CAGE酗酒篩檢問卷在台灣基層醫療院所的跨文化效度測試

郭千哲 [1,3]　　陳爲堅 [2]　　鄭泰安 [1,*]

CHIAN-JUE KUO[1,3], WEI J. CHEN[2], ANDREW T.A. CHENG[1]

[1] 中央研究院生物醫學科學研究所，台北市南港區研究院路二段128號
Institute of Biomedical Sciences, Academia Sinica, No. 128, Yen Chiun Yuan Rd Sec 2, Taipei, Taiwan.
[2] 國立台灣大學公共衛生學院流行病學研究所
Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei, Taiwan.
[3] 台北市立療養院
Taipei City Psychiatric Center, Taipei, Taiwan.
*通訊作者Correspondence author.　E-mail:bmandrew@ccvax.sinica.edu.tw

　　本研究的目的在於檢驗CAGE酗酒篩檢問卷在台灣地區的跨文化效度。研究樣本(N=198)來自一基層醫療院所門診個案的隨機抽樣。研究者加入與台灣飲酒文化相關，且在跨文化心理語言表達上與原CAGE的"Cut-down" (C_1, C_2)和"Annoyance" (A_1)對等的三題，形成實驗性中文版CAGE (the experimental Chinese CAGE, C-CAGE); 該問卷可以用來評估五種計分方法：CAGE, C_1A_1GE, C_1AGE, C_2A_1GE and C_2AGE。 所有樣本在完成C-CAGE之後均接受臨床醫師的半結構性臨床面談，以獲得酗酒的診斷。和過去西方國家的報告比較，原CAGE在本研究所表現的篩檢效能位於下限值。C_2AGE的特異度和整體準確性，顯著優於原CAGE，而C_2A_1GE則有較好的敏感度和較大的ROC曲線下面積值。以上均顯示從心理語言性的對等考量下，原CAGE仍有進一步改進的空間。(中華衛誌 1999；**18**(2)：87-94)

　關鍵字：跨文化差異，酒精傷害性使用，酒精依賴，基層醫療。

Validity of the CAGE questionnaire in a primary care setting in Taiwan: a cross-cultural examination

　　An easy to use screening instrument for drinking behavior or alcohol-related illness has not been established in Taiwan. The authors examined the cross-cultural validity of the CAGE questionnaire in Chinshan, Taipei County in Taiwan. Study subjects (N=198) were outpatients randomly drawn from the primary care setting in the township. Three psycholinguistic items, equivalent to "Cut-down" (C_1, C_2) and "Annoyance" (A_1) of CAGE, thought to be culturally relevant for people in Taiwan, were added to form an experimental Chinese CAGE (C-CAGE). This allowed for tests of five types of scoring combination, CAGE, C_1A_1GE, C_1AGE, C_2A_1GE and C_2AGE. Each subject received a semi-structured clinical interview for diagnosis of alcoholism after he or she completed a C-CAGE questionnaire. The performance of CAGE in our sample was at the lowest limit among those reported in Western countries. Compared to CAGE, C_2AGE was superior in robustness for specificity and overall accuracy. C_2A_1GE tended to be better in sensitivity and performance for the value of the area measured by receiver operating characteristics curve analysis. This indicates that CAGE has room for further improvement under the assumption of psycholinguistic equivalence. *(Chin J Public Health. (Taipei): 1999; **18**(2):87-94)*

Key words: Cross-cultural difference, Harmful use of alcohol, Alcohol dependence, Primary care.

The prevalence of alcoholism has increased prominently in Taiwan in the past four decades. An epidemiological study conducted in 1946-1948 found that the prevalence of alcoholism detected by clinical interview among Han Chinese was 0.01% [1]. On the basis of a survey using the Diagnostic Interview Schedule (DIS) in 1982-1985, the prevalence of alcoholism were 4.9% in metropolitan area and 10.0% in townships in Taiwan [2]. Despite the considerable differences in research methodology between the two surveys, the evidence does imply a substantial increase of alcoholism in Taiwan. How to detect individuals suffering from alcoholism for timely intervention is thus an important public health issue.

Individuals suffering from alcohol-related illnesses often seek help at primary care settings. However, the detection rate of alcoholism by primary care physicians is low [3]. An easy to use screening instrument would help alert physicians of patients' drinking problem. In this regard, self-administered questionnaires are less invasive, cheaper, more rapid and efficient than biological markers in cases identification [4]. Two popular screening questionnaires, CAGE [5] and the Michigan Alcoholism Screening Test (MAST) [6], have been demonstrated to have excellent validity among psychiatric patients, general practice attenders and other risk groups [7]. In general, MAST is less frequently used than CAGE due to its intimidating natures and a complicated scoring system [8].

The performance of these instruments for alcoholism screening, however, are not equally satisfactory across gender and ethnic subgroups [9]. Furthermore, applying of a well-developed instrument to a different culture may encounter problems related to culture-specific occurrence or expression of symptoms as well as those due to a language barrier [10]. Several researchers

have modified the original CAGE or MAST for particular populations, e.g., females [11], Swedish males [12], and Japanese [13]. However, the impact of cross-cultural application on the validity of CAGE was less investigated. In this study, we examined the validity of CAGE among Han Chinese in a primary care setting in Taiwan. Meanwhile, the prevalence of alcoholism among attenders in this setting was also estimated for the first time.

METHODS

The sample and fieldwork

Subjects of this study were sampled from the outpatient clinic at the Chinshan Health Station, which is the only primary care center in Chinshan Township of Taipei County. One hundred and ninety-eight individuals aged 15 and above were systematically drawn from a total of 1432 outpatients according to daily registered number (one per seven). A research assistant first asked each subject about the consumption of alcohol. Subjects who had ever drunk alcohol were asked to complete the experimental Chinese CAGE (C-CAGE). Those could not to read were interviewed by the assistant. For those who had never consumed alcohol, only sociodemographic data were collected. The alcoholic status of the subjects who answered the screening questionnaire was assessed blindly by a psychiatrist (C.J.K.) with a semi-structured clinical interview for alcoholism.

The case-finding instrument used in the semi-structured clinical interview was described in detail in a previous report with satisfactory reliability of the instrument [14]. The diagnosis of harmful use of alcohol and alcohol dependence were made according to ICD-10 Classification of Mental and Behavioral Disorders: Diagnostic Criteria for Research (ICD-10-DCR) [15].

The Chinese version of the CAGE

A two-stage translation of the original four CAGE items was first exercised. Then three additional items thought to be psycholinguistically equivalent to original items C and A in Chinese culture were added to form the C-CAGE. Thus, the C-CAGE contains the following items: (1)Have you ever felt you ought to cut down on your drinking?(denoted as C) (2)Have people annoyed you by criticizing your drinking? (denoted as A) (3)Have you ever felt bad or guilty about your drinking?(denoted as G) (4)Have you ever had a drink first thing in the morning to steady your nerves or get rid of a hangover? (denoted as E) (5)Have you ever felt that you drink too much? (denoted as C_1) (6)Have you ever failed to control the amount of your drinking beyond your original intention?(denoted as C_2) And (7)have your friends or family members ever tried to persuade you to drink less for your own good? (denoted as A_1)The construction of the last three new items was based on our previous experience of clinical interviews about drinking behavior and problems in Chinese communities. The Chinese culture emphasizes face, and people are often reluctant to admit their mistakes or their failure to do what they ought to do in front of others. In the event of alcohol drinking, this may imply a person deliberately drinks too much to cause problems, and to be criticized or feel guilty, as well as a failure to control the amount of drinking. Hence, the new items were carefully worded to avoid direct confrontation to respondents' drinking problems.

The meaning of item "Cut-down" in CAGE is similar to that of item C_1 as well as item C_2. Item C_1 emphasizes in another way about inquiring the opinion of "feeling drink to much". Those who feel ought to cut down their drinking may have experiences of failure to control. Item C_2 was constructed by pointing out their intrapsychic conflicts. In spite of similarity in content between item A_1 and item "Annoyance" of CAGE, the former item lays emphasis on the concerns from friends or family for their own good instead of bothering from others' criticism. Regarding items "Guilt" and "Eye-opener" of the CAGE, there were no further items added in this study. However, in our Chinese translation of item G, the word "Guilt" was substituted by a psycholinguistic equivalent "shameful" to diminish a potential link between Chinese word "guilt" and criminality[10].

Statistical analysis

The original CAGE were compared with various partially replaced Chinese versions, which contain A_1 for A and C_1 (or C_2) for C in statistics. The two-point scoring method (0,1) was used and therefore all the scores of the CAGE, C_1A_1GE, C_1AGE, C_2A_1GE and C_2AGE ranged from 0 to 4. The significance of the differences between tests was examined by the non-parametric McNemar test for two related samples. We then applied receiver operating characteristic curve (ROC) analysis to compare the overall performance among different versions of CAGE. A computer program, ROCFIT [16], was used to estimate the area under the ROC curve and its standard errors non-parametrically [17-18] for this purpose. A cut-off point that is the closest to the left-upper point of the ROC figure (sensitivity=1.0, false positive rate=0.0) was chosen as the optimal cut-off point. A $p < 0.05$ was considered significant.

RESULTS

There was no difference in sex and age distribution between the total outpatients (n = 1432) and the selected subjects (n = 198). Among the selected subjects, women had a lower education level and were more frequently widowed and unemployed than men (Table 1). Seventy-three of the selected subjects had consumed

alcohol. Among them, there was no sex difference in demographic distribution except a significant female excess of unemployment. Out of these 73 subjects, 70 (95.9%) completed both the screening questionnaire and clinical interview (95.2% for men and 100% for women).

Table 2 show the distribution of C-CAGE score under the diagnosis according to ICD-10-DCR in the past one year and lifetime. According to ICD-10-DCR, the number of subjects who met the criteria of harmful use of alcohol and alcohol dependence during the preceding year was 11 and 2, respectively (Table 2). If lifetime diagnosis was considered, the numbers were 19 and 3, respectively. The crude one-year and lifetime prevalence of alcoholism (including both harmful use of alcohol and alcohol dependence) were 6.7% and 11.3%, respectively.

Using the optimal cut-off point derived from the ROC curve, the validity indices of the five different versions of CAGE are listed in Table 3. Under the 1-year morbidity, as compared with C_2A_1GE (84.6%), CAGE had a lower (69.2%) sensitivity but not reaching statistical significance. The similar result was also noted by lifetime prevalence. In specificity, CAGE had a similar result with C_2A_1GE, but significantly in statistics better than C_1AGE and less superior than C_2AGE. Regarding the overall accuracy, CAGE was inferior to C_2AGE non-parametrically, but greater than C_1AGE. CAGE resembles C_2A_1GE in the overall accuracy by 1-year prevalence, but had an inferior tendency under lifetime morbidity. In brief, as compared with CAGE, C_2A_1GE inclined towards better in sensitivity, and C_2AGE was superior significantly in

Table 1　Demographic characteristics of total subjects selected from a primary care clinic (N=198) and a subset who ever drink (N=73).

	Total sample subjects		Subset of subjects who ever drink	
	men	women	men	women
Variables	N (%)	N (%)	N (%)	N (%)
Age(years)				
15-34	4(5)	7(6)	1(2)	1(10)
35-54	33(40)	36(31)	27(44)	4(36)
55	45(55)	73(63)	34(55)	6(55)
Marital status				
Single	8(10)	1(1)	4(7)	0(0)
Married	69(84)	87(75)	55(89)	9(9)
Others	5(6)	28(24)**	3(5)	2(18)
Education(years)				
0	25(31)	49(42)	18(29)	4(36)
1-6	30(37)	50(43)	24(39)	6(55)
7	27(33)	17(15)*	20(32)	1(9)
Job				
Without	30(37)	97(84)	20(32)	8(73)
With	52(63)	19(16)**	42(68)	3(27)*
Total	82(100)	116(100)	62(100)	11(100)

Sex difference(x^2 test): * p<0.05, ** p<0.01.

Table 2   Distribution of C-CAGE[a] score by the alcoholic status according to ICD-10-DCR[b] among the subjects who completed the C-CAGE and psychiatric interview.

C-CAGE score	Non-alcoholic	Harmful use	Dependence
	One-year prevalence		
0	12	0	0
1	7	0	0
2	14	2	1
3	15	1	0
4	2	4	1
5	5	0	0
6	2	4	0
7	0	0	0
Total	57	11	2
	Lifetime prevalence		
0	12	0	0
1	6	1	0
2	12	4	1
3	13	3	0
4	0	6	1
5	4	1	0
6	1	4	1
7	0	0	0
Total	48	19	3

[a] the experimental Chinese CAGE
[b] ICD-10, Diagnostic Criteria for Research (WHO, 1993) [15].

Table 3   Validity indices (values in %) of the five versions of the CAGE against alcohol use disorders according to ICD-10-DCR[a].

Instrument	Optimal Cut-point	Sensitivity	Specificity	Overall accuracy
	One-year prevalence			
CAGE	2	69.2	64.9	65.7
C_1A_1GE	2	76.9	68.4	70.0
C_1AGE	1	77.0	50.9**	54.3**
C_2A_1GE	2	84.6	63.2	67.1
C_2AGE	2	61.5	86.0**	81.4**
	Lifetime prevalence			
CAGE	2	54.5	64.6	61.4
C_1A_1GE	2	68.2	72.9	71.4
C_1AGE	1	68.2	52.1*	57.1**
C_2A_1GE	2	77.3	68.8	71.4
C_2AGE	2	45.5	87.5**	74.3**

[a] ICD-10, Diagnostic Criteria for Research (WHO, 1993) [15].
C_1, C_2: psycholinguistic equivalent items to "Cut-down" of CAGE; A_1: equivalent item to "Annoyance" of CAGE.
Significant differences (McNemar) concerning CAGE (binomial two-tailed probability): *p<0.05; **P<0.01.

specificity as well as the overall accuracy.

On the basis of ROC analysis using 1-year prevalence, the value of area under ROC curve was the lowest in C_1AGE and CAGE had a lower value as compared with C_2A_1GE (Table 4). There were no significant differences between the five versions of questionnaires. The figures for lifetime prevalence were similar, although the prevalence according to C_2A_1GE was higher than that according to CAGE at a borderline significance level (p=0.067).

DISCUSSION

Estimated prevalence of alcohol use disorders

Previous work among attenders seeking primary care in Western countries, varying in instruments used and sample characteristics, reported an estimated six month prevalence of 10.9% for DSM-III alcohol abuse [19], and 9.3% for at risk drinking screened by CAGE [8]. This study found a relatively lower prevalence of alcohol use disorder (6.7% in terms of one-year prevalence), which is in accordance with the lower prevalence of such morbidity reported in a previous survey using DIS in Chinese communities in Taiwan [2].

Validity of the CAGE and C-CAGE

The present study examined the cross cultural validity of the CAGE in a primary care setting in Taiwan. We have designed three psycholinguistic equivalents for the original items C and A of the CAGE based on our field experience in epidemiological studies of alcohol use disorders here. We then tested whether they can improve the performance of the CAGE in our sample. The results indicate a tendency of better validity of this modified screening tool when specific equivalent substitute the original one, but has not yet reached statistical significance, due to perhaps the limitation of sample size in this study.

In previous reports, sensitivity of CAGE ranged from 70 to 85% and specificity from 68 to 95% [8, 20-22]. Corresponding figures in this study fell at the lowest range of previous reports. In this regard, CAGE still did have cross-cultural validity in Taiwanese population and can be used for international comparison. Nevertheless, given the magnitude of validity indices, the CAGE has room for further improvement. Although our results were limited by small sample size, there were several clues under the assumption of psycholinguistic equivalence that might shed new light on the direction for improvement. C_2AGE expresses better validity than CAGE in specificity and overall accuracy, and these figures are in favor of a replacement of the C item of CAGE with item C_2. The area

Table 4　The value of the area under ROC curve and standard error of the five versions of the CAGE against alcohol use disorders according to ICD-10-DCR[a].

	CAGE	C_1A_1GE	C_1AGE	C_2A_1GE	C_2AGE
One-year prevalence					
Area under ROC curve	0.736	0.765	0.692	0.810[†]	0.767
Standard error	0.084	0.064	0.086	0.060	0.080
Lifetime prevalence					
Area under ROC curve	0.685	0.749	0.646	0.784[††]	0.706
Standard error	0.068	0.061	0.074	0.059	0.071

[a] ICD-10, Diagnostic Criteria for Research (WHO, 1993) [15].

Difference of area under ROC curve between C_2A_1GE and CAGE: [†]p=0.27; [††]p=0.067.

under ROC curve, i.e., the index of overall performance, of C_2A_1GE tends to be larger than CAGE and similar result was noted in sensitivity. The reasons that the differences do not reach a statistic significance may be due to the small number of cases with alcoholics. Therefore, C_2A_1GE still has the potential to be applied.

It is therefore implied that the validity of Chinese CAGE can be improved if psycholinguistically equivalent items C_2 and A_1 substitute the original C and A items. It is worthwhile to note that item C_2 and item A_1 put much emphasis on the questions related to family's concern about the health of respondents. Similar emphasis on family relationship has been observed in developing the Chinese Health Questionnaire (CHQ) for screening general psychiatric morbidity [10], which was modified from the General Health Questionnaire [23] in a similar way. The fact that C_1AGE was inferior to CAGE significantly in validity indices indicates that item C_1 is not a suitable equivalent item to improve the validity of CAGE.

Some researchers have pointed out that a general limitation of alcoholism screening instruments is that they elicit a lifetime prevalence of alcohol abuse problems and do not allow for a distinction between current and past abuse [24]. We have observed the similar limitation in this study. Nevertheless, our data showed that all five versions of CAGE had equivalent validity in screening for either current or lifetime alcoholism.

For prevention, it is important to detect alcoholics at an early stage. Subjects might have abnormal drinking patterns but still not yet meet diagnostic criteria for alcoholism. Although the performance of CAGE in detecting alcoholism in general is satisfactory, its performance in detecting hazardous patterns of alcohol drinking tend to be less impressive [7]. Similar decrement in the validity of either C_2A_1GE or C_2AGE in detecting risk drinking (eight or more drinks/day) also has been observed in this study. For 1-year morbidity, sensitivity and specificity of the C_2A_1GE at the optimal cut-off point were 60.0% and 50.4%, while the corresponding values of the C_2AGE were 40.0% and 78.5%, respectively. Recent studies that added a tolerance item (e.g., how many drinks does it take to make you feel high?) to the screening questionnaire have shown considerable promise in detecting heavy alcohol drinking in the primary care setting [25-26]. Whether this applies to our study population warrants further investigation.

ACKNOWLEDGMENTS

REFERENCES

1. Lin TY. A study of the incidence of mental disorder in Chinese and other cultures. Psychiatry 1953;**16**:313-36.
2. Helzer JE, Canino GJ, Yeh EK et al. Alcoholism - North America and Asia: A comparison of population surveys with the Diagnostic Interview Schedule. Arch Gen Psychiatry 1990;**47**:313-9.
3. Rydon P, Redman S, Sanson-Fisher RW et al. Detection of alcohol-related problems in general population. J Stud Alcohol 1992;**53**:197-202.
4. Bernadt MW, Mumford J, Taylor C et al. Comparison of questionnaire and laboratory tests in the detection of excessive drinking and alcoholism. Lancet 1982;**1**:325-8.
5. Mayfield D, Mcleod G, Hall P. The CAGE Questionnaire: validation of a new alcoholism screening instrument. Am J Psychiatry 1974;**131**: 1121-3.
6. Selzer ML. The Michigan Alcoholism

Screening Test: the quest for a new diagnostic instrument. Am J Psychiatry 1971;**127**:1653-8.

7. Maisto SA, Conners GJ, Allen JP. Contrasting self-report screens for alcohol problems: a review. Alcohol Clin Exp Res 1995;**19**:1510-6.

8. King M. At risk drinking among general practice attenders: validation of the CAGE questionnaire. Psychol Med 1986;**16**:213-7.

9. Cherpitel CJ. Analysis of cut points for screening instruments for alcohol problems in the emergency room. J Stud Alcohol 1995;**56**:695-700.

10. Cheng TA, Williams P. The design and development of a screening questionnaire (CHQ) for use in community studies of mental disorders in Taiwan. Psychol Med 1986;**16**:415-22.

11. Spak F, Hällström T. Screening for alcohol dependence and abuse in women: description, validation, and psychometric properties of a new screening instrument, SWAG, in a population study. Alcohol Clin Exp Res 1996;**20**:723-31.

12. Kristenson H, Trell E. Indicators of alcohol consumption: Comparisons between a questionnaire (Mm-MAST), interviews and serum gamma -glutamyltransferase (GGT) in a health survey of middle-aged males. Br J Addict 1982;**77**:297-304.

13. Saito S, Ilegami N. KAST (Kurihama Alcoholism Screening Test) and its application. Arukoru Kenkyuto Yakubutu Ison 1978;**13**:229-37.

14. Cheng TA, Chen WJ. Alcoholism among four aboriginal groups in Taiwan: high prevalences and their implications. Alcohol Clin Exp Res 1995;**19**:81-91.

15. World Health Organization. The ICD-10 Classification of Mental and Behavioral Disorders: Diagnostic Criteria for Research. Geneva: World Health Organization, 1993.

16. Metz CE, Wang PL, Kronaman HB. ROCFIT. Department of Radiology and the Franklin McLean Memorial Research Institute: University of Chicago, 1984.

17. Hanley JA, Mcneil BJ. The meaning and use of the area under a receiver operating characterisitic(ROC) curve. Radiology 1982;**143**:29-36.

18. Hanley JA, Mcneil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;**148**:839-43.

19. Coulchan JL, Zetter-Segal M, Block M et al. Recognition of alcoholism and substance abuse in primary care patients. Arch Intern Med 1987;**147**:349-52.

20. Bush B, Shaw S, Cleary P et al. Screening for alcohol abuse using the CAGE questionnaire. Am J Med 1987;**82**:231-5.

21. Buchsbaum DG, Buchanan RG, Welsh J et al. Screening for drinking disorders in the elderly using the CAGE questionnaire. J Am Geriatr Soc 1992;**40**:662-5.

22. Fertig JB, Allen JP, Cross GM. CAGE as a predictor of hazardous alcohol consumption in US army personnel. Alcohol Clin Exp Res 1993;**17**:1184-7.

23. Goldberg DP. Manual of the General Health Questionnaire. Windsor: NFER Publishing Co., 1978.

24. Umbricht-Schneiter A, Santora P, Moore RD. Alcohol abuse: comparison of two methods for assessing its prevalence and associated morbidity in hospitalized patients. Am J Med 1991;**91**:110-8.

25. Russell M, Martier SS, Sokol RJ et al. Screening for pregnancy risk-drinking. Alcohol Clin Exp Res 1994;**18**:1156-61.

26. Chan AWK, Pristach EA, Welte JW et al. Use of the TWEAK test in screening for alcoholism/heavy drinking in three populations. Alcohol Clin Exp Res 1993;**17**:1188-92.