

Comparison of four analytic strategies for complex survey data: a case-study of Spanish data

Mohd Masood^(1,2,3), Tim Newton⁽²⁾, Daniel D Reidpath⁽³⁾

(1) Centre of Studies for Population Oral Health and Clinical Prevention, Faculty of Dentistry, Universiti Teknologi MARA, Shah Alam, Malaysia.

(2) Division of Population and Patient Health, King's College London, United Kingdom.

(3) Global Public Health, Jeffrey Cheah School of Medicine and Health Science, Monash University, Bandar Sunway, Malaysia.

CORRESPONDING AUTHOR: Mohd Masood - Lecturer, Centre of Studies for Population Oral Health and Clinical Prevention, Faculty of Dentistry, Universiti Teknologi MARA, Shah Alam, Malaysia - Phone:0060162209941 - Fax: 0060355435803 - drmasoodmohd@gmail.com

DOI: 10.2427/11584

Accepted on March 3, 2016

ABSTRACT

Background: The aim of this secondary data analysis was to investigate the effect of four different analytical strategies: Model Based Analysis (MBA), Design Based Analysis (DBA), Multilevel Model Based Analysis (MMBA) and Multilevel Design Based Analysis (MDBA), on the model estimates for complex survey data.

Methods: Using data from the World Health Survey-Spain explanatory models for the outcome, Metabolic Equivalent of Task (METs) were calculated using MBA, DBA, MMBA and MDBA. Regression coefficients, standard errors (SE) and the Akaike Information Criterion (AIC) from all the models were compared.

Results: DBA gave the highest estimates for most of the variables, including consistently higher SE than all other models - 20% to 48% higher than estimates for MBA, 10% to 37% for MMBA and 23% to 35% for MDBA. The SE for MDBA were 2.5% to 13% higher than estimates derived from MMBA in level 1 predictors, but SE in MMBA was higher by 18% for level 2 predictors. Values of AIC suggested the model derived by MDBA was the best fit and DBA the poorest fit of the four models.

Conclusion: The MDBA appeared to be the most appropriate approach to analyse complex survey data on the basis that it had the lowest AIC. To confirm the findings of the present study a simulation study with hypothetical data would be required.

Key words: Design based, multilevel, complex survey

INTRODUCTION

Large epidemiological surveys almost universally employ multistage complex sampling procedures for data

collection, where clusters (or primary sampling units – PSUs) are sampled at the first stage, sub-clusters at the second stage, etc., until final units (typically an individual) are sampled at the final stage [1, 2]. The World Health

Survey is an example of multistage complex sampling where districts were selected as PSUs, enumerated area as secondary sampling units (SSUs) and households as tertiary sampling units (TSUs) [3]. Complex sampling strategies are used because they often make the process of estimation more efficient by reducing the cost of data collection for a given level of precision [4, 5]. Complex sampling strategies are particularly useful in the case of a population that is geographically dispersed, where a simple random sample would entail traveling significant distances and require greater time and effort for data collection. Complex sampling approaches almost inevitably result in unequal probabilities of individual selection, giving rise to the so called design features of complex survey data [1].

A complex sampling strategy, however, also imposes a multilevel or hierarchical structure on the data. For instance, in the World Health Survey described, above, a multilevel structure is present whereby individuals are embedded within a household, households are embedded within an enumerated area and enumerated areas within a district. Data from such complex surveys are therefore the product of both an underlying multilevel structure and the design features. If a multilevel structure has been imposed on the data through the sampling strategy, then that multilevel structure can itself become the focus of research. For example, multi-level analysis was used in a study of health care expenditure in which the authors estimated the simultaneous effects of individual-level and cluster-level characteristics on maternal health care spending [6, 7]. The rise of interest in multilevel analyses of hierarchically structured data introduces another dimension to consider in the analysis of complex survey data.

The combination of the presence of design features and multilevel data suggests four possible approaches to the analysis of data from complex survey design. The first approach is to analyse the data as if it was a simple random sample derived from the population ignoring both design features and multilevel structure of the data. This analysis can be termed a "model based analysis" (MBA) [8], for example the application of ordinary least squares regression (e.g., [9]). The second approach is to take account of the design features and the clustering in the data, while still treating all predictors as if they are measured at the lowest level – a "design-based analysis" (DBA) (e.g., [10]). This would involve including the weighted sample to provide unbiased estimates of the independent variables in the regression model [11-13]. The third approach would be to ignore the design features but instead focus on the multilevel nature of the data, allowing interpretations of individual and area level effects on individual outcomes using multilevel analysis (e.g., [14]). Such an approach would explain variation in the dependent variable at one level as a function of variables defined at other levels, plus interactions within and between levels [11], this could be described as a "multilevel model based analysis"

(MMBA). Like its non-multilevel counterpart, the model-based analysis may lead to biased estimates when employed in samples that include design features in the data [15]. Finally, the fourth approach is an analysis in which both the design features and the multilevel nature of the data are taken into account – a "multilevel design-based analysis" (MDBA) (e.g., [16] and [17]).

Previous research has studied the effect on model estimates of ignoring a design-based analysis of data from surveys employing complex sampling strategies [12, 18, 19] and of ignoring multilevel structure in multilevel data [11]. However, there has never been a systematic comparison of the effect of the four different modeling strategies (MBA, DBA, MMBA and MDBA) on the model estimates, when the data are collected using a complex survey design. Therefore, this study will investigate the effect of the MBA, DBA, MMBA and MDBA analytic strategies on model estimates from Spanish, World Health Survey (WHS) data.

METHODS

This study comprised the secondary analysis of a publicly available data set. Model estimates derived from four analytical strategies were compared: MBA, DBA, MMBA and MDBA.

Data source

The World Health Survey (WHS) is a large cross-sectional survey, that was administered in 70 countries between 2002–2003 to assess healthcare expenditure, adult mortality, birth history, risk factors, chronic health conditions and the coverage of health interventions [20]. The WHS adopted several steps to ensure standardisation and comparability across diverse sites and times, including extensive interviewer training, standardised measurement tools and techniques, an identical questionnaire and instrument pretesting. The WHS's sampling frame covered 100% of a Spain's eligible population and no ethnic groups nor geographic areas were excluded from the sampling frame. The target population included any adult, male or female, aged 18 years living in private households, who were not out of the country during the survey period. The WHS used a multistage stratified design in most countries including Spain with each elementary unit having a defined probability of selection [20]. WHS data is made freely available by the World Health Organization for secondary analysis by the research community.

The multistage sampling in the WHS used statistical enumeration areas as the Primary Sampling Units (PSUs). These were identified in the WHO sampling

documentation as naturally occurring groupings with clear, non-overlapping boundaries [21, 22]. The strata chosen varied by country and reflected local conditions. Some examples of the factors that were used for stratification were geography (e.g. North, Central, South), level of urbanization (e.g. urban, rural), socio-economic zones, provinces (especially if health administration is primarily under the jurisdiction of provincial authorities), or presence of health facilities in the area. The PSUs were used as a clustering or grouping variable in the current analysis. In WHS, stratification was done at the first stage of the sampling. Once the strata was chosen and justified, all stages of sample selection were conducted separately in each stratum. More detailed information on the sampling approach can be found elsewhere [23].

Data from Spain was selected for this study based on the sample size (n=6364) and the number of PSUs (997). The Spanish WHS also had an extremely high response rate (95.5%) compared with other WHS countries. After excluding cases with missing data, the final sample for analysis was 6079 individuals.

Variables

The outcome variable used in this study was a measure of physical activity per week in units of Metabolic Equivalent of Task (METs). One MET is defined as the energy spent sitting quietly (equivalent to 4.184 kJ per hour per kilogram of body mass) [24]. In the WHS, to assess physical activity respondents were asked to report the number of days during the last week on which they engaged in vigorous and moderate walking and the duration of such activities. Taking the different intensities and duration of the activities into account, a measure of energy expenditure per individual was estimated [25]. METs were selected for this analysis specifically because of the high intra-class correlation for this variable (ICC=0.23).

Explanatory variables

Age, sex, education, occupation, fruit and vegetable intake, body mass index (BMI), household income and setting (urban/rural) were the key explanatory variables. They were selected on the basis of factors identified in previously published research exploring predictors of activity as assessed by METs [26-28].

Age was measured in years of life. Education was measured in number of years of schooling. Occupation was a categorical variable distinguishing “employed”, “housewife”, “retired” and “not working”. BMI, defined as mass in kilograms divided by height in meters squared, was based on self-reported height and weight. The WHS did not contain a comprehensive nutrition survey measuring whole diets, but rather sought measurement of fruit and vegetable intake only. Two questions were used: “How many servings of fruit do you eat on a typical day?” and “How many servings of vegetables do you eat on a typical day?” [29]. Household wealth was defined in terms of ownership of material possessions, with each individual assigned a wealth score on the basis of ownership of a range of household goods. Factor analytic procedures were used to provide a wealth score for each household and households were then divided into quintiles of wealth. The urban-rural nature of the PSU was provided in the WHS dataset based on local definitions. These urban/rural PSUs were used as area level or level 2 predictors for multilevel analysis. Table 1 describes the explanatory variables used in all four models.

Analytical strategies

Four analytical strategies were developed. The first model was a model based analysis (MBA) which assumed the data were drawn as a simple random sample from the population. All predictors were treated as individual

TABLE 1. Summary of the statistical methods and variables used in the analyses

	MODEL BASED (MBA)	DESIGN BASED (DBA)	MODEL BASED MULTILEVEL (MMBA)	DESIGN BASED MULTILEVEL (MDBA)
OUTCOME VARIABLE	METs	METs	METs	METs
INDEPENDENT VARIABLES (LEVEL 1)	Age, Sex, Education, Occupation, Household income, Fruits and Vegetables intake, BMI, Setting (Urban, Rural).	Age, Sex, Education, Occupation, Household income, Fruits and Vegetables intake, BMI, Setting (Urban, Rural).	Age, Sex, Education, Occupation, Household income, Fruits and Vegetables intake, BMI.	Age, Sex, Education, Occupation, Household income, Fruits and Vegetables intake, BMI.
SURVEY DESIGN	-	Survey Design: PSU, Strata, Individual Weights.	-	Survey Design: Individual Weights, weights at other levels
LEVEL 2	-	-	Level 2: PSU, Setting (Urban, Rural).	Level 2: PSU, Setting (Urban, Rural).

level attributes and no account was taken of the design features or the clustering of the data. The second model was a design based analysis (DBA) which took account of design features of the data, whereas all the predictors were treated as individual level attributes. The estimation was based on inverse probability weighting and design based standard errors. The third model was a multilevel model based analysis (MMBA). In this third model all the predictors were treated as level 1 predictors except the urban-rural predictor which was treated as a level 2 (i.e., PSU) predictor. Design features were not applied to the data, however, the multilevel nature of the data was considered where individuals were clustered within PSU. The fourth model was a multilevel design based analysis (MDBA). The analysis took account of clustering as well as the design features.

All analyses conducted using the M-Plus statistical package [30]. DBA was performed using command "Analysis Type = COMPLEX" with input of "CLUSTERING", "STRATIFICATION" and "WEIGHTS" variables. MMBA was performed using command "Analysis Type = TWOLEVEL" with input of "CLUSTERING" variable. MDBA was performed using command "Analysis Type = COMPLEX TWOLEVEL" with input of "CLUSTERING"

"STRATIFICATION" and "WEIGHTS" variables and urban and rural setting at level 2. Regression estimates and standard errors for all four models were compared. Additionally the AIC (Akaike Information Criterion) was also calculated for all the four models to measure the relative goodness of fit and to compare the best fitted model among the four models.

RESULTS

The descriptive statistics (weighted and unweighted) for the outcome variable (METs) and each of the explanatory variables are shown in Table 2.

Table 3 summarises the four models of METs using the 8 predictors. The parameter estimates, standard errors and level of significance for each model is shown as well as the AIC.

The average age of the population assuming a model based design is about 2 years older than the design based sample; the model based analysis also reported about 8% fewer METs. There was little difference in the estimated BMI. Differences may similarly be observed in years of school, the percentage in occupation and the percentage

TABLE 2. Unweighted and weighted descriptive statistics of outcome and explanatory variables.

	UNWEIGHTED		WEIGHTED	
CONTINUOUS VARIABLES				
	Mean	SD	Mean	SD
MET	3068.7	60.05	3322.9	103.7
Age	52.7	0.23	50.3	0.33
Education (School years)	8.98	0.068	9.40	0.11
Fruits and Vegetable intake	3.44	0.022	3.01	0.032
BMI	26.04	0.054	25.93	0.081
CATEGORICAL VARIABLES				
	n	%	N	%
Gender				
Female	3677	58.6	18434123	58.0
Male	2598	41.4	13337321	42.0
Occupation				
Employed	2462	39.2	13280952	41.8
Housewife	1703	27.1	8428540	26.6
Retired	1673	26.7	7337952	23.1
Others	437	7.0	2713999	8.5
Household Income				
Lowest quintile	1330	21.5	5216859	16.7
2nd quintile	1172	18.9	5741298	18.3
3rd quintile	1257	20.3	6137899	19.6
4th quintile	1241	20.0	6988845	22.2
Highest quintile	1196	19.3	7290751	23.2
Setting				
Rural	1780	28.3	7150839	22.5
Urban	4495	71.6	24620605	77.5

in each wealth quintile. The highest difference of nearly 6% was seen in urban or rural settings, average BMI and level of fruit and vegetable intake.

There were both consistencies and variations in the results of the analyses across the four models. At a superficial level, predictors that were identified as statistically significant in one model were, with few exceptions, identified as statistically significant in the other models. Gender, for instance was a significant effect across the four models and urban-rural setting was not a significant effect in any model. Education level was an exception – statistically significant in all models except the MBA.

The regression estimates did not show a consistent pattern across the models. DBA gave the highest estimates among the four models for most of the variables but it showed lowest estimates for gender. That is, the effect of being male was about 32% less in the DBA than it was for the MMBA or MDBA. The estimates for the effect of BMI were lower for the MMBA and MDBA models and higher for the MBA and DBA models. The

lowest variation between the models (15%) was seen in age. The urban-rural variable was not significant and showed extreme variation in the estimates.

For the significant effects of gender, occupation and fruit-vegetable intake, the estimates from the two multilevel models (MMBA and MDBA) were more consistent with each other than they were with the single level models (MBA and DBA). For the age and years of education estimates, however, the two design based analyses were more consistent with each other than they were with the model based analyses. For the estimate of gender, the model based and design based (non-multilevel) were reasonably consistent.

Some interesting patterns were observed in standard error values across the four models. As one would expect, DBA estimates showed consistently higher standard errors than the other models: 20% to 48% higher than the MBA analysis, 10% to 37% higher than the MMBA and with the MDBA by 23% to 35%. On the other hand, MDBA had consistently higher standard errors by 2.5% to 13%

TABLE 3. Multivariate linear regression analysis showing MET association with various micro and macro level explanatory independent variables, with and without consideration of sampling design.

	MODEL BASED (MBA)		DESIGN BASED (DBA)		MULTILEVEL MODEL BASED (MMBA)		MULTILEVEL DESIGN BASED (MDBA)	
	Estimate	SE	Estimates	SE	Estimates	SE	Estimates	SE
Intercept	5180.9	600.9*	5828.9	879.7*	3336.3	632.4*	4512.0	619.0*
Age	-32.0	5.1*	-38.1	7.3*	-33.1	4.7*	-37.0	5.3*
Gender	Reference Group							
Female	Reference Group							
Male	1146.2	144.6*	1084.2	215.5*	1504.4	143.9*	1499.1	153.6*
Education (School years)	-21.6	14.0	-58.8	21.8*	-47.6	15.1*	-57.5	16.0*
Occupation	Reference Group							
Employed	Reference Group							
Housewife	-1192.7	189.2*	-1197.5	241.6*	-839.8	169.6*	-785.0	176.5*
Retired	-1348.0	208.6*	-1200.4	283.0*	-1619.0	193.9*	-1519.1	200.5*
Others	-1220.0	249.6*	-1453.8	313.5*	-1062.1	234.8*	-1078.6	240.9*
Household Income	Reference Group							
Lowest quintile	Reference Group							
2nd quintile	-227.1	193.0	-101.2	266.8	33.6	178.5	-57.8	204.6
3rd quintile	70.4	189.6	176.0	310.7	236.4	188.4	198.1	209.6
4th quintile	61.8	190.1	152.7	289.0	164.4	205.1	28.1	219.4
Highest quintile	-112.1	193.7	415.3	312.8	55.6	212.4	17.8	227.4
Fruits and Vegetable intake	181.1	35.4*	251.8	59.9*	117.4	37.2*	94.0	38.9*
BMI	-45.6	15.3*	-61.2	21.6*	-20.7	13.9	-27.4	14.9*
Setting	Reference Group							
Rural	Reference Group							
Urban	-247.8	138.5	-128.2	267.3	-290.1	241.2	-47.6	196.6
	AIC= 116323.8		AIC=116635		AIC= 116068		AIC=115911	

* p-value <0.05; SE Standard Error

in comparison to the MMBA model in level 1 predictors, but the standard error in the MMBA model was higher by 18% for the level 2 predictor. The MDBA also had a higher standard error when compared to the MBA model but the variation was comparatively smaller (from 3.7% to 12.5%). In comparison the MDBA model had a lower standard error by 6.8% in the occupation 'housewife' category. The AIC value was lowest for the MDBA model and highest for DBA.

DISCUSSION

Four possible methods (MBA, DBA, MMBA, MDBA) of data analysis for complex survey data were compared. On the basis of the fit of the models (AIC) the best analysis is, in order of best fit: MDBA, MMBA, MBA, followed by the DBA.

Data collection incorporating multi-stage sampling and design features is common in large epidemiological surveys. In the past it was relatively difficult to take account of the complex survey design in the data analysis, but recent advances in statistical software have made design based analysis accessible. Notwithstanding the availability of the improved software, judgments still need to be made about the best approach to take with the data. The results of this analysis raise important questions about how researchers should approach data derived from surveys with complex sampling designs. Although the model-based methods have gained popularity over the design-based methods as they can be readily implemented using standard commercial software, there is a consensus among statisticians that a straightforward MBA is inappropriate, since the common observation is that such an approach underestimates the uncertainty of the estimate. The analysis here, however suggests that the estimates themselves can vary substantially in magnitude although not in significance. The standard errors are, unsurprisingly, higher for the design-based approach in all the explanatory variables. This generally did not affect the statistical significance of the results, except in one case where the MBA estimate was non-significant.

In general the multilevel models tended to show greater agreement with each other than with the other models. Most of the estimates for MDBM were closer to MMBA as compared to other models, moreover, the AIC of the MDBA was closest to that of the MMBA model. The MMBA explicitly models the clustered nature of the data which should narrow the standard errors and eventually increase inferential accuracy. However incorporation of design features in MMBA increases standard error. Thus combining both design features and multilevel modeling leads to a standard error estimate that falls in between DBA and MMBA. Regarding the reliability of the estimates, the standard errors are lower with the MMBA for all of the variables. Therefore, the design-based analysis estimates are, overall, more precise than those from the model-based analysis.

It could be argued that the significance of variables in the regression models appears to be largely invariant, that more complex analytical methods offer little over simple MBA. Indeed, traditionally the availability of software and computing to conduct complex analyses was limited. If an analytic strategy is imperfect (as they all are) but yields the same general interpretation, does it matter if it is strategy A or strategy B is selected? However we have seen that the accuracy of such estimates will vary with the analytical method chosen which may be important in reducing overall error.

To confirm the findings of the present study future work could be performed on simulation data where hypothetical population data fitting a multilevel model could be created. A sample could then be drawn from this hypothetical population using complex survey sampling designs to compare the regression parameters derived from the four analytical approaches.

CONCLUSION

The four analytic strategies for complex survey data provide substantially different model estimates, standard errors and AIC. The lowest AIC was derived from the Multilevel Design Based analysis, which appears therefore to be the most appropriate approach to analyse complex survey data.

Previous research has studied separately the effect of complex survey design and hierarchical data structures on model estimates.

This study provides a systematic comparison of the combined effect of including consideration of the complex survey design and the hierarchical structure of the data on model estimates when the data are collected using a complex survey design with multilevel elements.

One limitation of this study is that the data were derived from a real life survey. Simulation data could yield a more accurate comparison of the methods, through generating a hypothetical population fitting the multilevel model.

Competing Interests

The authors declare that they have no competing interests.

Funding

There was no funding for this project

Acknowledgements

None

Authors' contributions

All the authors contributed in the manuscript and approved the final draft of the manuscript. MM contributed to conceptual framework, data analysis and manuscript writing; DR contributed to the conceptual framework and write up and TN contributed in write up.

Ethical Approval

The analysis described in this study utilised a public data set containing anonymous data collected with informed consent, made freely available for secondary data analysis by the World Health Organization. Therefore institutional ethics committee clearance was neither required nor sought for this analysis.

References

- Rabe-Hesketh S, Skrondal A. Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2006;169(4):805-27.
- Masood M, Reidpath DD. Multi-country health surveys: are the analyses misleading? *Curr Med Res Opin* 2014.
- Eaton DK, Kann L, Kinchen S, et al. Youth risk behavior surveillance - United States, 2011. *MMWR Surveill Summ* 2012;61(4):1-162.
- Kish L. *Survey sampling*: pp. xvi. 643. John Wiley & Sons: New York, 1965.
- Reidpath DD, Masood M, Allotey P. How much energy is locked in the USA? Alternative metrics for characterising the magnitude of overweight and obesity derived from BRFSS 2010 data. *International journal of public health* 2013.
- Leone T, James KS, Padmadas SS. The Burden of Maternal Health Care Expenditure in India: Multilevel Analysis of National Data. *Matern Child Health J* 2012.
- Masood M, Masood Y, Newton T. Impact of national income and inequality on sugar and caries relationship. *Caries Res* 2012;46(6):581-8.
- Kalton G. *Introduction to survey sampling*. Beverly Hills, London: Sage, 1983.
- Harling G, Msisha W, Subramanian SV. No association between HIV and intimate partner violence among women in 10 developing countries. *PLoS one* 2010;5(12):e14257.
- Merikangas KR, Jin R, He JP, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Archives of general psychiatry* 2011;68(3):241-51.
- Diez-Roux AV. Multilevel analysis in public health research. *Annu Rev Public Health* 2000;21:171-92.
- Ghosh S, Pahwa P, editors. *Using Design-based and Model-based method to study the relationship of asthma and smoking using longitudinal survey data*. ASA Biometrics Section, 2006. Alexandria: American Statistical Association.
- Reiter JP, Zanutto EL, Hunter LW. Analytical Modeling in Complex Surveys of Work Practices. *Industrial and Labor Relations Review* 2005;59(1):82-100.
- Subramanian SV, Perkins JM, Ozaltin E, Davey Smith G. Weight of nations: a socioeconomic analysis of women in low- to middle-income countries. *The American journal of clinical nutrition* 2011;93(2):413-21.
- Carle AC. Fitting multilevel models in complex survey data with design weights: Recommendations. *BMC Med Res Methodol* 2009;9:49.
- Antai D, Moradi T. Urban area disadvantage and under-5 mortality in Nigeria: the effect of rapid urbanization. *Environ Health Perspect* 2010;118(6):877-83.
- Goldhaber-Fiebert JD, Jeon CY, Cohen T, Murray MB. Diabetes mellitus and tuberculosis in countries with high tuberculosis burdens: individual risks and social determinants. *International journal of epidemiology* 2011;40(2):417-28.
- Masood M, Yusof N, Hassan MI, Jaafar N. Assessment of dental caries predictors in 6-year-old school children - results from 5-year retrospective cohort study. *BMC public health* 2012;12:989.
- Masood M, Yusof N, Hassan MI, Jaafar N. Longitudinal Study of Dental Caries Increment in Malaysian School Children: A 5-Year Cohort Study. *Asia-Pacific journal of public health / Asia-Pacific Academic Consortium for Public Health* 2012.
- WHO. *World Health Survey Reports 2006*. Available from: <http://www.who.int/healthinfo/survey/en/index.html>.
- Duncan C, Jones K, Moon G. Context, composition and heterogeneity: using multilevel models in health research. *Soc Sci Med* 1998;46(1):97-117.
- Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet* 2007;370(9590):851-8.
- WHO. *World Health Survey Reports 2003*. Available from: <http://www.who.int/healthinfo/survey/whsresults/en/index.html>.
- Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Medicine and science in sports and exercise* 2000;32(9):S498-504.
- Guthold R, Ono T, Strong KL, Chatterji S, Morabia A. Worldwide variability in physical inactivity a 51-country survey. *Am J Prev Med* 2008;34(6):486-94.
- Sarrafzadegan N, Kelishadi R, Esmailzadeh A, Mohammadifard N, Rabiei K, Roohafza H, et al. Do lifestyle interventions work in developing countries? Findings from the Isfahan Healthy Heart Program in the Islamic Republic of Iran. *Bull World Health Organ* 2009;87(1):39-50.
- Hamer M, Kivimaki M, Steptoe A. Longitudinal patterns in physical activity and sedentary behaviour from mid-life to early old age: a substudy of the Whitehall II cohort. *J Epidemiol Community Health* 2012;66(12):1110-5.
- Tesfaye F, Byass P, Wall S. Concurrent comparison of energy intake and expenditure among adults in Butajira District, Ethiopia. *Public Health Nutr* 2008;11(7):675-83.
- Hall JN, Moore S, Harper SB, Lynch JW. Global variability in fruit and vegetable consumption. *Am J Prev Med* 2009;36(5):402-9 e5.
- Muthén LK, Muthén BO. *Mplus User's Guide*. In: Sixth, editors. Los Angeles, CA: Muthén & Muthén, 2011.