparison of SCAN for alcohol dependence had a higher concordance coefficient ($\kappa = 0.76$ for ICD-10 *vs* 0.73 for DSM-IV) than alcohol abuse ($\kappa = 0.35$ for ICD-10 *vs* 0.60 for DSM-IV). Calculations on the reliability of SCAN at the item and criterion levels for alcohol showed that the 'impaired control' questions were associated with the lowest test-retest agreement ($\kappa = 0.57$).

Wilson Compton introduced SCAN to Thailand in 2003 at Siriraj (Bangkok) and Srinagarind (Khon Kaen) Hospitals. Two psychiatrists (SP and TK) were trained to use SCAN at a standard 5-day training course in the UK. When they returned to Thailand, they commenced validity and reliability research on a translated Thai version of SCAN. The authors' sub-study is part of that validity assessment research with an aim to determine the inter- and intra-rater reliability of the Thai version of SCAN at the item level.

Material and Method

Subjects and setting

The present study focused on 15 alcohol- and tobacco-dependent subjects either in-/out-patients from Srinagarind University Hospital, Khon Kaen, Thailand, between October 2003 and August 2004. Another 15 sex- and age-matched subjects, with no history of alcohol or tobacco use, were selected from among the friends and/or relatives of the patients. The study protocol was reviewed and approved by the Human Research Ethics Committee at Khon Kaen University and all respondents signed informed consent before participating. The subjects were given 500 Baht to cover their travel and time.

Instruments

SCAN is a semi-structured interview conducted by a clinically-experienced psychiatrist. The use of the Comprehensive SCAN Glossary is essential as symptoms are explored and entered so that a diagnosis can be generated using computerized algorithms. The two psychiatrists (SP and TK) were trained in the use of SCAN at a standard 5-day training course in the UK.

Various control measures were applied including: the distribution of field-tested interview materials and data entry software, which included logic and consistency checks, central editing, and co-rating the interviews. Translation of SCAN into Thai and back

translation (to English) were done by SP et al. (2003). The comprehensibility of the language was pre-tested in the four (ethnic cum linguistic) regions of Thailand (i.e. the Center at Srithunya Hospital in Nonthaburi; the North at Suanprung Hospital in Chiang Mai; the South at Suansarqanrom Hospital in Suratthani; and the Northeast at Jittavej-KhonKaen- Rachanagharind Hospital in Khon Kaen.

The resulting comments and suggestions for possible modifications were summarized and re-discussed among groups representing people from those regions until unanimous agreement was reached for the final Thai wording of each item. Back translation was used to assess conservation of the original English meaning in the final Thai version. With permission from WHO, the final Thai version was incorporated into the SCAN-Shell computer program. The authors' then did validity testing of Sections 11 (Use of Alcohol) and 12 (Use of Psychoactive Substance other than Alcohol - we specified tobacco use).

Design

The authors' sample comprised 30 subjects (15 alcohol- and tobacco-use-disordered patients and 15 normal persons as control group). The ICD-10 or DSM-IV criteria were used to identify the cases from among patients being served at In-/Out-Patient Services.

All of the subjects had to be Thais, 18 years of age or over. The two interviewers (SP and TK) conducted the inter-rater reliability test. One interviewer was video-recorded as he interviewed each subject; the other interviewer independently rated these recorded interviews. For a comparison of intra-rater reliability, the psychiatrists re-rated the video material two weeks later.

Statistical analysis

The chance-corrected, degree of agreement on each symptom-question for SCAN Sections 11 and 12 was calculated using the kappa statistic (Fleiss, 1981). Both the inter- and intra-rater agreements were computed. Kappa values under 0.40 indicated 'poor' agreement; 0.41-0.60 'fair'; 0.61-0.80 'good'; and > 0.81 'excellent'.

Results

The participants (76.7% male) averaged 42 $\pm$ 9 years of age (range, 20-61). Most were married, had an average of 11.3 years of schooling, and had a middle socioeconomic rank (Table 1).

The κ values for both inter- and intra-rater reliability, which showed 'excellent' concordance between the raters for all items of tobacco use disorders (mean inter-rater $\kappa = 0.84$), and 'good' concordance for alcohol use disorders (mean inter-rater $\kappa = 0.66$). The concordance of a single rater rating two times (after a two-week interval) was 'excellent'(mean intra-rater $\kappa = 0.82$-0.87) for both sections (Table 2).

The itemized κ_s of both inter- and intra-rater reliability about alcohol dependence symptoms were 'good' except items D6 'Activities limitation (Interest neglect)' and D7 'Use despite knowledge of psychological/physical problems (continue use)' that produced only 'fair' agreement. The items about physical and mental health problems due to drinking produced 'poor' agreement (Table 3).

The κ values for both inter- and intra-rater reliability for items about tobacco dependence criteria were 'good' (Table 4) except for items about 'activity limitation' (interest neglect), which were 'fair'. Items about *"feel your heart rate slow down and having heart trouble due to tobacco"* had lower intra- than inter-rater agreement.

The 'cascade of reliability' by itemized κ_s, item 11.023 'physical health problems due to alcohol drinking' had the lowest inter-rater agreement; not withstanding, the inter-rater agreement for items on tobacco dependence were mostly 'good' to 'excellent', except for the item on criteria for activity limitation ('restriction of social activities due to smoking' and 'salience of smoking-related activities'). The intra-rater agreement for items asking about tobacco and alcohol use were 'good' to 'excellent' except for items asking about symptoms of heart trouble due to tobacco (Table 5).

Discussion

The authors investigated the reliability-testing of the Alcohol and Tobacco Use Section of the Thai version of SCAN. The size of the sample was kept small because of limited time and money. All of the cases were men while the controls were both men and women. Therefore, the results from the present study cannot be generalized to the general population because there was no information of alcoholic or smoking women. However, the authors completed the essential reliability and validity evaluations of the alcohol and tobacco sections of the Thai version of SCAN, and examined ways of improving the classification and assessment systems.

The Thai version of SCAN was reliable with

respect to inter-rater diagnosis of both alcohol and tobacco dependence, but its reliability was lower for diagnosis of alcohol abuses. The reason may be denial defense of alcohol abuser who did not recognize the adverse effect from alcohol but the interview psychiatrist had recognized it by direct observation. There was some significant variation in the inter-rater κ_s on items about alcohol and tobacco dependence. Tobacco dependence had a higher agreement than alcohol dependence because there are fewer variations in the pattern of smoking than drinking. Hence, the translation of diagnostic criteria should be made more detailed in the pattern of alcohol use for better operational assessments.

Inter-rater κ_s for each item and criteria suggested inferences about the validity of the concept of the 'dependence syndrome' according to Thai culture and vernacular. A lower inter-rater reliability occurred when detecting heavier daily alcohol use (*i.e.* the D6 criteria on limitations to activity such as 'narrowing drinking pattern', 'restriction of social activities due to smoking' and 'salience of smoking-related activities') and D7 on continued use (*i.e.* 'despite being aware of psychological/physical problems') perhaps because of: 1) unclear questioning; 2) difficulty in understanding the item questions; 3) cultural differences in interpreting the severity of drinking or defensiveness of the respondent.

Muthen et al (1993) studied the DSM-III-R and draft DSM-IV alcohol dependence criteria items among US alcohol users and he developed two models, the first, that detected heavy social drinking and, the second, a relevant problematic drinking.

Nelson et al (1999) studied the factor-structures for DSM-IV substance use disorder criteria among users of alcohol, cannabis, cocaine and opioids and they found that the DSM-IV criterion D7 'use despite problems' was central to only alcohol dependence; indicating that even with the newly attempted diagnostic criteria of the ICD-10 and DSM-IV, common conceptualization and operationalization have not yet reached a satisfactory level. Perhaps this is due to the subjectivity of social criteria of (*i.e.* 'harm' and 'social activities in daily life') has variable under-

Table 1. Sample characteristics

Main substance use	N	Mean age	Mean years of education	% Male	% unemployed	% married	Sample Control group	Patient group	Cultural description
Alcohol, Nicotine	30	42	11.3	76.7	0	70	15	15	Lower and middle socioeconomic status, Northeastern Thai, Buddhism

Table 2. Kappa values of both inter- and intra-rater reliability for all items in Chapters 11 (Use of alcohol) and 12 (only items on use of tobacco)[a]

	Inter-rater reliability		Intra-rater reliability	
	Chapter 11 Alcohol use disorder	Chapter 12 Tobacco use disorder	Chapter 11 Alcohol use disorder	Chapter 12 Tobacco use disorder
Total number of items	32	34	32	34
Mean κ	0.66	0.84	0.82	0.87
Median κ	0.68	0.85	0.80	0.92
Mode κ	0.60	1.00	1.00	1.00
Standard deviation	0.19	0.14	0.10	0.14
Minimum	0.17	0.48	0.64	0.47
Maximum	0.95	1.00	1.00	1.00

a = Fleiss 1981

Table 3. Inter- and intra-rater reliability of items asking about ICD-10/DSM-IV alcohol use disorder

Criteria D (dependence): A (abuse)	Item number	Reliability			
		Inter-rater		Intra-rater	
		Kappa	95%CI	Kappa	95%CI
Compulsive use		0.74		0.87	
- frequency of drinking	11.002	0.90	0.71-1.00	0.95	0.76-1.00
- estimated usual daily amounts of alcohol	11.003	0.60	0.45-0.75	0.75	0.59-0.91
- heavier daily alcohol	11.004	0.56	0.41-0.71	0.67	0.52-0.83
- frequency of heavier daily of alcohol	11.005	0.91	0.72-1.00	1.00	0.81-1.00
- most recent drink	11.006	0.91	0.71-1.00	0.87	0.68-1.00
- subjective need for alcohol (Past year)	11.007	0.79	0.54-1.00	0.84	0.57-1.00
- current and past drinking (Past year)	11.032	0.51	0.35-0.67	1.00	0.80-1.00
D1 Tolerance of alcohol (Past year)	11.018	0.73	0.51-0.94	0.95	0.74-1.00
D2 Withdrawal		0.81		0.90	
- alcohol withdrawal problems (Past year)	11.025	0.73	0.52-0.94	0.86	0.66-1.00
- drinking to relieve or avoid withdrawal symptoms (Past year)	11.026	0.89	0.61-1.00	0.94	0.66-1.00
D3 More than intended (impaired control)					
- impaired capacity to control drinking once started (Past year)	11.009	0.77	0.56-0.98	0.69	0.48-0.89
D4 Unsuccessful cut-backs (impaired control)					
- impaired capacity to abstain or cut drinking (Past year)	11.008	0.95	0.70-1.00	0.95	0.70-1.00
D5 Great deal of time (interest neglect)					
- time involved in drink-related activities (Past year)	11.016	0.83	0.55-1.00	0.83	0.55-1.00
D6 Activities limitation (Interest neglect)		0.62		0.81	
- salience of drink-related activities (Past year)	11.015	0.70	0.40-1.00	0.82	0.52-1.00
- narrowing of drinking pattern (Past year)	11.017	0.54	0.31-0.78	0.80	0.55-1.00
D7 Continue use despite knowledge of psychological/ physical problem		0.49		0.81	
- persistent drinking after social and legal harm (Past year)	11.012	0.55	0.23-0.87	0.89	0.53-1.00
- psychological/mental health problems due to drinking	11.019	0.38	0.20-0.56	0.80	0.58-1.00
- circumstances of onset of alcohol induced MH problem	11.020	0.43	0.28-0.59	0.70	0.52-0.88
- interference due to alcohol induced symptoms (Past year)	11.021	0.64	0.44-0.84	0.78	0.58-0.98
- persistent drinking after mental harm (Past year)	11.022	0.78	0.46-1.00	0.92	0.60-1.00
- physical health problems due to drinking (Past year)	11.023	0.17	-0.01-0.35	0.70	0.46-0.93
- persistent drinking after physical harm (Past year)	11.024	0.37	0.09-0.65	1.00	0.64-1.00
- interference activities because of alcohol drinking (Past year)	11.037	0.60	0.40-0.80	0.69	0.49-0.89
Harmful use (Abuse)		0.72		0.78	
A1 failure to fulfill major role obligations (Past year)	11.013	0.84	0.57-1.00	0.83	0.56-1.00
A2 risk taking behavior with alcohol (Past year)	11.014	0.66	0.40-0.92	0.77	0.50-1.00
A3 legal problems due to drinking (Past year)	11.011	0.64	0.38-0.90	0.79	0.49-1.00
A4 social problems due to drinking (Past year)	11.010	0.72	0.51-0.93	0.72	0.51-0.93
Pathological or idiosyncratic reaction to alcohol (Past year)	11.027	0.57	0.33-0.82	0.80	0.53-1.00
Acute toxic effects of alcohol, uncomplicated (Past year)	11.028	0.83	0.66-1.00	0.79	0.63-0.96
Trauma, bodily injury due to alcohol (Past year)	11.029	0.38	0.18-0.57	0.73	0.48-0.98
Hematemesis, aspiration, other complications (Past year)	11.030	0.42	0.20-0.65	0.64	0.40-0.87
Change of consciousness or perceptual abnormal due to alcohol	11.031	0.86	0.62-1.00	0.75	0.53-0.96

Table 4. Inter-and intra-rater reliability of items asking about ICD-10/DSM-IV tobacco use disorder

Criteria	Item number	Relability			
		Inter-rater		Intra-rater	
		Kappa	95%CI	Kappa	95%CI
Compulsive use		0.87		0.90	
- Lifetime use of tobacco	12.060	0.93	0.57-1.00	0.92	0.57-1.00
- Cigarettes (number of cigarettes)	12.061	0.95	0.76-1.00	0.95	0.76-1.00
- Cigarettes (number)	12.066	0.95	0.78-1.00	0.95	0.78-1.00
- Subjective need for tobacco (Past year)	12.073	0.66	0.41-0.92	0.77	0.51-1.00
D1 Tsolerance to tobacco (Past year)	12.080	0.66	0.37-0.96	0.77	0.51-1.00
D2 Withdrawal		0.87		0.90	
- craving for tobacco	12.090	0.94	0.61-1.00	1.00	0.67-1.00
- get irritable and angry (tobacco)	12.091	0.86	0.54-1.00	0.93	0.61-1.00
- get anxious (tobacco)	12.092	0.89	0.59-1.00	1.00	0.69-1.00
- have difficulty in concentration (tobacco)	12.093	0.88	0.60-1.00	0.88	0.60-1.00
- become restless (tobacco)	12.094	0.74	0.44-1.00	0.91	0.61-1.00
- have increased appetite and gain weight (tobacco)	12.095	0.77	0.46-1.00	0.78	0.46-1.00
- feel your heart rate slow down (tobacco)	12.096	0.84	0.55-1.00	0.53	0.24-0.82
- headaches, malaise, weakness (tobacco)	12.097	0.79	0.51-1.00	1.00	0.72-1.00
- drowsiness or insomnia (tobacco)	12.098	0.79	0.51-1.00	1.00	0.71-1.00
- hands shaking (tobacco)	12.099	1.00	0.72-1.00	1.00	0.72-1.00
- feeling of depression (tobacco)	12.100	0.79	0.51-1.00	0.79	0.51-1.00
- upset stomach or mouth ulceration	12.101	1.00	0.71-1.00	0.79	0.51-1.00
- increased cough after cessation	12.102	1.00	0.64-1.00	1.00	0.64-1.00
- smoking in order to avoid withdraw alsymptoms	12.103	0.84	0.52-1.00	0.92	0.60-1.00
D3 More than intended (impaired control)					
- impaired capacity to control smoking once started (Past year)	12.075	0.70	0.46-0.93	0.70	0.46-0.93
D4 Unsuccessful cut-backs (impaired control)					
- impaired capacity to abstain or reduce tobacco use (Past year)	12.074	0.88	0.61-1.00	0.94	0.63-1.00
D5 Great deal of time (interest neglect)					
- time involved in smoking-related activities (Past year)	12.079	1.00	0.72-1.00	1.00	0.72-1.00
D6 Activities limitation (Interest neglect)		0.59		0.74	
- restriction of social activities due to smoking (Past year)	12.076	0.59	0.30-0.87	0.77	0.46-1.00
- salience of smoking-related activities (Past year)	12.078	0.48	0.27-0.70	0.66	0.39-0.92
- restriction of physical activities due to smoking (Past year)	12.081	0.69	0.40-0.97	0.79	0.51-1.00
D7 Continue use despite knowledge of psychological/ physical problems		0.88		0.87	
- persistent smoking after social restriction (Past year)	12.077	0.67	0.37-0.98	1.00	0.67-1.00
- cough due to smoking	12.082	0.86	0.54-1.00	0.78	0.47-1.00
- susceptible to respiratory infections (tobacco)	12.083	0.80	0.50-1.00	0.91	0.59-1.00
- bronchitis due to tobacco	12.084	1.00	0.68-1.00	0.91	0.60-1.00
- high blood pressure due to tobacco	12.085	1.00	0.71-1.00	1.00	0.71-1.00
- heart trouble due to tobacco	12.086	1.00	0.72-1.00	0.47	0.19-0.75
- cancer of the lung due to tobacco	12.087	1.00	0.64-1.00	1.00	0.64-1.00
- toxic effects of nicotine	12.088	0.74	0.45-1.00	0.79	0.49-1.00
- persistent tobacco use after physical harm	12.089	0.81	0.49-1.00	1.00	0.67-1.00

Table 5. Summary of inter-and intra-rater kappa of SCAN interview items

Kappa value / Substance	Inter-rater kappa					Inter-rater kappa				
	Slight 0.00-0.20	Poor 0.21-0.40	Fair 0.41-0.60	Good 0.61-0.80	Excellent 0.81-1.00	Slight	Poor	Fair	Good 0.61-0.80	Excellent 0.81-1.00
Alcohol	11.023	11.019 11.024 11.029	11.003 11.004 11.012 11.017 11.020 11.027 11.030 11.032 11.037	11.007, 11.009 11.010, 11.011 11.014 11.015 11.018 11.021 11.022 11.025	11.002 11.005 11.006 11.008 11.013 11.016 11.026 11.028 11.031	-	-	-	11.003,11.004 11.009,11.010 11.011,11.014 11.017,11.019 11.020,11.021 11.023,11.027 11.028,11.029 11.030,11.031 11.037	11.002,11.005 11.006,11.007 11.008,11.012 11.013,11.015 11.016,11.018 11.022,11.024 11.025,11.026 11.032
Nicotine			12.076 12.078	12.073 12.075 12.077 12.080 12.081 12.083 12.088 12.094 12.095 12.097 12.098 12.100	12.060 12.061 12.066 12.074 12.079 12.082 12.084 12.085 12.086 12.087 12.089 12.090 12.091 12.092 12.093 12.096 12.099 12.101 12.102 12.103			12.086 12.096	12.076,12.078 12.073,12.075 12.080,12.081 12.088,12.095 12.100,12.082 12.101	12.060,12.061 12.066,12.074 12.077,12.079 12.083,12.084 12.085,12.087 12.089,12.090 12.091,12.093 12.094,12.097 12.098,12.099 12.102,12.103

standable content than the more observable criteria of withdrawal, tolerance, and craving. The item-criteria cascade of inter-rater reliability gives some clues on how these operationalizations could be improved to give better reliability in our assessment, *viz.:* better understanding of the item concepts, validity in questioning, examples of focused symptoms and frequent discussion about the respondent's answers.

Conclusion

The Thai version of SCAN provided high inter-rater reliability for diagnosis of both alcohol and tobacco dependence, but low inter-rater reliability for diagnosis of alcohol abuse. If the interviewer using SCAN understands the concept of each item, give examples of the focus symptoms, and discuss with the respondents about their understanding, it would improve the reliability and accuracy of this diagnostic test.

Acknowledgements

The present study was done with guidance from the WHO-SCAN Training and Reference Center in London (UK). The authors wish to thank Khon Kaen University Research Affairs Division for funding support and Mr. Bryan Roderick Hamman for his assistance with the English-language presentation of the manuscript.

References

1. Compton WM, Cottler LB, Dorsey KB, Spitznagel EL, Mager DE. Comparing assessments of DSM-IV substance dependence disorders using CIDI-SAM and SCAN. Drug Alcohol Depend 1996; 41: 179-87.
2. Fleiss J. Statistical methods for rates and proportions. 2nd ed. New York: Wiley; 1981.
3. Muthen BO, Grant B, Hasin D. The dimensionality of alcohol abuse and dependence: factor analysis of DSM-III-R and proposed DSM-IV criteria in the 1988 National Health Interview Survey. Addiction 1993; 88: 1079-90.
4. Nelson CB, Rehm J, Ustun TB, Grant B, Chatterji S. Factor structures for DSM-IV substance disorder criteria endorsed by alcohol, cannabis, cocaine and opiate users: results from the WHO reliability and validity study. Addiction 1999; 94: 843-55.
5. Sirivonge na Ayudhya A, Sai-ngarn J. Alcohol in Thai social: situation of alcohol use and the alcohol policy. Abstract of 2nd National conference on substance abuse, Khon Kaen Province, Thailand, 31st August - 2nd September 2004: 17-9.
6. Schottenfeld RS, Pantalon MV. Assessment of the patient. In: Galanter M, Klebler HD, editors. The American psychiatric press textbook of substance abuse treatment. 2nd ed. Washington, DC: American Psychiatric Press; 1999: 114-5.
7. Ustun B, Compton W, Mager D, Babor T, Baiyewu O, Chatterji S, et al. WHO study on the reliability and validity of the alcohol and drug use disorder instruments: overview of methods and results. Drug Alcohol Depend 1997; 47: 161-9.

ความเที่ยงของผู้ประเมินสองคนและความเที่ยงของผู้ประเมินคนเดียวกัน ณ ต่างเวลา เมื่อใช้เครื่องมือ SCAN ฉบับภาษาไทยเกี่ยวกับการใช้สุราและบุหรี่

สุวรรณา อรุณพงค์ไพศาล, ธวัชชัย กฤษณะประกรกิจ, สุชาติ พหลภาคย์, จิราพร เขียวอยู่

วัตถุประสงค์: *เพื่อทดสอบความเที่ยงของการให้คะแนนจากผู้ประเมินสองคน (inter-rater reliability) และความเที่ยงของการให้คะแนนจากผู้ประเมินคนเดียวกันเมื่อให้คะแนนต่างเวลากัน (intra-rater reliability) แก่คำตอบต่อคำถามแต่ละข้อของ SCAN ฉบับภาษาไทยหมวดการใช้สุราและบุหรี่*

วัสดุและวิธีการ: *การวิจัยนี้เป็น diagnostic agreement study กระทำในระหว่าง ตุลาคม พ.ศ. 2547 ถึง มีนาคม พ.ศ. 2548 คณะผู้วิจัยได้แปลบทสัมภาษณ์หมวดการใช้สุราและบุหรี่ของ WHO Schedules for Clinical Assessment in Neuropsychiatry (SCAN) Version 2.1 ให้เป็นภาษาไทย หลังจากนั้นมีการตรวจสอบและแก้ไขภาคภาษาไทยอีกครั้งหนึ่งเพื่อให้มีความหมายแม่นตรงกับความหมายเดิมด้วยการตรวจสอบความหมายในภาคภาษาอังกฤษที่แปลกลับจากภาคภาษาไทยว่ามีความหมายตรงกับภาคภาษาอังกฤษต้นฉบับหรือไม่ แล้วนำไปทดลองใช้กับอาสาสมัครจาก 4 ภาคของประเทศไทยภาคละ20คนรวม ใช้เวลา 2 เดือน แล้วปรับปรุงสำนวนเป็นครั้งสุดท้าย จากนั้นจึงทดสอบความเที่ยงของ SCAN Interview ฉบับภาษาไทยที่ได้ โดยจิตแพทย์ 1 คนใช้ SCAN Interview book ฉบับภาษาไทยหมวดการใช้สุราและบุหรี่สัมภาษณ์อาสาสมัครที่เป็นผู้ป่วยที่มีปัญหาจากการใช้สารเสพติดชนิดใด ๆ ตามเกณฑ์ ICD-10 โดยคัดเลือกมาจากผู้ป่วยนอกหรือผู้ป่วยในที่มาแผนกจิตเวชจำนวน15 คน และสัมภาษณ์ญาติผู้ป่วยที่เป็นปกติอีก 15 คน รวมทั้งสิ้น 30 คน มีการถ่ายบันทึกการสัมภาษณ์ไว้ในแถบวีดิทัศน์ด้วย อาสาสมัครทั้งหมดเป็นคนไทยอายุ 18 ปีขึ้นไป ทุกคนยินดีลงชื่อแสดงความสมัครใจเข้าร่วมกับการวิจัย จากนั้นจิตแพทย์คนเดิมนำแถบวีดิทัศน์มาให้คะแนนใหม่อีกครั้งหนึ่งห่างจากครั้งแรก 2 สัปดาห์โดยไม่ทราบคะแนนเดิม เพื่อหาค่า intra-rater reliability ส่วนการหาค่า inter-rater reliability ได้ให้จิตแพทย์คนที่ 2 ดูแถบวีดิทัศน์ ข้างต้น แล้วให้คะแนนแก่ข้อมูลโดยไม่ทราบคะแนนของจิตแพทย์คนก่อน

สถิติที่ใช้: *Kappa และ Weighted Kappa ค่า Kappa ที่ต่ำกว่า 0.40 แสดงว่า 'poor' agreement ถ้าอยู่ระหว่าง 0.41-0.60 แสดงว่า 'fair'agreement ถ้าอยู่ระหว่าง 0.61-0.80 แสดงว่า 'good' agreement และถ้า สูงกว่า 0.81 แสดงว่า 'excellent' agreement*

ผลการศึกษา: *อาสาสมัครส่วนใหญ่เป็นเพศชาย (คิดเป็นร้อยละ 80) อายุเฉลี่ย 42 $\pm$ 9 ปี (ช่วงอายุ 20-61 ปี) พบค่าความเที่ยงในจิตแพทย์คนเดียวกันดีมาก(mean intra-rater k = 0.82-0.87)ทั้งหมวดการใช้สุราและบุหรี่ ค่าความเที่ยงระหว่างจิตแพทย์ 2 คนในทุกข้อคำถามของหมวดการใช้บุหรี่ดีมาก (mean inter-rater k = 0.84) แต่หมวดสุรามีน้อยกว่า (mean inter-rater k = 0.66) ข้อคำถามทั้งสองหมวดที่มีความเที่ยงในระหว่างจิตแพทย์พอใช้ คือคำถามเกี่ยวกับการใช้เวลาจำนวนมากไปกับการใช้สุรา/บุหรี่ หรือกิจกรรมอื่น ๆ ถูกจำกัด(mean inter-rater k = 0.59-0.62) และ ยังคงใช้สุราอย่างต่อเนื่องแม้ได้รู้แล้วว่ามีปัญหาทางสุขภาพจิต/กายข้อใดข้อหนึ่ง(mean inter-rater k = 0.49) ข้อคำถามที่ค่าความเที่ยงระหว่างจิตแพทย์ 2 คนไม่ดี คือคำถามเกี่ยวกับปัญหาทางสุขภาพกายและปัญหาทางสุขภาพจิตอันเนื่องมาจากการดื่มสุรา (mean inter-rater k = 0.17-0.38)*

สรุป: *แบบสัมภาษณ์ SCAN ฉบับภาษาไทย หมวดการใช้สุราและการใช้บุหรี่มีความถูกต้องด้านเนื้อหา และมีความน่าเชื่อถือ มี intra และ inter-rater reliability สูงโดยเฉพาะการวินิจฉัยโรคติดสุราหรือโรคติดบุหรี่แต่สำหรับการวินิจฉัยโรคดื่มสุราแบบอันตรายจะมีความน่าเชื่อถือน้อยกว่า แนวทางการแก้ไข เพื่อเพิ่มความน่าเชื่อถือและถูกต้องคือผู้ใช้ SCAN ควรพยายามทำความเข้าใจเกี่ยวกับกรอบแนวคิดของข้อคำถามแต่ละข้ออย่างละเอียด เน้นการยกตัวอย่างชัดเจนและมีการอภิปรายร่วมกันระหว่างผู้ใช้*