The Validation of Questionnaire on Risk Perception of Developing Five Most Common Non-Communicable Diseases in Malaysia

Hafizah P.[1], Zaleha M. I.[2] , Shamsul A.S.[3]

[1,2,3]*(Department of Community Health, National University of Malaysia, Kuala Lumpur,Malaysia)*

Abstract : *The use of questionnaire as a tool to measure public perception on their risk of developing disease has been widely used by many researchers. However, in Malaysia, there is no validated questionnaire developed for such purpose yet. This study aim to develop and validate the questionnaire on risk perception of developing five most common non-communicable diseases (NCDs) in Malaysia. A cross sectional study was carried out among staff and clients of Kuala Rompin Health Clinic, Pahang aged 18 years and above from different ethnicities and age groups. This study was divided into 4 phases: i) developing the questionnaire involving related specialists and pre testing, ii) improving and editing the questionnaire and redistribution to 130 respondents for the pilot study, iii) factor analysis and confirmatory factor analysis, and iv) internal consistency reliability testing of the questionnaire . Factor analysis performed showed that factors loading values are all more than 0.4 and items in each domain were correlated with items in the same domain, but not with items in the other domain; findings which support the convergence and discriminant validity of the questionnaire. Confirmatory factor analysis came up with one five factors best fit model. Internal consistency reliability analysis on health related behavior, special practices, information seeking and perceived probability of diseases have good Cronbach α values of between 0.5 to 0.8, whereas intra-class correlation (ICC) for the perceived severity has a value of between 0.1 to 0.2. Thus, the questionnaire is successfully validated and considered as a useful tool to be used in measuring public perception on their risk of developing NCDs in Malaysia. However, improvement will still be necessary from time to time to ensure its relevance especially if it is to be used by different populations or countries of different background in Malaysia, taking into account the new changes in disease pattern.*

Keywords *- Malaysia, non-communicable diseases, questionnaire, risk perception, validation*

I. INTRODUCTION

For many years, researchers have been keen to develop a suitable tool which enables them to measure public perception. Clearly, this is no easy task to do. However, for those who opt to do so, their enthusiasm were based on what has been mentioned in Health Belief Model, whereby in order for an individual to change his/ her behavior towards a better health, he/ she must have his/ her own perception on the likelihood of experiencing a condition that would adversely affect their health, they concerned about the severity of the condition and believed that taking appropriate action towards the prevention of the disease or condition will outweigh the adverse effect of the action[1]. Thus, one of the possible methods that can be used to measure public perception is by using questionnaire.

The use of questionnaire as a tool to measure public perception has been widely used by many researchers. This is because; the use of questionnaire is one of the easiest and cheapest methods for data collection. Thus, it is one of the most widely used research tools especially in studies involving large number of respondents, enabling statistical analysis that will give bigger power compare to other methods [2],[3].

On top of that, questionnaire too had a significant role in epidemiological data collection. Most of the time, questionnaire will be partially or extensively used as source of information in an epidemiological investigation for history of exposure to certain chemicals, outcomes, modifiers and confounders. Obviously, the use of an invalid questionnaire is simply a waste of time and money and an adequate preparation and validation of questionnaires is thus essential for the quality of data [4]. Moreover, questionnaire is also one of the readily accepted methods by public as compared to the more invasive and long experimental study which involved procedure such as blood taking or require them to take medicine.

In order to rationalize the importance of having a questionnaire which is able to measure public perception on developing NCD in this country, we have to relate them to the current disease scenario. In the year 2002, data had shown that almost 60.0 percent of the cause of death and 46.0 percent of the global burden of disease occur in developing countries. The same scenario happened in the year 2009 whereby, for every 10 deaths occurred worldwide, six will be related to the NCD conditions, three due to communicable diseases, reproduction or nutrition and one is injury related. This phenomenon was believed to be due to the rapid changes

in the economic, social and demographic structure especially in the developing countries. On top of that, the ever increasing number in tobacco users, unhealthy dietary practice, lack of physical activity and other unhealthy life style are among the well-known causes for developing NCD [5], [6].

The five main diseases which are the most common NCD in Malaysia are hypertension, diabetes, heart disease, cancer and stroke [7], [8]. The evidence for their important role can be seen in the report by the Ministry of Health, whereby these five diseases are among the 10 most common causes of government hospital admissions. Reviewing the result of the third National Health and Morbidity Survey (NHMS) 2006, it shows that there was an increase in the prevalence of NCDs and its risk factors. For example, there was a twofold increase in the prevalence of diabetes from 6.3 percent in NHMS I in the year 1986 to 14.9 percent in NHMS III in the year 2006. The same goes for hypertension in which it increases from 32.9 percent in NHMS II in 1996 to 42.6 percent in NHMS III. The risk factors for NCD showed the same pattern in NHMS III whereby it was estimated that there were about 2.8 millions of Malaysians aged 18 years old and above who were current smokers, 5.5 million were physically inactive, 3.6 millions are overweight and 1.7 million were obese [9].

A study done at the state level regarding the risk factors of cardiovascular disease among the respondents aged 18 years and above from a few villages around Kuantan, Pahang showed that the prevalence of hypertension was 35.6 percent and the prevalence of diabetes was 54.8 percent [10]. Even though the percentages are lower compared to that of NHMS III, and the data might not represent the adult population of Pahang, it has to be taken seriously as this might be an important sign for NCDs potential to escalate in this state. Thus, this shows how important it is to have the right tool to measure public perception on the probability of them developing NCDs in the future so that the right action can be taken to overcome the problem. Therefore, the aim of the study is to develop and validate the questionnaire on risk perception of developing five most common non-communicable diseases in Malaysia.

II. METHODOLOGY

2.1 Questionnaire development

The development of the questionnaire involved the Public Health Specialists from the Department of Community Health, National University of Malaysia Medical Centre, consisting of Epidemiologist, Health Management, Health Promotion and Family Health Specialist to ensure the content validity of the questionnaire. Regarding the face validity of the questionnaire, pre- testing was conducted by distributing the form to the staff and visitors of the Kuala Rompin District Library in order to ensure its comprehensibility as well as its readability. Modifications to the questions were then conducted after taking into account all the inquiries and suggestions given by the respondents.

The original form of the questionnaire which had been modified were then redistributed to 130 respondents, consisting of the staff and clients of Kuala Rompin Health Clinic aged 18 years and above from various ethnicities and age groups. However, the respondent has to meet the criteria set by the researcher before they are eligible to be involved in this study. The criteria are: i) able to understand and communicate in Bahasa Melayu, and ii) have never been diagnosed with one or more of the five most common NCDs in Malaysia, namely hypertension, diabetes, heart disease, cancer and stroke. Data collected from this cross-sectional study were then analysed for the validation process of the questionnaire.

A total of six sections (Section A to F) and 24 items were designed for the questionnaire. Section A was the sociodemographic section whereby respondents will need to answer in a string form. The fifteen behaviors, special practice, information seeking and perceived severity (Section B, C, D and E) which were included were designed using a Likert scale. Each item was scored on a five point scale, such that a respondent most frequently doing the healthy activities or least frequently doing the unhealthy activities will give a score of '5'. Conversely, a respondent most frequently doing the unhealthy activities or least frequently doing the healthy activities will give a score of '1'. Questions in B, C and D sections were summed to obtain an overall high, moderate and low risk behavior; special practice towards better health was classified as high, moderate and low; and also health information seeking was categorized as good, moderate or low. As for the last section (Section F), which is the perceived probability of disease, it was also designed using a Likert scale but in this case, each item was scored on a ten point scale, such that a respondent who perceived himself/ herself to be in high probability of acquiring diseases listed will give a score of 10 while those who perceived themselves to be in low probability of acquiring diseases listed will give a score of 1. This section was also summed to obtain an overall high, moderate and low perceived probability of disease.

Internal consistency will determine whether each item listed in the questionnaire is able to represent each of the concepts to be measured. As for validity, it will determine whether it measures what it aims to measure and this can be examined by two methods which are construct validity and content validity. Content validity shows whether items in a questionnaire cover the intended topics clearly; whereas construct validity which is the extent to which the questionnaire supports pre-defined hypotheses, is assessed by whether it produces an anticipated set of relationships with other variables such as clinical evidence [11].

2.2 Construct validity

Construct validity was tested by performing Factor Analysis (FA) on the data collected. The objective of performing factor analysis is to see whether a smaller set of variables could be used to summarise the data, with little loss of information [12]. In this study, factor analysis was performed by setting the factor solution into five factors because during the questionnaire development phase, five domains for FA have been identified; namely behavior, special practice, information seeking, perceived severity and perceived probability of disease. However, unlike the other domains, behavior domain only requires respondents to answer based on the coded scale rather that the real Likert scale. Thus, this domain was dropped from the analysis.

Apart from FA, Confirmatory Factor Analysis (CFA) was also performed on the collected data. This is because, CFA helps to fit in the data collected to the factor model developed. Thus, CFA was also used to assess the convergent and discriminant validation of the measurement [13]. In addition, given the fact that the proposed model was based on the technical input from the respective specialists, previous empirical research and theoretical findings, the CFA approach was considered as the most appropriate method to statistically [14] confirmed the proposed factors of risk perception on developing the NCDs under study.

2.3 Internal consistency reliability analysis

Internal consistency reliability analysis was tested by using Cronbach's alpha (α). The higher the α coefficient, the more consistent is the questionnaire items in measuring the variables under study. The Cronbach α was set at 0.5 for this current study.

III.   RESULT

Socio-demographic characteristics of the respondents are shown in Table 1.

Table 1: Respondents socio-demographic characteristic:

Socio-demographic characteristic	n	Mean (Standard deviation)	Percentage
Age (years)	130	31.77(9.44) (Minimum:18, Maximum: 71)	
Gender			
Male	42		32.3
Female	88		67.7
Ethnic group			
Malay	113		86.9
Aborigines	10		7.7
Indian	3		2.3
Others	4		3.1
Highest formal education attained			
Not schooling	5		3.8
Primary school	13		10.0
Secondary school	59		45.4
College/ University	53		40.8
Working class			
Male:			
Government	19		45.2
Private company	7		16.7
Self employed	7		16.7
Not working/ student	9		21.4
Female:			
Government	36		40.9
Private company	10		11.4
Self employed	3		3.4
Not working/housewives	39		44.3
Household income (RM)			
Median:	130	1,500	
Interquartile range:			
25%		1,000	
50%		1,500	
75%		3,500	

Household income class		
<RM1,000	27	20.8
RM1,000-RM3,000	68	52.3
RM3,001-RM5,000	28	21.5
>RM5,000	7	5.4
Have close relatives (parents, / siblings) with one or more of the listed diseases:		
Hypertension	Yes: 59	45.4
	No: 71	54.6
Diabetes	Yes: 36	27.7
	No: 94	72.3
Heart disease	Yes: 23	17.7
	No: 107	82.3
Cancer	Yes: 10	7.7
	No: 120	92.3
Stroke	Yes: 8	6.2
	No: 122	93.8

All the respondents completely answered the questionnaire, thus all of the 130 data obtained were analyzed. Results showed that majority of the respondents are female (67.7 percent), Malay (86.9 percent) with the mean age of 31.77 years old (standard deviation 9.44). Most of the respondents have attained formal education up to the secondary school level (45.4 percent), work as the government servant (45.2 percent) for male respondents and not working/ housewives (44.3 percent) for female respondents; and household income of between RM 1,000 to RM 3,000 a month (52.3 percent).

With regards to reported family members with any of the NCDs listed, it follows the NCD diseases pattern in Malaysia whereby the highest reported disease is hypertension (Yes: 45.4 percent). This is followed by diabetes (Yes: 27.7 percent), heart disease (Yes: 17.7 percent), cancer (7.7 percent) and stroke (6.2 percent).

3.1 Construct validity

The result of FA performed based on four factors solution showed that one of the items in the special practice domain (recite dua') has factor loadings value of 0.391 (not shown here). This value is lower than 0.4 which is considered as the cut-off point for a good factor loading [15]. Thus, in order to improve the construct validity of this questionnaire, the 'recite dua' item was dropped. The data was then reanalyzed and this time without the dropped item. Result showed that the percentage of variance explained for this questionnaire has improved from 51.2 percent to 52.9 percent. On top of that, the Scree Plot (Figure 1) shows that the plot levels off at factor six. This shows that at least five factors are suitable to be retained.
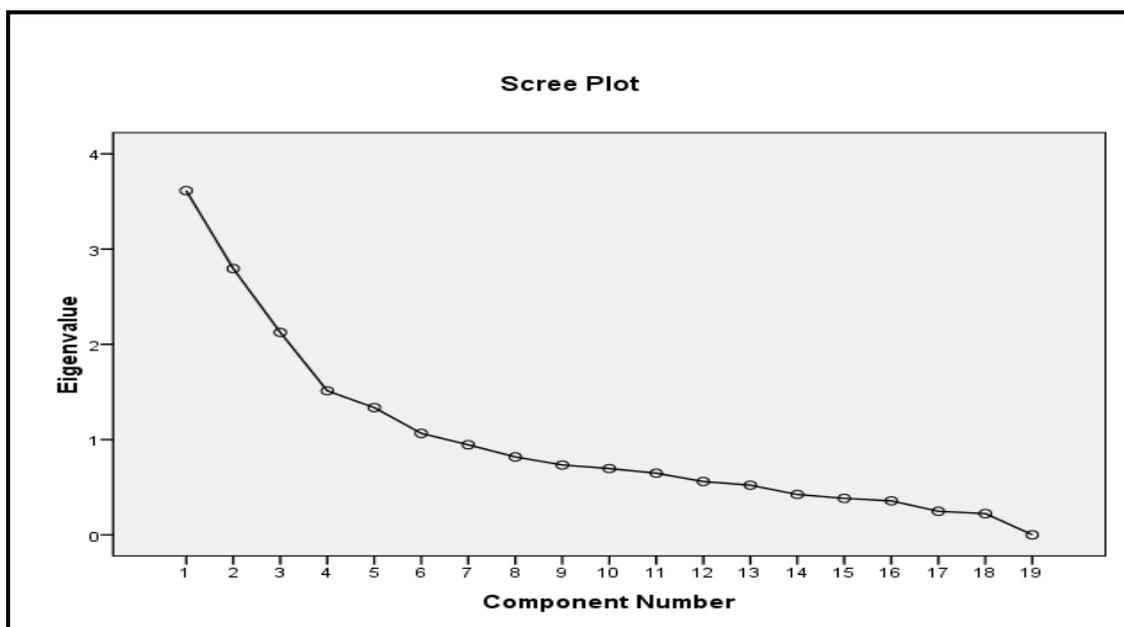

Figure 1: Scree Plot

Table 2: Rotated Component Matrix

Items	Disease probability factor	Information seeking factor	Perceived severity factor	Special practice factor
Herbs	-	-	-	0.57
Nutritional labeling	-	-	-	0.53
Health screening	-	-	-	0.53
Recreation	-	-	-	0.53
Newspaper/ magazine	-	0.51	-	-
Radio/ television	-	0.75	-	-
Internet surfing	-	0.69	-	-
Consulting Health Personnel	-	0.61	-	-
Consulting others	-	0.47	-	-
Perceived Severity of				
Hypertension	-	-	0.82	-
Diabetes	-	-	0.75	-
Heart disease	-	-	0.66	-
Cancer	-	-	0.76	-
Stroke	-	-	0.55	-
Risk perception of				
Hypertension	0.68	-	-	-
Diabetes	0.75	-	-	-
Heart Disease	0.84	-	-	-
Cancer	0.78	-	-	-
Stroke	0.82	-	-	-

The factor loading of < 0.40 is suppressed for presentation

As for the rotated component matrix (Table 2), result shows that all of the factor loadings showed a value of more than value set which is 0.4. At the same time, items in each domain were correlated with items in the same domain, but they are not correlated with items in the other domain. CFA was then conducted on the data collected to get the best fit measurement Structural Equation Modeling, SEM model. Results of the analysis indicate that one model (Figure 2) provide a good fit to the data with parameters 16 of; chi-square (df) = 0.052 (3), p value= 0.997, relative chi-square= 0.017, GFI= 1.000, AGFI= 0.999, CFI= 1.000, PRATIO= 0.300 and RMSEA= 0.000.
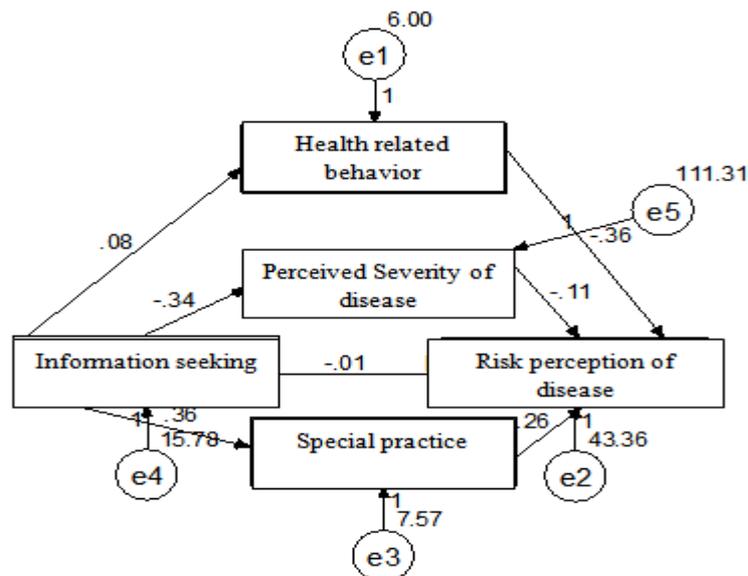
Chi-square (df) = .052 (3); P value (>=0.05) = .997;
Relative Chi-Sq (<=2) = .017;
GFI(>=0.95) = 1.000; AGFI(>=0.9) = .999;
CFI(>=0.9) = 1.000; Pratio = .300;
RMSEA(<=0.08) = .000.
(Unstandardized estimates)

Figure 2: SEM Model for risk perception of developing five most common non-communicable diseases

3.2 Internal consistency reliability analysis

Cronbach's α reliability analysis was computed for the 20 items in the health related behavior, special practice, information seeking and perceived probability of disease sections to measure internal consistency. Results of the analysis indicate that Cronbach's α was between 0.5 to 0.8 for all factors which is satisfactory except for the health related behavior which had a Cronbach's α of < 0.5 (α =0.2). As for the perceived severity section, intra-class correlation analysis revealed a value of between 0.1 to 0.2 (Table 3).

Table 3: Internal consistency reliability analysis result

Factor	Mean (standard deviation)	Cronbach α
Exercise/physical activity	2.92 (1.17)	
Smoking	4.55 (1.14)	
Alcohol	4.97 (0.17)	0.2
Fast food	3.62 (1.10)	
Snacking	3.81 (1.14)	
Herbs	1.81 (1.03)	
Nutritional labeling	3.14 (1.40)	
Health screening	1.98 (1.16)	0.5
Recreation	2.62 (1.31)	
Newspaper/ magazine	3.51 (1.22)	
Radio/ television	3.02 (1.21)	
Internet surfing	2.33 (1.30)	0.7
Consulting Health	2.52 (1.19)	
Personnel	2.63 (1.18)	
Consulting others		
Perceived Severity of	3.75 (1.26)	0.06(-0.11,0.23)*
Hypertension	3.42 (1.17)	0.13(-0.04,0.30)*
Diabetes	2.29 (1.07)	0.24(0.07, 0.39)*
Heart disease	1.95 (1.33)	0.03(-0.14,0.20)*
Cancer	3.58 (1.21)	0.16(-0.02,0.32)*
Stroke		
Risk perception of	3.17 (2.12)	
Hypertension	2.85 (1.92)	
Diabetes	2.44 (1.77)	0.8
Heart Disease	2.05 (1.53)	
Cancer	1.97 (1.40)	
Stroke		

IV.   Discussion

FA performed on the data collected shows that the number of factor solution is four. This result fulfilled the criterion for the number of factors that should be obtained which are: 1) obtain all factors with eigenvalue ≥ 1 (Kaiser's criterion or rule), 2) Scree plot: The Cattell rule is to pick all factors prior to where the plot levels off and 3) comprehensibility: though not a strictly mathematical criterion, limiting the number of factors to those who are readily comprehensible and in this study it is set at four based on the number of domains [15]. Furthermore, because items in each domain were correlated with items in the same domain, but not with items in the other domain; and the factors loading value are all more than 0.4 after dropping the item recite dua', this finding supports the convergence and discriminant validity of the questionnaire.

As we move to the model developed from the CFA, only one best fit model that fulfilled the criterion for the Structural Equation Modeling, SEM [16] parameters for best fit model is successfully developed. This five factors model is viewed as the best model because apart from being developed by including all the five factors under study, it also supported the Health Belief Model [17] that highlighted the factors influencing individual perception on probability of diseases which in turn govern the future behavioral changes of an individual. According to the Health Belief Model, those factors are perceived severity, special practice and current health related behavior that were also influenced by information seeking.

Discussing on the reliability analysis, the items for the special practice, information seeking and probability perception have a satisfactory Cronbach α of 0.5, 0.7 and 0.8 respectively. This fulfilled the α value set for this study which is 0.5. However, the behavior section item has a Cronbach α value of 0.2. This is lower than the value set for this study. One of the possible explanations for this is because the αvalue is sensitive to the number of item for each section. This too was mentioned by previous researcher [18], [19] whereby a Cronbach

α value of less than 0.7 is rather common for the study with items of less than 10 for each section. Thus, this value can be improved by increasing the item numbers in the questionnaire. Nevertheless, since three out of four sections in this questionnaire have a Cronbach α ≥ 0.5, thus it can be concluded that this questionnaire has a satisfactory internal validity [20].

As for the perceived severity section, the intra-class correlation values were rather lower (Table 3) than the value set for this study which is 0.7 (between 0.1 to 0.2). The possible explanation for this is because for a score derived from questionnaire (soft measures), it is very difficult to achieve high ICC value as compared to the 'hard measurement' such as body weight or skin thickness. Thus, even though one of the options is to drop the item which caused low ICC, but if it can affect the coverage of the questionnaire that is the content validity, then the researcher should reconsider before dropping it [15] .

V.   CONCLUSION

Even though the results of the validation process carried out on the questionnaire are limited for the few items in it, nevertheless it is successfully validated and is considered as a useful tool to be used in measuring public perception on their risk of developing NCDs in Malaysia. However, improvement will still be necessary from time to time to ensure its relevance, especially if it is to be used by various populations or countries of different background from Malaysia and also taking into account the new changes in diseases pattern.

REFERENCES

[1]   J. Hubley and J.Copeman. Understanding health and illness behavior, in *Practical Health Promotion*, pg. 68. Cambridge: Polity Press. 2008

[2]   J.H. McMillan and S.Schumacher. *Research in Education: A Conceptual Introduction* (5th Ed.), New York, NY: Longman. 2001

[3]   M.J. Nieuwenhuijsen. Design of exposure questionnaires for epidemiological studies. *Occup Environ Med 2005;62:272–280. doi: 10.1136/oem.2004.015206*. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1740990/. 2005.Accessed on October 18, 2011.

[4]   The IEA European Questionnaire Group.*Epidemiology deserves better questionnaires*.. www.iea-europe.org/download/Questionnaires.pdf. 1998. Accessed on October 18, 2011.

[5]   World Health Organization.*World health report: chapter 3: perceiving risk*. www.who.int/entity/whr/2002/chapter3/en/index.htm. 2002.Accessed on September 5, 2010.

[6]   World Health Organization.*World health statistics: table 2: cause-specific mortality and morbidity*. www.who.int/whosis/whostat/EN_WHS09_Table2.pdf. 2009Accessed on September 5, 2010.

[7]   Health Fact.*Health informatics centre planning and development division Ministry of Health Malaysia*. 2008

[8]   Health Fact.*Health informatics centre planning and development division Ministry of Health Malaysia*. 2009

[9]   I. M. Fesisul. Ppt: National strategic plan for non-communicable diseases (NSP-NCD). www.malaysianheart.org/file_dir/17730830514a51a6f6b9f87.pdf . 2009.Accessed on March 7, 2011.

[10]   S. F.U. Akter, A. R. M. Fauzi, M.S. Nordin, S. Satwi, A. Mohamed, M.A. Aznan and D. Samsul. Prevalence of cardiovascular risk factors in a selected community at Kuantan, Pahang, Malaysia. *International Journal of Medicine and Medical Sciences Vol. 2(10), pp. 322-328, October 2010*. http://www.academicjournals.org/ijmms . 2010. Accessed on March 7, 2011.

[11]   J. Dawson, R.Fitzpatrick and A. Carr. Questionnaire on the perceptions of patients about shoulder surgery. web.jbjs.org.uk/cgi/content/abstract/78-B/4/593. 1996Accessed on October 18, 2011

[12]   J. DeCoster. Overview of factor analysis. . http://www.stat-help.com/notes.html 1998.Accessed on November 9, 2011.

[13]   Y. H. Chan.Biostatistics 302: principal component and factor analysis. *Singapore Med J 2004 Vol 45(12 : 558*. www.sma.org.sg/smj/4512/4512bs1.pdf .2004.Accessed on May 11, 2011.

[14]   A.H. Mohammed, N. Hartley and T.Ward. Measuring banks' automated service quality: a confirmatory factor analysis approach. *Marketing Bulletin, 2005, 16, Article 1*. http://marketing-bulletin.massey.ac.nz/V16/MB_V16_A1_AlHawari.pdf. 2005.Accessed on December 2, 2011

[15]   L. Naing. Intermediate to advance biostatistics in medical sciences workshop 28-30 jan. Accessed on January 22, 2010.

[16]   J.L. Arbuckle. AMOS 6.0 user's guide: appendix c. United States of America.Amos Development Corporation. 2005

[17]   University of Twente. Health belief model. Theory cluster: Health Communication. http://www.utwente.nl/cw/theorieenoverzicht/Theory%20clusters/Health%20Communication/Health_Belief_Model.doc/   .2010. Accessed on March 7, 2011.

[18]   G. Turconi, M. Celsa, G. Biino, M.A. Sartirana and C. Roggi. Reliability of a dietary questionnaire on food habits, eating behaviour and nutritional knowledge of adolescents. *European Journal of Clinical Nutrition (2003) 57, 753–763*.http://www.nature.com/ejcn/journal/v57/n6/full/1601607a.html. 2003.Accessed on January 9, 2012.

[19]   J. Pallant. SPSS: survival manual. Crows Nest: Allen & Unwin. http://www.allenandunwin.com/spss/further_resources.html .2004.Accessed on January 9, 2012.

[20]   J.M.Bland and D.G. Altman. Statistics notes: Cronbach's alpha. BMJ 1997;314:572. http://www.bmj.com/highwire/filestream/324955/field_highwire_article_pdf/0.pdf .1997.Accessed on January 8, 2012.