

# On the application of structural equation modeling for the construction of a health index

Ferra Yanuar · Kamarulzaman Ibrahim ·  
Abdul Aziz Jemain

Received: 22 October 2009 / Accepted: 3 March 2010 / Published online: 6 April 2010  
© The Japanese Society for Hygiene 2010

## Abstract

**Objective** The health of an individual is influenced by many factors. These could include factors that are related to the economy and the environment, as well as social and biological factors. Many studies have been carried out to study the effect of these factors on health, in terms of the individual factors or combined factors. The main purpose of this study was to demonstrate the value of structural equation modeling for the construction of an index to describe the health status of an individual.

**Methods** Structural equation modeling was applied to data obtained from 5035 respondents in a survey conducted in the district of Hulu Langat, Malaysia, in the year 2001 by the Department of Community Health, Medical Faculty, Universiti Kebangsaan Malaysia, Malaysia. The survey involved the gathering of information on the respondents' demography, lifestyles, mental health condition, and biomarkers.

**Results** Socio-demography and mental health condition were found to have a direct effect on the health index. However, lifestyle had an indirect effect on the health index, as mediated by the mental health. Based on the indicator of model fit, the proposed model fits the data well.

**Conclusions** Structural equation modeling was found to be pertinent to be used for analyzing the health index of a particular individual.

**Keywords** Health index · Structural equation modeling · Socio-demography · Lifestyle · Mental health

## Introduction

A health index is often used as an indicator of the level of health of a community or a country. These indexes are influenced by the health level of individuals within the community. Besides being often used as an indicator of the state of welfare of a country, the health index may also indicate the level of productivity and economic growth of a particular country. A high proportion of healthy people in the community could contribute to a greater average health level as well as the prosperity of the country.

According to the Malaysian Ministry of Health (2006–2012), there are many key indicators of health, including average life expectancy at birth, crude death rates, and the prevalence of health risk factors such as tobacco smoking, alcohol consumption, overweight, physical inactivity, and others. Because the health index may reflect the level of economic development, factors that are related to the health of an individual need to be identified. Health authorities should take into account the factors that influence the level of health among the public when deciding on plans of action for improving the health index of an individual. Having identified the factors, such as the risk factors for a particular disease, the health authorities could embark on a specific program which aims at reducing the risk due to those factors.

---

F. Yanuar (✉)  
Department of Mathematics, Faculty of Mathematics  
and Natural Sciences, Universitas Andalas, Kampus Limau  
Manis, Padang 25163, Indonesia  
e-mail: ferrayanuar@yahoo.co.id

K. Ibrahim · A. A. Jemain  
School of Mathematical Sciences, Faculty of Science  
and Technology, Universiti Kebangsaan Malaysia,  
43600 Bangi, Selangor, Malaysia

Because the determinants of health may vary between communities, and with respect to time, determination of the health index is not a clearcut process. In addition, some of the determinants such as lifestyle, socio-demography, and mental health are not directly observable or are latent. This makes the process of determining the health index complicated; thus, a suitable technique has to be used when modeling the health index. Various studies have been carried out to identify the factors that are related to health. Torrance et al. [1], for example, classified the health status of an individual using a multi-attribute utility function based on seven attributes: sensory, mobility, emotion, cognitive, self care, pain, and fertility. Ramirez [2] found that income, education level, ill health, socioeconomic factors, and the institutional health system were the determinants of health status, based on an order probit method. Meanwhile, Dhargupta [3] developed a cognitive index based on the information from respondents on mental health, childcare, pregnancy management, family planning, immunization care, nutritional requirement, general treatment of diseases, and health education.

Some studies have demonstrated that a structural equation modeling approach is particularly suitable for modeling the interrelationship between observable and unobservable factors which describe the health status. Boniface and Tefft [4] used the structural equation modeling approach to identify the relationship between an unobservable factor, i.e., lifestyle, and observed variables, such as cigarette smoking, physical activity, and alcohol intake. Chern et al. [5] considered a structural equation model to model the relationship between the expenditures on health and the indicators of health. Stafford et al. [6] applied a structural equation approach to model potential neighborhood determinants of health, such as socio-relational characteristics, the built environment, and neighborhood amenities.

The relationships in structural equation modeling are usually formulated by linear regression equations, and can be graphically expressed by path diagrams. The distinct advantages of this technique compared to other statistical methods are the ability to test construct level hypotheses and provide many interpretations and much information based on the analysis. An overview of structural equation modeling is presented in the next section. In this article, our main purpose is to demonstrate the value of structural equation modeling for the construction of an index to describe the health status of an individual living in the district of Hulu Langat, Malaysia. The relationship of the indicator variables, obtained based on a questionnaire survey, with their respective latent variables, i.e., socio-demographic characteristics, lifestyle, and mental health was determined and we allowed for the interrelationship among these latent variables and their relationship to the health index.

## Materials and methods

### Data

The data set used for the analysis was based on the Health Risk Assessment survey conducted in Hulu Langat in the year 2001 by the Department of Community Health, Medical Faculty, Universiti Kebangsaan Malaysia (UKM) [7]. The main purpose of the survey was to maintain health through promotional and proper disease-prevention activities supported by information technology. Based on the information obtained from the Health Risk Assessment (HRA) of Hulu Langat District [7], about 5.7% of the people living in Hulu Langat had chronic diseases; the main diseases identified were gastritis (9.1%), diabetes (4.3%), and asthma (4.2%). Considering the high prevalence rate of these diseases, Hospital Universiti Kebangsaan Malaysia (HUKM), as the center for the coordination of activities involving healthcare services in the area of Hulu Langat, is continuously monitoring the health status of people in Hulu Langat.

More than 7000 respondents were involved in this survey. However, this analysis is based on the 5035 respondents who provided complete information in the survey. A questionnaire form was used to obtain information from the respondents, such as socio-demographic status, lifestyle, and mental health condition. In addition, medical screening for measuring body height and weight, blood pressure, cholesterol level, high-density lipoprotein (HDL), and blood glucose was done either during a house visit or during consultation at the clinic at HUKM. Respondents were also asked about the number of health problems experienced and their perceptions about their health condition at the time of the interview. Complete information regarding the data can be obtained from the HRA of Hulu Langat District [7].

### Factors related to health

There are many factors which are related to health; some are observable and some are not. Biomarkers, such as blood pressure level, cholesterol level, etc., can be considered as examples of observable measures which can be used to describe the level of health of an individual. The factors that are not directly observed, such as lifestyle, socio-demography, and mental health condition can be measured through indicator variables. Following is a brief explanation of each unobservable variable considered in this study and its respective indicators.

Socio-demographic status is an important determinant of a health index. Several studies [8–11] have found that people who are poorer, less well educated, or have a lower-status job tend to have a lower health index. The indicators

of interest considered in this study for assessing the socio-demographic factor were education level, employment status, gender, and age. Respondents were asked about their education level and the responses were indicated by 1, 2, 3, and 4 as “never attended school”, “elementary school”, “high school”, and “college/university”, respectively. With respect to employment status, the responses obtained were coded as 1 and 2 for “unemployed” and “employed”, respectively. For coding the gender of the respondent, 1 denoted “male” and 2 denoted “female”. The ages of the respondents were recorded and treated as a continuous variable.

The important role of lifestyle as a contributory factor in influencing the level of health of an individual is well known. An unhealthy lifestyle, such as being a smoker or rarely having physical exercise, could have a negative influence on the health condition of an individual [8–12]. The lifestyle indicators considered in this study, based on the list of health-related behaviors as suggested by Nakayama et al. [12] (a list that has also been used by Boardman [8]) were physical exercise, working hours, fiber and/or vegetable consumption, and number of sleeping hours. With respect to these indicators, respondents were asked about physical exercise using the question “how many times a week on average do you have physical exercise?” The responses for this question ranged from 1, indicating “none” to 4, referring to “more than 3 times in a week”. Working and sleeping hours were considered as continuous variables, as obtained from the question asking about the number of hours spent on the average by the respondents for working and sleeping during a particular day. Meanwhile, to know whether or not the respondents consumed fruits and/or vegetables, the respondents were asked the question “do you consume food which contains a high level of fiber such as vegetable or fruit every day?” The responses were “no” or “yes”, denoted as 1 and 2, respectively.

Apart from lifestyle, mental health is often recognized as one of the major health determinants. A study by Nakayama et al. [12] showed that health was influenced by many factors, including mental health. There was a significant correlation between physical health and mental health. Hence, it is reasonable to assume that mental health is a contributory factor for determining the health index. Nakayama et al. [12] and Boardman [8] have suggested the use of stress level and the number of experiences of serious problems as an indicator for mental health. They found that people who have a high stress level, as well as those who have experienced health problems, possess a low level of mental health. The question used in the present study to know the stress level was “what is your stress level now?” Responses to this question were 1 for “low”, 2 for “middle”, and 3 for “high”. To know whether respondents had

experienced serious problems or not, they were asked the question “did you have any serious problems or difficulties in the past year?” The responses were “no problems”, “one problem”, and “2 or more problems”, denoted by 1, 2, and 3, respectively.

#### Measures of health index

In this study, as based on the literature, it was assumed that socio-demography, lifestyle, and mental health were latent factors that were related to the health index of an individual. The health index could also be measured directly based on certain indicators such as body mass index (BMI), blood pressure, cholesterol level, blood glucose level, the number of common health problems experienced by the respondent, and the subjective judgment of the respondents on their health condition. The BMI of an individual is the value for body weight divided by the square of his/her height. The measurement and classifications of BMI as normal, underweight, overweight, and obese were based on the categories suggested by the Centers for Disease Control and Prevention (CDC 2009). The classifications that were used for blood pressure level were normal, prehypertension, and high blood pressure, based on the American Heart Association (2008). The classifications for total cholesterol level were decided using the guidelines provided by the National Institutes of Health (2001). The classifications of blood glucose level were based on the classification from the American Diabetes Association (2007). In the HRA survey conducted in 2006, the respondents were asked whether or not they had ever experienced any one or more of the 14 common health problems that were listed in the questionnaire. In the present study, the respondents who had fewer than two health problems were classified as healthy, those who had two to four health problems were considered less healthy, and those with more than four health problems were considered as unhealthy. The general health condition of the respondent was found based on the answer to the question “What would you say about your health in general?” The responses to this question were: (1) “unhealthy”, (2) “less healthy”, and (3) “healthy”.

#### Analytical technique

Structural equation modeling was applied to the data obtained from the survey. Before the structural equation modeling was applied, a path model was constructed based on the hypothesized relationships as described above. The hypothesized relationships involved the interrelationships between the latent variables, which included socio-demography, lifestyle, mental health, and the health index, as well as the relationships between the indicators and each individual latent variable.

The process of modeling in structural equation modeling could be thought of as a four-stage process: model specification, model estimation, model evaluation, and model modification. To achieve the first process, structural equation modeling involves two major components; namely, the measurement equation and the structural equation. A mathematical expression of the measurement equation model is specified as follows:

$$x = \Lambda_x \xi + \delta \tag{1}$$

$$y = \Lambda_y \eta + \varepsilon \tag{2}$$

where  $x$  is an  $m \times 1$  vector of indicators describing the  $q_1 \times 1$  exogenous latent vector  $\xi$  and  $y$  is an  $n \times 1$  vector of indicators that proxy the  $q_2 \times 1$  endogenous latent vector  $\eta$ ,  $\Lambda_x$  and  $\Lambda_y$  are  $m \times q_1$  and  $n \times q_2$  matrices of the loading coefficients as obtained from the regressions of  $x$  on  $\xi$  and  $y$  on  $\eta$ , respectively, and  $\delta$  and  $\varepsilon$  are  $m \times 1$  and  $n \times 1$  random vectors of the measurement errors, respectively. The random vectors  $\delta$  and  $\varepsilon$  are assumed to be independent and also uncorrelated with the latent variables  $\xi$  and  $\eta$ , respectively. Given the observed data for describing the vectors  $x$  and  $y$ , the measurement equations appropriately group together the correlated indicator variables to form the latent variables in  $\xi$  and  $\eta$ . This is done by assigning fixed parameters and defining unknown parameters in  $\Lambda_x$  and  $\Lambda_y$ .

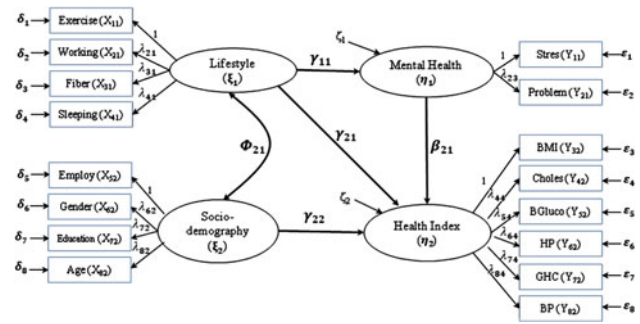
The interrelationship among the latent factors is explained through a structural equation model. It is expressed in mathematical formula as given by:

$$\eta = B\eta + \Gamma\xi + \zeta \tag{3}$$

where  $\eta$  is  $q_2 \times 1$  endogenous random vector of latent variables,  $\xi$  is  $q_1 \times 1$  exogenous random vector of latent variable,  $B$  is  $q_2 \times q_2$  matrix of coefficients relating the endogenous latent variables to one another in the structural relationship,  $\Gamma$  is  $q_2 \times q_1$  matrix of coefficients relating the exogenous latent variables to the endogenous latent variables, and  $\zeta$  is  $q_2 \times 1$  random vector of error measurements in the equations. It is assumed that  $B$  has zeros in the diagonal, and  $(I - B)$  is required to be nonsingular,  $\xi$  is uncorrelated with  $\zeta$ .

The hypothesized model with structural and measurement components for describing the health index is illustrated in Fig. 1. It is also reasonable to hypothesize that the socio-demographic status of respondents may be correlated to lifestyle [10], as indicated in Fig. 1 by the double-headed arrow connecting the two latent variables.

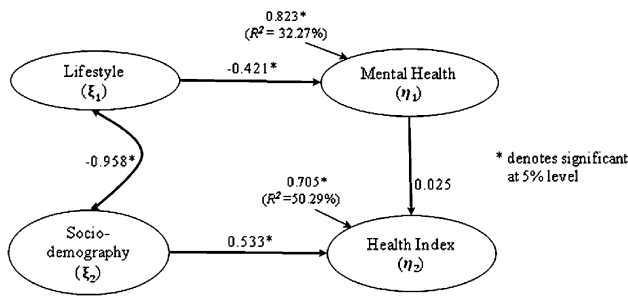
After the model specification component is performed, the population parameters are then estimated. The goal of estimation is to minimize the difference between the hypothesized matrix which is a function of parameter  $\theta$ , a vector that includes all the unknown parameters, denoted as



**Fig. 1** A diagrammatic illustration of the health index using structural equation modeling. *BMI* body mass index, *Choles* cholesterol, *BGluc* blood glucose, *HP* health problem, *GHC* general health condition, *BP* blood pressure

$\Sigma(\theta)$  and the sample covariance matrix, denoted as  $S$ . To measure the closeness between these two variance covariance matrices  $S$  and  $\Sigma(\theta)$ , a suitable fitting function, denoted by  $F(S, \Sigma(\theta))$  is utilized. In this study, an approach called robust weighted least square (RWLS) [13] is applied to handle non-normal and/or non-continuous data. This estimation method is not only preferable for data which are non-normal but also for data with small and large sample sizes [13]. The RWLS estimates parameter, standard errors,  $\chi^2$ , and fit indices based on diagonal elements of the weight matrix derived from the asymptotic variances of the thresholds and latent correlation estimates. One of the software packages available for this estimation is Mplus [14].

The next process in structural equation modeling is model evaluation of sample parameters. The criteria for estimation of fit include examination of the solution, measure of overall fit, and detailed assessment of fit. Firstly, parameter estimates with the right sign and size, standard errors within reasonable ranges, correlations of parameter estimates, and squared multiple correlations are commonly used to check for the appropriateness of each variable. Then, the overall model fit is evaluated to see how well the specified model fits the data. A complete discussion of model fit is outside the scope of this article. The result of the indicators of model fit that are presented by Mplus, such as the root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker Lewis Index (TLI) will be used to indicate whether or not the model fits the data. RMSEA measures the degree of model adequacy based on population discrepancy in relation to degrees of freedom; a value less than 0.08 is preferred (acceptable). CFI and TLI are examples of an incremental fit index. These types of index compare the improvement of the fit of the proposed model over a more restricted model, called an independence or null model, which specifies no relationships among variables. CFI and TLI ranges from 0 to 1.0, with values closer to 1.0 (or more



**Fig. 2** The fitted model, which includes the standardized parameter estimates of the latent variables and measurement errors

than 0.80) indicating a better fit [15]. Mplus also presents the  $\chi^2$  and weighted root mean square residual (WRMR). But neither of these indicators is considered to determine the model fit in this study because the  $\chi^2$  test is sensitive to the sample size, and WRMR is an experimental fit measure and does not always behave well.

The last step in structural equation modeling is model modification, which is done for at least two reasons: to test the hypothesis (in theoretical work) and to improve the fit (especially in exploratory work). There are three basic methods used in modifying a structural equation modeling model, i.e.,  $\chi^2$  difference, Lagrange multiplier (ML), and Wald tests. Not all of these three methods are discussed in this article; only the  $\chi^2$  difference test is applied because it is the most popular of the three [16].

**Results**

The data gathered from the survey were fitted to the proposed model structure and the results obtained are provided in Fig. 2 and Table 1.

The findings show that the proposed model fitted reasonably well to the data (RMSEA = 0.064, CFI = 0.829 and TLI = 0.820). There was a high correlation of  $-0.958$  between socio-demography and lifestyle, as represented by the line with the two-headed arrows in Fig. 2. These findings mean that people of a higher socio-demographic status tended to practice less healthy lifestyle behaviors. For example, people who had a higher education level were found to practice physical exercise less often as opposed to those at a lower level, and/or they rarely consumed food with fiber as compared to those in the lower education level. The regression coefficient of lifestyle on mental health was  $-0.421$  and was found to be significant. This indicates that people with a healthy lifestyle tend to have a good condition of mental health. Figure 2 also tells us that the amount of explained variance ( $R^2$ ) in this relationship is 32.27%. Researchers determine the amount of explained variation ( $R^2$ ) by squaring the standardized error associated

**Table 1** Parameter estimates, standard error estimates, and  $R^2$  for each measurement equation

Latent variable	Indicator variable	Estimate	Standard error	$R^2$
Lifestyle ( $\xi_1$ )	Exercise ( $X_{11}$ )	0.268	0.017*	7.2
	Working ( $X_{21}$ )	-0.728	0.020*	52.9
	Fiber ( $X_{31}$ )	0.013	0.016	0.01
	Sleeping ( $X_{41}$ )	0.248	0.041*	6.1
Socio-demography ( $\xi_2$ )	Employment ( $X_{52}$ )	0.922	0.012*	74.9
	Gender ( $X_{62}$ )	-0.585	0.015*	34.3
	Education ( $X_{72}$ )	0.441	0.015*	19.4
Mental health ( $\eta_1$ )	Age ( $X_{82}$ )	-0.341	0.013*	11.6
	Stress ( $Y_{13}$ )	0.593	0.085*	35.2
	Problem ( $Y_{23}$ )	0.236	0.040*	5.6
Health index ( $\eta_2$ )	BMI ( $Y_{34}$ )	0.270	0.031*	7.3
	Choles ( $Y_{44}$ )	0.282	0.032*	7.9
	BGlucose ( $Y_{54}$ )	0.153	0.028*	2.3
	HP ( $Y_{64}$ )	0.101	0.044*	1.0
	GHC ( $Y_{74}$ )	0.180	0.064*	3.2
	BP ( $Y_{84}$ )	0.998	0.173*	77.1

BMI body mass index, Choles cholesterol, BGlucose blood glucose, HP health problem, GHC general health condition, BP blood pressure

\* Significant at 5% level

with latent variables and subtracting the value obtained from 1 [17]. In this relationship,  $R^2 = 1 - \zeta^2 = 1 - 0.823^2 = 0.323$ . After allowing for the modification indices, this study found that lifestyle had an indirect effect on the health index, as mediated by the mental health. Therefore,  $\gamma_{21}$  in Fig. 1 should not be estimated for the final model. Mental health was associated with the health index, although the association was not significant ( $\beta = 0.025$ ,  $p = 0.497$ ). Figure 2 also shows that socio-demographic status had a significant relationship to the health index. One can conclude here that socio-demographic status correlates positively with health condition, which implies that people of good socio-demographic status tend to experience a better health condition. The estimated structural equation model that addressed the relationship of the health index with socio-demography and mental health was:

$$\eta_2 = 0.533\xi_1 + 0.025\eta_1.$$

Hence, socio-demographic status had a greater effect on the health index than mental health. Altogether, socio-demography and mental health explained 50.29% of the variation of the health index.

The value of the standardized coefficient of factor loading and the associated standard error for each indicator variable are presented in Table 1. Almost all indicator variables that we hypothesized as predictors were significantly related to their respective latent factors, except for

fiber (indicator of lifestyle). In this case, if fiber is excluded from the model, the model will not converge. So, we should not exclude this indicator from the model. The last column in Table 1, i.e., the  $R^2$  value, indicates the proportion of variation in the latent factor as explained by each particular indicator. It was found that working hours described lifestyle better than other indicators. The employment status explained the greatest proportion of variation in socio-demographic status as opposed to other indicators. However, both indicators for mental health managed to explain only a relatively small proportion of variation in this latent factor. The blood pressure (BP) level described the health index best when compared to other indicators.

## Discussion

The main purpose of the present study was to apply a structural equation modeling technique in modeling the health index of the Hulu Langat community. Structural equation modeling has the ability to describe a set of relationships between independent variables and dependent variables, where some of these variables are latent or unobserved. This technique is the only analysis that could be used to do complete and simultaneous tests of such relationships [16]. The analysis in the present study was implemented under Mplus version 5.2, a flexible tool which allows one to examine the relationship involving continuous, ordered categorical, and dichotomous variables.

Although there have been many previous studies on the modeling of a health index, there are a limited number of studies on the application of structural equation modeling for a health index, particularly with allowances for information on respondents' socio-demography, lifestyle, mental health, and biomarkers. The indicators for latent factors used in the present study were as follows. The socio-demographic indicators were employment status, gender, education level, and age. Lifestyle was explained by the frequency of engaging in physical exercise, number of working hours, consumption of food that contained fiber and number of sleeping hours. Stress level and experience of serious problems that respondents have had were used as indicators to measure mental health condition. It seems that some other indicators should also be used to describe mental health, because the contribution of stress level and the experience of serious problems in reducing the total variation was relatively low. This is probably why mental health was not significant in explaining the health index. However, it was found that the goodness of fit of the model was adequate when mental health was included as one of the latent factors. Thus, it is reasonable to consider mental health as one of the latent factors in the model; this idea is

supported by some previous work, such as the study by Nakayama et al. [12]. In addition to its relationship with the other latent variables, the health index is also explained by the biomarkers BMI, blood pressure, cholesterol level, blood glucose, the number of health problems, and general health condition. The latent factors, including socio-demography and mental health, were found to have a direct effect on the health index.

However, structural equation modeling also has some limitations. The structural equation modeling technique is quite complex because structural equation modeling simultaneously tests the relationships between directly observed and unobserved variables, and the relationships among unobserved variables. The user should be careful when constructing a relationship that is usually built based on the literature. Structural equation modeling also requires a large sample size, generally several hundred observations, as the precision of the estimates is affected by the sample size. The requirement of a large sample size incurs a higher cost. Another limitation of structural equation modeling is that there is no single "gold standard" for goodness-of-fit assessment. The user may refer to the literature to decide on the cutoff criteria for the goodness-of-fit indexes for their model.

## Conclusions

Our enthusiasm for structural equation modeling is based on its ability to enhance our understanding of health index modeling, as both latent and observed variables are considered. Through the use of structural equation modeling, one can understand better how direct and indirect factors affect the health index. In addition, the indicators which correlated significantly to these factors could be identified. Structural equation modeling is quite flexible as it allows one to use combinations of various types of data, such as dichotomous, ordered categorical, and continuous. Because ordered categorical and dichotomous data may not come from a normal distribution, the commonly used approach of maximum likelihood is not applied. Instead, the RWLS method is used for parameter estimation. In the present study, structural equation modeling was found to be pertinent to be used for analyzing the health index of a particular individual.

As with any statistical modeling techniques, interpreting the results of structural equation modeling depends on the quality of the measured data. A good fit of the specified measurement or structural model to the observed data indicates that the model is consistent with the relationships within the observed data. However, there may be other models that could specify the data as well. Whenever possible, the relative fit of alternative theoretically plausible models should be considered.

Structural equation modeling is not without its critics or problems, and we would not suggest otherwise. In this study, although we may never fully capture the intricacies of health index modeling, structural equation modeling has at least moved us in the direction of understanding a process of modeling interrelated variables more fully and completely.

**Acknowledgments** This research was supported by a grant from Universiti Kebangsaan Malaysia (UKM-ST-06-FRGS0011-2007). We wish to thank the Department of Community Health, Medical Faculty, Universiti Kebangsaan Malaysia, which provided us with the health survey data that were used in this study.

**References**

1. Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q. Multiattribute utility function for a comprehensive health status classification system: health utilities index mark 2. *Med Care*. 1996;34:702–22.
2. Ramirez M, Gallego JM, Sepulveda CE. The determinants of the health status in developing countries: results from the Colombian case. *Economia*. 2004;41:1–15.
3. Dhargupta A, Sen M, Goswami A, Mazumder D. Development of cognitive index to measure health status. *Indian Res J Ext Edu*. 2008;8:26–8.
4. Boniface DR, Tefft ME. The application of structural equation modeling to the construction of an index for the measurement of health-related behaviours. *Statistician*. 1997;46:505–14.
5. Chern JY, Wan TTH, Begun JW. A structural equation modeling approach to examining the predictive power of determinants of individuals' health expenditures. *J Med Syst*. 2002;26:323–36.
6. Stafford M, Sacker A, Ellaway A, Cummins S, Wiggins D, McIntyre S. Neighbourhood effects on health: a structural equation modeling approach. *Schmollers Jahrbuch*. 2008;128:1–12.
7. Department of Community Health, Medical Faculty, Universiti Kebangsaan Malaysia (UKM) (2002) Health risk assessment of Hulu Langat District. Bangi: UKM Press.
8. Boardman JD. Stress and physical health: the role of neighborhoods as mediating and moderating mechanisms. *Soc Sci Med*. 2004;58:2473–83.
9. Cheadle A, Pearson D, Wagner E, Psaty BM, Diehr P, Koepsell T. Relationship between socioeconomic status, health status, and lifestyle practices of American Indians: evidence from a plains reservation population. *Public Health Rep*. 1994;109:405–13.
10. Uitenbroek DG, Kerekovska A, Festchieva N. Health lifestyle behaviour and socio-demographic characteristics. A study of Varna, Glasgow and Edinburgh. *Soc Sci Med*. 1996;43:367–77.
11. Shi L. Socio-demographic characteristics and individual health behaviors. *Sociodemogr Factors Health Behav*. 1998;91:933–42.
12. Nakayama K, Yamaguchi K, Maruyama S, Morimoto K. The relationship of lifestyle factors, personal character, and mental health status of employees of a major Japanese electrical manufacturer. *Environ Health Prev Med*. 2001;5:144–9.
13. Olsson UH, Foss T, Troye SV, Howell RD. The performance of ML GLS, and WLS estimation in structural equation modeling under conditions of misspecification and non-normality. *Struct Equ Modeling*. 2000;7:557–95.
14. Muthén LK, Muthén BO. *Mplus user's guide*. 5th ed. Los Angeles: Muthén & Muthén; 2007.
15. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6:1–55.
16. Ullman JB. Structural equation modeling: reviewing the basics and moving forward. *J Pers Assess*. 2006;87:35–50.
17. Weston R, Gore PA Jr. A brief guide to structural equation modeling. *Counseling Psychol*. 2006;35:719–51.