

The development of a Chinese version of the tobacco use subscale of the behavioral risk factor surveillance system (BRFSS)

Yi-Hua Chen^a, Hung-Yi Chiou^a, Ping-Ling Chen^{b,*}

^a School of Public Health, College of Public Health, Taipei Medical University, No. 250, Wu-hsing Street, Taipei (110), Taiwan

^b School of Nursing, College of Nursing, Taipei Medical University, No. 250, Wu-hsing Street, Taipei (110), Taiwan

Available online 14 February 2008

Abstract

Objectives. To provide a culturally and linguistically specific survey instrument for tobacco consumption in Taiwan, we evaluated the reliability and validity of a Chinese translation of the 1993 US Tobacco Use Subscale of Behavioral Risk Factor Surveillance System (TU-BRFSS).

Methods. An integrative translation was followed by a pilot study of 100 randomly selected adults from throughout Taiwan. Telephone interviews took place in July, 2004. Validity was assessed by Content Validity Index (CVI) computed on the basis of expert review and the averaged scores of back-translation.

Results. Of 29 questions, 25 met the CVI criteria for retention in the instrument. In the back-translation assessment, 85% of the average scores taken from the expert evaluations were above 4 (scale of 1–5). Three of four percent agreements between the referent question and 4 other questions were 100%.

Conclusions. The Chinese version of the TU-BRFSS, with appropriate content, semantics, and conceptual equivalence, appears valid and reliable for future surveillance and research in Taiwan and other Chinese populations.

© 2008 Elsevier Inc. All rights reserved.

Keywords: BRFSS; Tobacco use subscale; Reliability; Validity; Taiwan; Smoking

Introduction

Tobacco use is a global epidemic and is considered to be the single most preventable cause of premature morbidity and mortality among men and women (Peto and Lopez, 2001). In Taiwan, tobacco control is considered one of the most important health policies by the administration.

Based upon U.S. tobacco control from 1965 to 2001, the percentage of adult smokers aged 18+ in the general population declined from 42.4% to 22.8%, both for males (52.0% to 25.5%) and for females (34.0% to 21.5%; Centers for Disease Control and Prevention, 2003). This indicates that it takes a sustained effort and a long period of time to reduce smoking rates. According to surveys conducted by the Taiwan Wine and Tobacco Monopoly Bureau between 1973 and 1996, the overall

adult smoking rate was 55%–63% for males and 2.3%–4.6% for females (Department of Health, 2006). In a 2002 national survey, approximately 50.0% of men and 5.8% of women reported that they smoke every day, which is higher than what has been observed in other countries (Department of Health, 2002).

World Health Organization (WHO) urged member states to immediately take action on nationwide strategies of tobacco control to prevent further morbidity and mortality caused by tobacco use. With the establishment of national tobacco surveillance systems toward tobacco use prevention and control, a standardized and reliable structure and capacity could be applied to assess and monitor tobacco profiles within a country. To promote tobacco control actions globally, it is essential for countries to be aware of similarities and differences among groups. Key to the assessment of international comparisons and the evaluation of a country's tobacco control efforts in light of the experiences of other countries is a set of standardized instruments and administration for surveillance data collection.

* Corresponding author. Fax: +886 2 2738 4831.

E-mail addresses: yichen@tmu.edu.tw (Y.-H. Chen), hychiou@tmu.edu.tw (H.-Y. Chiou), plchen@tmu.edu.tw (P.-L. Chen).

As a member in the global village, previous surveys in Taiwan on smoking prevalence and behaviors might not be sufficient for effective international comparison because of differences in the instruments and methods utilized in Taiwan and in other countries. Thus, a more appropriate assessment instrument should be applied in Taiwan to facilitate comparison with the global population.

The Behavioral Risk Factor Surveillance System (BRFSS, available for download at www.cdc.gov/brfss), which is the most widely used random digit dial telephone survey, was designed to measure prevalence and time trends for health-related behaviors in the U.S. population in 1993 (Centers for Disease Control and Prevention, 2006). Telephone surveys are an appropriate method for investigating the prevalence of health risk behaviors among populations. Thus, given the cost advantage and feasibility for long-term administration, the Tobacco Use Subscale of BRFSS (TU-BRFSS), a widely adopted instrument in surveillance system in literature, was selected for translation and use for tobacco surveillance and research in Taiwan.

The TU-BRFSS was designed to evaluate adult smoking prevalence, age of smoking initiation, behaviors for smoking cessation, exposure to and policy for environmental tobacco smoke. Reliability testing has been reported for the core BRFSS questions (Shea et al., 1991; Stein et al., 1993, 1995; Brownson et al., 1994). Although an original instrument might have been established with sound psychometric properties, it is critical to evaluate the reliability and validity of translated versions (Polit and Hungler, 1999).

Thus, the aims of this study were to translate and evaluate the Chinese version of the TU-BRFSS, to conduct a comprehensive pilot study, and to assess the validity and reliability of using this instrument to assess smoking behaviors of adults aged 18+ years in Taiwan.

Methods

In order to evaluate the content, semantics, and conceptual equivalence of the instruments in both the source and target languages, Guillemin et al. (1993) recommended the adoption of translation, back-translation, expert review and a pilot study as guidelines for cross-cultural adaptation of health-related measures. This study was administered based upon these guidelines.

Translation, back-translation, and expert review

Back-translation, the most commonly recommended and adopted manner of instrument translation (Brislin et al., 1973; Flaherty et al., 1988; Jones and Kay, 1992), was used to assess the semantics and content equivalence of the target language version. While the semantic equivalence assesses whether the meaning of each question remains the same compared with the source language version, the content equivalence ensures that the content of each item maintains consistent cultural relevance between the two versions. The validity and conceptual equivalence of the target language was assessed through expert review.

Thus, an integrative translation method was developed based upon the methods of translation, back-translation, and expert review as follows.

Step I: Two bilingual experts translated the TU-BRFSS instrument from English into Chinese, with Chinese Version I generated from the combined agreement on translation.

Step II: Five experts in tobacco-related fields assessed the validity, suitability, cultural equivalence, conceptual equivalence, and applicability of the

instrument for international comparisons of Chinese Version I. They were asked to rate each question on a four-point scale based on relevance, clarity, and importance.

Step III: Suggestions for questions and wording modifications that were collected from expert review were then taken into consideration while editing the BRFSS into Chinese Version II.

Step IV: Two bilingual English instructors of university-level English classes translated Chinese Version II back into English to ensure that the Chinese Version II retained the same meaning as the English version.

Step V: Two additional bilingual English instructors independently compared the original English instrument and the version translated back from Chinese Version II to ensure the equivalence and cultural relevance. Each question was thus scored from “1” meaning “very inappropriate” to “5” meaning “very appropriate.”

Step VI: The results from Step V were edited and further modified based upon expert evaluation. This Chinese version III of the TU-BRFSS was used in a pilot study to assess reliability.

Pilot study

Sample/location

A pilot study was conducted in July, 2004. The sample was selected from adults in Taiwan with at least one residential telephone in the household, as the household telephone coverage rate in Taiwan in 1999 was up to 99% (Ministry of the Interior, 1999). A sample size of at least 25 (preferably 50) is required if the purpose of a pilot study is to examine whether measurements are reliable and valid (Lancaster et al., 2004). Because our study was a nationwide investigation, 100 adults were predetermined for investigation. Phone numbers were proportionately selected and called by random digit sampling. While about 39% of calls were not answered, 33% were not eligible (e.g., non-residential household), and 10% refused to participate, 100 adults were questioned, with an overall average completion rate of 18% (17.2%, 18.7%, 20.8% and 15.8% in the Northern, Central, Southern, and Eastern areas of Taiwan, respectively).

Procedure

For each of the Northern, Central, Southern, and Eastern areas of Taiwan, 25 area codes were randomly selected. Then, the last four digits of the telephone number were randomly selected to ensure that all residents of Taiwan with a home phone had a chance of being selected, regardless of their number's presence in a phone book. Up to four calls during three different calling periods were made to contact a selected household before a replacement number was

Table 1

The content validity index for the Chinese Version I of the BRFSS, the Tobacco Use Subscale Instrument (TU-BRFSS)

Traits	No. of questions	Percentage (%)
CVI score for keeping		
1	22	75.9
0.8	3	10.3
0.6	4	13.8
CVI score for modification		
1	4	13.8
0.8	6	20.7
0.6	9	31.0
0.4	2	6.9
0.2	7	24.1
0	1	3.4
Keep the question?		
No	4	13.8
Yes	25	86.2
Unmodified ^a	(10)	(40)
Modified ^a	(15)	(60)

Data from nationwide Taiwanese respondents (2004).

^aThe numbers in parentheses represent number and percentages of retained questions only.

Table 2

The percentage distribution of the averaged scores reported by two experts who compared the two versions of the instrument (i.e., the source version and English version translated back from the Chinese version II) in the back-translation

Score average	No. of questions ^a	Percentage (%)
5	6	22.2
4.5	9	33.3
4	8	29.6
3.5	1	3.7
3	2	7.4
2.5	1	3.7
Total	27	100

Data from nationwide Taiwanese respondents (2004).

^aThis indicates “number of questions receiving the given score average.”

generated and dialed. Each interview was administered by trained Public Health graduate students and lasted approximately 5–10 min. All interviewees who completed the phone call provided full responses.

Statistical analyses

Data from the integrative translation method

The validity of the instrument was assessed by both the Content Validity Index (CVI) calculated from expert review and the averaged scores estimated from back-translation. The CVI is derived from a four-point ordinal rating scale with “1” meaning a very inappropriate element and “4” a very appropriate item (Lynn, 1986). The CVI score for retaining a question was defined as the proportion of questions that received a rating of 3 or 4 by the experts, while the CVI for modification was calculated as the proportion that received a rating of 4. Generally, questions that had a CVI over 0.8 remained, while those with CVI scores for modification lower than 0.6 were further edited or deleted based upon the experts’ opinions.

Data from the pilot study

Because the instrument was not designed with a set of questions with consistent ordinal or continuous coding, Cronbach’s alpha was not applicable. Alternative methods, such as percent agreement, were adopted for reliability examination in the pilot study. SAS 8 (SAS Institute, Cary, NC) statistical software was used for data analysis.

Results

The CVI assessment of the Chinese TU-BRFSS version I is presented in Table 1. Of 29 questions, those with CVI for

keeping over 0.8 were retained, and the rest were discarded, resulting in a 25-question scale. By modifying 15 questions using experts’ opinions, the edited Version II instrument was then used for back-translation examination.

In Table 2, 85% of average scores taken from the expert evaluations comparing the two versions of the instrument were above 4 (scale of 1–5), indicating acceptable content and cultural equivalence between the original and translated versions. Questions, especially those with average scores below 4, were further modified using experts’ opinions. The revised Chinese version III was used in the pilot study that examined reliability.

In the pilot study, the gender ratio within the community sample was approximately equal. Over 60% of respondents were married, and over 75% had at least a high school diploma. Sixty-seven percent currently held a job, and more than half earned a family income of 1 million NT (i.e., equivalent to \$30,000 US) or less per year (not shown in table). Table 3 compares the sociodemographic distribution of national data, drawn from governmental statistics reported by the Department of Statistics, Ministry of the Interior, the Executive Yuan, and the data from our study. No statistically significant difference was found regarding the distribution of gender, age, employment status, and marital status (all p -values > 0.05) between groups.

Further data analysis assessed the logical consistency of an answer to one question with the response to a comparable question. If interviewees’ answers were reliable, their responses to the referent question and the other four questions in Table 4 should have been highly consistent. Results indicated that the percent agreement between the referent question and 3 out of the 4 questions was 100%. This high consistency demonstrated good reliability.

In summary, based upon the reliability and validity assessment, the TU-BRFSS instrument was further edited and modified as follows. Questions of “last smoked regularly” and “has sought medical care” were deemed unsuitable and were removed. For cultural applicability, four more questions were added; these include questions on “how often have you quit smoking?” “in the past 7 days, when you were at home, how many days has someone else smoked in front of you?” “in the

Table 3

The sociodemographic distribution of national data and our pilot study data for the BRFSS instrument, the Tobacco Use Subscale (TU-BRFSS)

		National data	Sample	Chi-square ^a
		No. (%)		
Gender	Male	11,541,585 (50.9%)	51 (51%)	0.97
	female	11,147,537 (49.1%)	49 (49%)	
Age	18–29	3,886,690 (23.3%)	27 (27.8%)	0.52
	30–64	10,672,210 (63.9%)	57 (58.8%)	
	65+	2,150,475 (12.9%)	13 (13.4%)	
Employment status ^b	Employed	9,786,000 (95.6%)	67 (95.7%)	0.95
	Unemployed	454,000 (4.4%)	3 (4.3%)	
Marital status	Single	4,429,153 (39.3%)	28 (28%)	0.08
	Married	6,022,846 (53.4%)	63 (63%)	
	Divorced/widowed	819,500 (7.3%)	9 (9%)	

Data are from nationwide Taiwanese respondents (2004).

^a Chi-square tests were carried out for the comparison of distributions between national data and sample data.

^b Employment status excluded students, materfamilias, retired people, or those who are unable to work.

Table 4
Logical analysis of questions on current smoking status and other smoking-related items in the Chinese version III instrument

Items	Percent agreement (%)
Reference question: Do you now smoke cigarettes every day, some days, or not at all? To be compared with the following four questions:	
Question 1: About how long has it been since you last smoked cigarettes regularly?	100
Question 2: During the past 30 days (1 month), on how many days did you smoke cigarettes?	90
Question 3: During the past 30 days (1 month), how many cigarettes did you usually smoke on the days you smoked?	100
Question 4: Which statement best describes the rules about smoking inside your home?	100

Data are from nationwide Taiwanese respondents (2004).

past 7 days, when you were working at your job, how many days has someone else smoked in front of you?" and "are you a native resident?" Question order and wording were modified to better fit the logical thought flow of the interview. The Chinese version was thus developed to assess adult smoking behaviors in Taiwan.

Discussion

The integrative translation method in this study presented a systematic and valid approach to translate an instrument. Empirical evidence from our study suggests that the Chinese version of the TU-BRFSS is valid and reliable for international comparison and native investigation. It possessed the appropriate semantics, content (assessed by back-translation), and conceptual (assessed by expert review) equivalence of the original language.

Cigarette smoking is a major health risk behavior worldwide. In an analysis of national vital statistics in developed countries from 1950 to 2000, Peto et al. (1994) estimated that more than 10 million people will die from smoking by 2025. In Taiwan, 13.9% of male mortality and 3.3% of female mortality could be attributed to cigarette smoking (Liaw and Chen, 1998). Reducing tobacco-related risk behaviors is a priority among the national and international health objectives.

Reliability and validity are essential characteristics to ensure that instruments are valid for smoking surveillance and monitoring. Previous studies have demonstrated that the BRFSS questionnaire is a valid tool to survey and conduct research in the US (Shea et al., 1991; Stein et al., 1993, 1995; Brownson et al., 1994). For example, Stein et al. (1993) demonstrated that reliability coefficients were over 0.7 for behavioral risk factors including smoking. In another study by Shea et al. (1991), Pearson or kappa correlations for questions concerning demographics and behavioral risk factors were more than 0.6 ($p < 0.001$), except for questions about diet. Test-retest reliability demonstrated acceptable to high question reliability on an individual level. In our study, the logical consistency of an answer to one question with a response to another comparable item was assessed. There was 100% agreement between smoking status

and smoking habits, smoking quantity per day, and smoking rules at home. This high consistency demonstrated that the Chinese version of the TU-BRFSS instrument is appropriate and reliable.

Content validity measures the comprehensiveness and representativeness of the content of a scale and could contribute to support the construct validity. The CVI is the most widely adopted approach to quantify content validity. Based upon recommendations by Lynn (1986), items in this study that did not achieve the required minimum agreement of the experts were eliminated or further edited. The final Chinese version of the TU-BRFSS thus demonstrated good content validity. Further, when comparing the two versions of the instrument (i.e., the source version and English version translated back from the Chinese version II), 85.1% of the average scores taken from the expert evaluations were above 4 and items were modified based upon experts' review. Thus, the Chinese version of the TU-BRFSS displayed good content and cultural equivalence with the source and target versions.

Study limitations and strengths

Results of the study directed the need for a culturally and linguistically specific survey instrument to address tobacco control issues in Taiwan. A procedure recommended by Guillemin et al. (1993) was adopted for an appropriate reliability and validity assessment of TU-BRFSS. A nationwide sample of 100 community adults was drawn for the administration of a pilot study. However, there were some limitations to our study. First, internal consistency reliability as assessed by Cronbach's alpha was inapplicable because the instrument was not designed with a series of questions that had a consistent ordinal or continuous scale of response choices. In this study, logical consistency was measured instead by percent agreement. In addition, test-retest reliability was not possible because of the anonymous nature of the telephone interview in the pilot study. Second, the CVI procedure to assess content validity might have been limited by the possibility of chance inflation (agreement) (Waltz and Bausell, 1981). However, five experts in our study were capable of providing a sufficient level of control for chance agreement (Lynn, 1986). Finally, a relatively small sample size and rather low response rate might restrict inferences from broader generalization. However, our pilot study recruited an appropriate sample size as a pilot and was representative of the national data reported by government statistics regarding the distributions of gender, age, employment, and marital status.

Future studies are needed to evaluate the Chinese version of the TU-BRFSS in a more demographically and geographically diverse population. With the recruitment of more participants, broader generalizations may be possible.

Conclusions

In summary, these preliminary findings support applying the Chinese version of the TU-BRFSS for surveillance and research to measure and monitor the prevalence of smoking-related issues in Taiwan and other Chinese populations as well. Based

upon the reliability and validity of the instrument, future work should be performed to provide the nation with a more accurate estimate of smoking prevalence and evaluation of effects on tobacco control intervention.

Acknowledgments

This research was supported by a grant from the Bureau of Health Promotion, Department of Health, Executive Yuan, Taiwan.

References

- Brislin, R.W., Lonner, W.J., Thorndike, R.M., 1973. *Cross-cultural Research Methods*. John Wiley & Sons, New York.
- Brownson, R.C., Jackson-Thompson, J., Wilkerson, J.C., Kiani, F., 1994. Reliability of information on chronic disease risk factors collected in the Missouri Behavioral Risk Factor Surveillance System. *Epidemiology* 5, 545–549.
- Centers for Disease Control and Prevention, 2003. Cigarette smoking among adults — United States, 2001 [erratum appears in *MMWR Morb Mortal Wkly Rep.* 2003 Oct 24;52(42):1025]. *MMWR Morb. Mort. Wkly. Rep.* 52, 953–956.
- Centers for Disease Control and Prevention, 2006. Behavioral Risk Factor Surveillance Factor Operational and User's Guide. Version 3.0. In: National Center for Chronic Disease Prevention and Health Promotion, (Ed.). Division of Adult and Community Health, Behavioral Surveillance Branch, Centers for Disease Control and Prevention, USA.
- Department of Health, 2002. Reports from knowledge, attitude, and behavior about health promotion in Taiwan. Bureau of Health Promotion, Department of Health, Executive Yuan, Taiwan.
- Department of Health, 2006. Taiwan Tobacco Control Annual Report, 2006. Bureau of Health Promotion, Department of Health, Executive Yuan, Taiwan.
- Flaherty, J.A., Gaviria, F.M., Pathak, D., et al., 1988. Developing instruments for cross-cultural psychiatric research. *J. Nerv. Ment. Dis.* 176, 257–263.
- Guillemin, F., Bombardier, C., Beaton, D., 1993. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines.[see comment]. *J. Clin. Epidemiol.* 46, 1417–1432.
- Jones, E.G., Kay, M., 1992. Instrumentation in cross-cultural research. *Nurs. Res.* 41, 186–188.
- Lancaster, G.A., Dodd, S., Williamson, P.R., 2004. Design and analysis of pilot studies: recommendations for good practice. *J. Eval. Clin. Pract.* 10 (2), 307–312.
- Liaw, K.M., Chen, C.J., 1998. Mortality attributable to cigarette smoking in Taiwan: a 12-year follow-up study. *Tob. Control* 7, 141–148.
- Lynn, M.R., 1986. Determination and quantification of content validity. *Nurs. Res.* 35, 382–385.
- Ministry of the Interior, 1999. Investigation on the Telecommunication Industry. The Ministry of the Interior, the Executive Yuan, Taiwan.
- Peto, R., Lopez, A.D., 2001. Future worldwide health effects of current smoking patterns. In: Koop, C.D., Pearson, C., Schwarz, M.R. (Eds.), *Critical Issues In Global Health*. Jossey-Bass, New York, NY.
- Peto, R., Lopez, A.D., Boreham, J., Thun, M., Health, C., 1994. Mortality from smoking in developed countries, 1950–2000: indirect estimates from national statistics. Oxford University Press, Oxford.
- Polit, D.E., Hungler, B.P., 1999. *Nursing research: Principles and methods*, 6th ed. Lippincott, Philadelphia.
- Shea, S., Stein, A.D., Lantigua, R., Basch, C.E., 1991. Reliability of the behavioral risk factor survey in a triethnic population. *Am. J. Epidemiol.* 133, 489–500.
- Stein, A.D., Courval, J.M., Lederman, R.I., Shea, S., 1995. Reproducibility of responses to telephone interviews: demographic predictors of discordance in risk factor status. *Am. J. Epidemiol.* 141, 1097–1105.
- Stein, A.D., Lederman, R.I., Shea, S., 1993. The Behavioral Risk Factor Surveillance System questionnaire: its reliability in a statewide sample [see comment] *Am. J. Public Health* 83, 1768–1772.
- Waltz, C.W., Bausell, R.B., 1981. *Nursing research: Design, statistics and computer analysis*. F. A. Davis, Philadelphia.