

# Inappropriate Survey Design Analysis of the Korean National Health and Nutrition Examination Survey May Produce Biased Results

Yangho Kim<sup>1</sup>, Sunmin Park<sup>2</sup>, Nam-Soo Kim<sup>3</sup>, Byung-Kook Lee<sup>4</sup>

<sup>1</sup>Department of Occupational and Environmental Medicine, Ulsan University Hospital, University of Ulsan College of Medicine, Ulsan;

<sup>2</sup>Department of Food and Nutrition, Hoseo University, Asan; <sup>3</sup>Institute of Environmental & Occupational Medicine, Soonchunhyang University, Asan;

<sup>4</sup>Korea Industrial Health Association, Seoul, Korea

**Objectives:** The inherent nature of the Korean National Health and Nutrition Examination Survey (KNHANES) design requires special analysis by incorporating sample weights, stratification, and clustering not used in ordinary statistical procedures.

**Methods:** This study investigated the proportion of research papers that have used an appropriate statistical methodology out of the research papers analyzing the KNHANES cited in the PubMed online system from 2007 to 2012. We also compared differences in mean and regression estimates between the ordinary statistical data analyses without sampling weight and design-based data analyses using the KNHANES 2008 to 2010.

**Results:** Of the 247 research articles cited in PubMed, only 19.8% of all articles used survey design analysis, compared with 80.2% of articles that used ordinary statistical analysis, treating KNHANES data as if it were collected using a simple random sampling method. Means and standard errors differed between the ordinary statistical data analyses and design-based analyses, and the standard errors in the design-based analyses tended to be larger than those in the ordinary statistical data analyses.

**Conclusions:** Ignoring complex survey design can result in biased estimates and overstated significance levels. Sample weights, stratification, and clustering of the design must be incorporated into analyses to ensure the development of appropriate estimates and standard errors of these estimates.

**Key words:** Korean National Health and Nutrition Examination Survey, Complex survey design, Design-based analysis, Ordinary statistical data analysis, Multi-stage stratified sampling

Received: November 13, 2012 Accepted: March 15, 2013

**Corresponding author:** Byung-Kook Lee, MD

179 Hyoryeong-ro, Seocho-gu, Seoul 137-870, Korea

Tel: +82-2-585-1584, Fax: +82-2-585-1584

E-mail: bkleee@kiha21.or.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## INTRODUCTION

The Korean National Health and Nutrition Examination Survey (KNHANES) is one of a series of health-related programs sponsored by the Korean Centers for Disease Control and Prevention [1]. The KNHANES has been used to assess the health and nutritional status of the non-institutionalized civilian population in Korea since 1998. Since the survey data became available for official public use in Korea, research papers using

KNHANES data have become more common and have increased in numbers year by year [1]. Public access to the survey data has generated new momentum for research organizations, universities, and healthcare providers, including clinical medical researchers who had experienced difficulties in obtaining the raw data [2-4].

The KNHANES is a complex, stratified, multistage, probability-cluster survey of a representative sample of the non-institutionalized civilian population in Korea. The stratified multistage cluster sampling design creates non-independence among units, along with disproportionate sampling where some groups may be oversampled or a probability proportional to size has been applied [5,6]. This creates two primary challenges that must be addressed when analyzing complex survey data: 1) homogeneity that results from the non-simple random sample (i.e., non-independence) and 2) disproportionate sampling that results in unequal selection probabilities (e.g., oversampling or adjustment for non-response) [7,8]. Failure to address these issues results in incorrect estimation of standard errors, which often results in underestimation and an increased probability of a type I error [6,9].

Given the inherent nature of multiple complex survey designs, the KNHANES and the US NHANES provide specific guidelines for analysis and reporting of sample weights, so the stratification and clustering of the design must be incorporated into analyses to ensure appropriate estimates and standard errors. These guidelines warn that ignoring the complex design leads to biased estimates and overstated significance levels [10,11]. Regrettably, most of the published research articles using KNHANES data do not incorporate survey design analysis into their statistical analysis. Instead, most use regular standard statistical analysis, treating KNHANES data as if they were collected using a simple random sampling method. The objectives of this study were to: 1) explore the serious problem of misuse of statistical analyses using KNHANES data, 2) provide several examples of differing results when KNHANES data are analyzed using ordinary statistics and survey design statistics, and 3) recommend survey design statistics to calculate non-biased estimates for descriptive or analytical data analysis.

## METHODS

### Experimental Design and Data Collection

The entire KNHANES program was designed to assess the health and nutritional status of adults and children in Korea.

The KNHANES program began in 1998 and was conducted as a series of regular surveys in 2001 and 2005. In 2007, the survey became a continuous program that has focused on a variety of health and nutritional measurements to meet public need. The survey is conducted annually, using a rolling sampling design involving a complex, stratified, multistage, probability-cluster survey of a representative sample of the non-institutionalized civilian population in Korea.

Data analysis using the KNHANES requires special statistical analysis to handle the multistage complex sampling survey design, which yields data that differ from survey data collected using simple random sampling. Because of the inherent nature of the complex sampling survey design, the KNHANES provides specific guidelines for data analysis that should be incorporated into any survey design statistics developed using SAS or SPSS survey design programs, SUDAAN, and other relevant programs. This study analyzed the proportion of research papers that have used an appropriate statistical methodology since the KNHANES survey data became available for official public use. This study was restricted to the research papers cited in the PubMed online system from 2007 to 2012 (<http://www.ncbi.nlm.nih.gov/PubMed>). First, the research papers were classified into two groups: ordinary statistical data analysis papers treating the survey data as simple random sampling data; and design-based analysis papers that incorporated survey design into their analyses. Model-based analyses using multilevel or hierarchical data analysis were classified as special cases of sampling weighted data analysis papers incorporating survey design characteristics. To calculate the proportion of ordinary statistical data analyses and design-based analyses of the KNHANES, we searched the PubMed website for all available articles using the search terms "KNHANES or Korean National Health and Nutrition Examination Survey." Once collected, all the available articles were reviewed and classified into two groups (ordinary statistical data analysis without sampling weight and design-based analysis with sampling weight).

We also compared differences in mean and regression estimates between the ordinary statistical data analyses and design-based data analyses using KNHANES 2008 to 2010, which represented the second and third years of KNHANES IV (2007 to 2009) and the first year of KNHANES V (2010 to 2012). Detailed information on the design of the survey has been provided elsewhere [12,13]. Briefly, the survey consists of three components: a health interview survey, a health examination

survey, and a nutrition survey. We used two different populations to compare differences in estimates based on ordinary statistical data analyses and design-based analyses. The first study population was restricted to female KNHANES 2008 to 2009 participants who completed the health examination survey, who had a bone mineral density measurement, and who were menopausal (including surgically-induced menopause) ( $n=567$ ). The second study population was restricted to KNHANES 2008 to 2010 participants  $\geq 20$  years of age who had completed the health examination survey, including having their blood levels of lead, cadmium, and mercury measured ( $n=5919$ ).

Information on age, education, smoking history, and alcohol intake was collected during the health interview. Height and weight were measured with the participants wearing light clothing and no shoes. The body mass index was calculated as weight in kilograms divided by the square of the height in meters. Residence area was categorized into urban areas (administrative divisions of a city) and rural areas (areas not classified as administrative divisions of a city). Detailed information on classification variables has been provided elsewhere [12,13].

### Measurement of Blood Pressure

Blood pressure was measured while subjects were in a sitting position following a 5-minute rest period. The blood pressure was measured on three occasions with a mercury sphygmomanometer on the right arm and averaged for a final blood pressure reading. Hypertension was defined as diastolic blood pressure  $\geq 90$  mmHg or systolic blood pressure  $\geq 140$  mmHg. Prehypertension was defined as a diastolic pressure  $> 80$  mmHg or systolic blood pressure  $> 120$  mmHg.

### Serum Creatinine and Hemoglobin Measurements

Serum creatinine was measured by the kinetic Jaffe method using an autoanalyzer (model 7600; Hitachi, Tokyo, Japan). Blood hemoglobin was measured with an XE-2100D (Sysmex, Tokyo, Japan).

### Measurement of Bone Mineral Density

Bone mineral density was measured in the femur and lumbar spine by dual-energy X-ray absorptiometry at the health examination site using a QDR Discovery (formerly QDR 4500A) fan beam densitometer (Hologic, Inc., Bedford, MA, USA) [14]. Considering the available KNHANES data, we used the total fe-

mur T-score for the analysis.

### Determination of Lead, Cadmium, and Mercury Levels in Whole Blood

To assess the levels of heavy metals in whole blood, 3-mL blood samples were drawn into standard commercial evacuated tubes containing sodium heparin (Vacutainer). Blood lead and cadmium were measured using graphite-furnace atomic absorption spectrometry with Zeeman background correction (Perkin Elmer AAnalyst 600; Perkin Elmer, Turku, Finland). Levels of blood mercury were measured using a gold-amalgam collection method with a DMA 80 (Milestone, Bergamo, Italy). The details of blood metal analyses have been provided elsewhere [12].

### Statistical Analysis

Statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC, USA) and SUDAAN release 10.1 (Research Triangle Institute, Research Triangle Park, NC, USA), a software package that incorporates sample weights and adjusts analyses for the complex sample design of the survey. Survey sample weights were used in all analyses to produce estimates that were representative of the non-institutionalized civilian Korean population.

The levels of lead, cadmium, and mercury in the blood were log-transformed because their distributions were positively skewed, and the un-weighted and weighted geometric means (GM) with 95% confidence intervals (CI) were calculated by gender using the Proc Descript function in SUDAAN.

Odds ratios (ORs) and 95% CI for hypertension and prehypertension were calculated for log-transformed blood lead, cadmium, and mercury (base 2) levels as a continuous variable after covariate adjustment using Proc Rlogist with un-weighted and weighted statistical procedures. Covariates were gender, age, residence area, education level, smoking and drinking status, hemoglobin, body mass index, and serum creatinine.

Finally, the ORs and 95% CI values for osteoporosis were calculated for log-transformed blood lead, cadmium, and mercury (base 2) levels as continuous variables and quartiles of blood mercury in postmenopausal women after covariate adjustment using Proc Rlogist with un-weighted ordinary statistical data analyses and design-based weighted statistical procedures.

## RESULTS

During the period between 2007 and June 2012, a total of 256 relevant research papers were cited in PubMed. Table 1 lists the proportions that used ordinary statistical data analyses and design-based weighted statistical analyses. The number of relevant articles published increased steadily by year, and increased more rapidly in recent years. Throughout the 5.5 year time period, 19.8% of the articles used design-based data analyses.

Table 2 lists the differences in point estimates and standard errors of means between the un-weighted and weighted analyses by gender and age group. Not only did the means differ between the two analyses, but the standard errors in the weighted analyses also tended to be larger compared to those in the un-weighted analyses.

ORs and 95% CIs for hypertension and prehypertension

**Table 1.** Number (%) of research articles published in peer-reviewed journals cited in PubMed using Korean National Health and Nutrition Examination Survey data during 2007 to 2012<sup>1</sup>

Published year	No. of research articles (%)		
	Ordinary statistical analysis	Design-based analysis	Total
2007	12 (92.3)	1 (7.7)	13
2008	15 (88.2)	2 (11.8)	17
2009	24 (88.8)	3 (11.2)	27
2010	40 (87.0)	6 (13.0)	46
2011	58 (76.3)	18 (23.7)	76
2012 <sup>1</sup>	57 (73.1)	21 (26.9)	78
Total	206 (80.2)	51 (19.8)	257

<sup>1</sup>Until the end of June.

**Table 2.** Comparison of geometric mean  $\pm$  standard error of blood metals between ordinary statistical data analysis and design-based analysis by gender, age groups in Korean adult population

Classification	n	Blood lead ( $\mu\text{g/dL}$ )		Blood cadmium ( $\mu\text{g/L}$ )		Blood mercury ( $\mu\text{g/L}$ )	
		Ordinary statistical analysis	Design-based analysis	Ordinary statistical analysis	Design-based analysis	Ordinary statistical analysis	Design-based analysis
Total	5919	2.310 $\pm$ 0.012	2.288 $\pm$ 0.015	0.969 $\pm$ 0.007	0.967 $\pm$ 0.010	4.352 $\pm$ 0.033	4.303 $\pm$ 0.052
Gender	Male	2.735 $\pm$ 0.019	2.700 $\pm$ 0.021	0.888 $\pm$ 0.010	0.886 $\pm$ 0.012	5.205 $\pm$ 0.056	5.173 $\pm$ 0.080
	Female	1.951 $\pm$ 0.014	1.943 $\pm$ 0.018	1.057 $\pm$ 0.011	1.055 $\pm$ 0.014	3.639 $\pm$ 0.034	3.587 $\pm$ 0.047
Age (y)	20-29	1.820 $\pm$ 0.020	1.816 $\pm$ 0.024	0.596 $\pm$ 0.011	0.597 $\pm$ 0.014	3.614 $\pm$ 0.053	3.652 $\pm$ 0.067
	30-39	2.129 $\pm$ 0.024	2.152 $\pm$ 0.025	0.874 $\pm$ 0.014	0.876 $\pm$ 0.018	4.313 $\pm$ 0.068	4.287 $\pm$ 0.076
	40-49	2.402 $\pm$ 0.029	2.421 $\pm$ 0.034	1.084 $\pm$ 0.016	1.078 $\pm$ 0.017	4.846 $\pm$ 0.081	4.859 $\pm$ 0.093
	50-59	2.690 $\pm$ 0.030	2.672 $\pm$ 0.030	1.218 $\pm$ 0.019	1.209 $\pm$ 0.022	5.022 $\pm$ 0.085	5.033 $\pm$ 0.100
	>60	1204	2.595 $\pm$ 0.032	2.502 $\pm$ 0.037	1.212 $\pm$ 0.017	1.243 $\pm$ 0.022	4.084 $\pm$ 0.074

were calculated for log-transformed blood metal levels after adjusting for covariates (Table 3). The covariates for models 1 and 2 were gender, age, residence area, education level, smoking and drinking status, hemoglobin, and serum creatinine. In a logistic regression analysis using log-transformed blood cadmium as a continuous independent variable after covariate adjustment, the ORs and 95% CIs for hypertension with a doubling of blood cadmium were 1.240 (1.123 to 1.367) and 1.249 (1.117 to 1.398) in the ordinary statistical data analyses and design-based data analyses, respectively. However, a logistic regression analysis using log-transformed blood cadmium as a continuous independent variable after covariate adjustment revealed that the ORs and 95% CIs for pre-hypertension with a doubling of blood cadmium were 1.295 (1.178 to 1.423) and 1.279 (1.147 to 1.426) in ordinary statistical data analyses and

**Table 3.** Comparison of logistic regression analysis<sup>1</sup> of blood metals on hypertensive status in adults between ordinary statistical data analysis and design-based analysis during 2008 to 2010

Independent variable	Ordinary statistical analysis	Design-based analysis
	Odds ratio (95% confidence interval) <sup>2</sup>	
Model 1. Dependent variable: hypertension		
Log blood lead ( $\mu\text{g/dL}$ )	0.983 (0.859, 1.124)	0.999 (0.854, 1.170)
Log blood cadmium ( $\mu\text{g/L}$ )	1.240 (1.123, 1.369)	1.249 (1.117, 1.398)
Log blood mercury ( $\mu\text{g/L}$ )	0.963 (0.882, 1.051)	0.968 (0.878, 1.068)
Model 2. Dependent variable: pre-hypertension		
Log blood lead ( $\mu\text{g/dL}$ )	1.137 (0.995, 1.298)	1.173 (1.001, 1.375)
Log blood cadmium ( $\mu\text{g/L}$ )	1.295 (1.178, 1.423)	1.279 (1.147, 1.426)
Log blood mercury ( $\mu\text{g/L}$ )	0.902 (0.827, 0.984)	0.924 (0.833, 1.023)

<sup>1</sup>Adjusted for sex, age, education, body mass index, smoking and drinking status, hemoglobin and serum creatinine.

<sup>2</sup>For doubling of blood metals.

**Table 4.** Comparison of logistic regression analysis<sup>1</sup> of blood metals on osteoporosis in menopausal women between ordinary statistical data analysis and design-based analysis

Independent variable	Ordinary statistical analysis	Design-based analysis
	Odds ratio (95% confidence interval)	
Model 1. Continuous blood metals		
Log blood lead (µg/dL)	1.771 (1.029, 3.047) <sup>2</sup>	2.371 (1.307, 4.301) <sup>2</sup>
Log blood cadmium (µg/L)	0.899 (0.575, 1.406) <sup>2</sup>	0.952 (0.584, 1.552) <sup>2</sup>
Log blood mercury (µg/L)	0.488 (0.320, 0.743) <sup>2</sup>	0.628 (0.376, 1.048) <sup>2</sup>
Model 2. Categorized blood mercury		
≤ 2.504	1 (reference)	1 (reference)
Blood mercury quartiles (µg/L) >2.504 to 3.487	0.863 (0.475, 1.569)	1.122 (0.537, 2.345)
>3.487 to 5.001	0.611 (0.342, 1.093)	0.944 (0.475, 1.877)
>5.001	0.450 (0.251, 0.808)	0.575 (0.280, 1.181)

<sup>1</sup>Adjusted for age, education, body mass index, smoking and drinking status, hemoglobin, and serum creatinine.

<sup>2</sup>For doubling of blood metals.

design-based analyses, respectively. While log-transformed blood lead did not exhibit any significant association with hypertension or prehypertension in logistic regression analyses regardless of type of analysis, log-transformed blood mercury exhibited a significant protective effect on prehypertension in un-weighted ordinary logistic regression analyses, while weighted design-based analyses did not reveal any significant associations between blood mercury and prehypertension.

Finally, Table 4 lists logistic regression analyses of blood metals on osteoporosis in menopausal women by ordinary statistical data analyses and design-based analyses. In the logistic regression analyses using log-transformed metals as a continuous independent variable after covariate adjustment, the ORs for osteoporosis with doubled blood lead were statistically significant between two types of analyses. The doubled blood mercury had a statistically significant OR (0.488: 0.320 to 0.743) in ordinary statistical data analyses, but no association was observed between blood mercury and osteoporosis in the design-based analysis. Similar findings were observed in logistic regression analyses using quartiles of blood mercury as an independent variable for osteoporosis. The OR (95% CI) of the highest quartile of blood mercury in ordinary statistical data analyses was 0.450 (0.251 to 0.808) with statistical signifi-

cance; in design-based analyses, it was 0.575 (0.281 to 1.181) with no statistical significance.

## DISCUSSION

Large-scale health surveys including the KNHANES and the US NHANES are increasingly being used to investigate associations between risk factors of diseases, nutritional status, and health-related outcomes [15-17]. Because these surveys record many variables for a large number of individuals, they enable many hypotheses to be addressed, some of which may not even be formulated until years after the surveys have been completed. Additionally, these surveys are frequently population-based, thereby avoiding the potential biases of hospital-based studies [3].

Generally, these large-scale health surveys use complex probability procedures, such as stratification, multiple stages of cluster selection, and oversampling to obtain a representative sample of the target population [5,18,19]. Because of these sampling procedures, analyses of these data using traditional statistical software (such as SPSS or SAS) that use ordinary and generalized least squares estimation techniques tend to result in an underestimated standard error, inappropriate confidence intervals, and misleading tests of significance [20]. Until a decade ago, many analysts of data collected in sample surveys felt comfortable using standard procedures in statistical software packages to analyze the data. These standard procedures were based on the assumption that the data being analyzed emerged from a simple random sample of the population of interest, where all sample respondents were randomly selected without replacement from the population, and all respondents had equal probability of being included in the sample [21]. During the last several decades, many survey data analysts have expressed their concerns that survey data were being analyzed using standard statistical procedures instead of specific procedures based on survey design [6,8].

In Korea, many research articles using KNHANES data have been published in peer-reviewed journals since the Korea Centers for Disease Control and Prevention released the KNHANES dataset to the research community in 2007, and the number of research articles using KNHANES data has increased remarkably over the last six years. The KNHANES provides a good source of data for studying the relationship between nutritional status, food habits, environmental factors, and health-related outcomes. Even though the proportion of weighted

research articles has increased from 7.7% in 2007 to 26.9% in 2012, the majority of peer-reviewed research articles using KNHANES data cited in PubMed did not use appropriate analyses. Standard statistical analyses that do not use weighted multiple complex survey data generally 1) yield biased point estimates of population parameters, 2) underestimate the standard error for point estimates, 3) produce overly-narrow confidence intervals on population parameters, and 4) yield tests of significance that are overly likely to reject the null hypothesis because standard error or variability of statistics is generally underestimated, resulting in more frequent type I errors [3,6,20,22].

This study compared ordinary statistical data analyses without sampling weight and design-based (data analyzed using special survey design statistical procedures) GMs and their standard errors by gender and age group (Table 2). As anticipated, the GMs differed between the two groups, and standard errors of data based on weighted analyses were larger than those based on un-weighted data analyses. Clearly, ordinary statistical un-weighted GMs and standard errors are not the same as design-based GMs and standard errors, and ordinary un-weighted GMs and standard errors cannot be used as an alternative to weighted GMs in analyses of survey data [23].

The failure of data analysts to use appropriate statistical software procedures for unbiased analyses of survey data essentially negates the considerable money and effort needed to collect the data, and can result in scientific publications that present skewed pictures of populations of research interest [21]. Un-weighted analyses of complex multi-stage cluster sampling data tend to have a narrow confidence interval, which creates biased and skewed results and generates a type I error. This study illustrated that un-weighted data analyses generated incorrect conclusions, for example that blood mercury is associated with a protective effect on prehypertension (Table 3, model 2) in an adult population and is associated with having osteoporosis among menopausal women (Table 4); these results were due to the narrow confidence intervals.

The cross-sectional nature of most survey designs like KNHANES makes it difficult to draw causal associations between health-related outcomes and certain independent variables, including heavy metals such as blood mercury. The survey data can only provide statistical associations through conservative and appropriate statistical analysis and plausible scientific justification. In this study, blood mercury was a significant predictor for the prevention of osteoporosis among meno-

pausal women using an un-weighted analysis. This finding reconfirmed a recent publication regarding this topic, which claimed that high blood mercury has a protective effect for osteoporosis among Korean menopausal women [24]. The authors analyzed 481 postmenopausal women who participated in KNHANES 2008 and underwent bone mineral density measurement. We reanalyzed the data on these subjects using the same analysis format, but were unable to recreate the results due to the total number of menopausal women who were eligible for their inclusion criteria (menopausal and measurement of bone mineral density). In 2008, KNHANES only measured the bone mineral density of the participants who were enrolled in the later part of the survey (about 50% of all participants). Therefore, we analyzed two years of data (2008 to 2009) including 567 postmenopausal women and obtained the same results of protective effect of mercury for osteoporosis using an ordinary statistical data analysis without sampling weight, but failed to obtain statistical significance in design-based statistical analysis. The KNHANES measured total blood mercury, which included mainly organic mercury with a small portion of inorganic mercury. Because most organic mercury comes from foods including various fishes, high levels of fish consumption may be beneficial from a nutritional standpoint, but may also increase blood mercury levels. Therefore, it is unclear whether high blood mercury has a protective effect on osteoporosis. It is possible that the blood mercury level is simply another unknown index of fish consumption.

Table 3, model 1, reveals no differences between the ordinary statistical data analysis and design-based analysis. This is the most desirable outcome for an un-weighted analysis of survey data. In the past, it was common to use ordinary statistical procedures due to the lack of availability of statistical software capable of more complex analyses. However, during the last two decades, suitable software has become available and is easily incorporated into ordinary statistical packages such as SAS, SPSS, and STATA. Therefore, analysts cannot claim the scarcity of software tools as a justification for using ordinary statistical packages, which treat the sample design as simple random sampling [6,8].

Two basic approaches are generally available for the analysis of survey design data: design- and model-based approaches [6,25]. Design-based analysis incorporates all the characteristics of multiple complex sampling designs (weighting, stratification, clustering, and post-stratification) and is the standard analysis method in the majority of national survey designs

[26-28]. In contrast, model-based analysis does not use sample weighting, but instead uses a standard statistical method that considers weighting, stratification, and clustering in the modeling. As long as the model describes every possible observation to explain relationships accurately, model-based analysis can ignore the complexity of survey design, even for weighted correlated data. Unfortunately, theoretically derived models known to hold for all observations often do not exist for survey situations, unlike those derived from the physical sciences [6,8].

Over the past 10 years, many researchers in the health sciences have become interested in conducting secondary analyses using data from complex sample surveys. These analyses are descriptive, analytical, hypothesis-generating, or model-building. Sample survey statisticians are aware that specialized software should be used to analyze complex sample survey data, particularly when analyses are descriptive or analytical and the survey design includes clustering [20]. Until recently, some analysts in Korea thought that un-weighted ordinary statistical data analysis was the same as model-based analysis, and did not require sample weighting [29-31]. The lack of awareness about the necessity of using weighted analysis in survey design analysis resulted in biased estimates, generating model misspecification and type I errors in the majority of published articles using the KNHANES in Korea.

Given the inherent nature of multiple complex survey design, the NHANES and KNHANES provide specific guidelines for analysis and reporting of sample weights, and state that the stratification and clustering of the design must be incorporated into the analysis to obtain accurate estimates and standard errors. The guidelines warn that ignoring the complex design will lead to biased estimates and overstated significance levels [10,11]. Hahs-Vaughn et al. [32] reviewed methodological considerations in using complex survey data from the Head Start Family and Child Experiences Survey. They warned that failing to address survey design issues resulted in underestimated standard errors and biased parameters. To avoid methodological problems in survey design data analyses, they recommended reporting practices for authors, editorial board members, and editors to assist in ensuring transparency and greater ease in replication when using complex sample data. Their recommendations for authors were as follows: 1) specify the survey weight applied during the analyses to correct for disproportionate sampling; 2) if a design-based approach is followed, specify how variances have been esti-

mated; 3) if the complex sample design is not addressed in the analyses, provide information to help the reader understand why that choice was appropriate and how it affects interpretation of the results.

The recommendations for authors should be implemented as part of the journal's manuscript submission guidelines. Their recommendations for journal editors were as follows: 1) require authors to report the survey weights and variance estimation method when their manuscript includes analyses of complex sample data; 2) providing resources to the editorial board, such as a manuscript or other primer [15] that provides information about complex survey data and can be a helpful tool when reviewing manuscripts that have analyzed this type of data. The analysis guidelines of data providers such as NHANES and KNHANES that incorporate a multiple complex sampling design must be followed; if not, any use of other analysis methods must be justified scientifically, as with model-based analysis, which uses standard statistical methods that consider survey design complexity, such as stratification, clustering, and non-response.

Although two basic approaches are generally available, it should be noted that there may be three possible statistical methods for handling the data of a multistage complex sampling survey design: the first is the un-weighted ordinary statistical method that takes into account neither complex sample survey design nor sample weights; the second is the weighted ordinary statistical method without consideration of the complex sample survey design; the last is incorporating sample weights and properly taking into consideration complex sample survey design. However, we were not able to identify the second method due to the lack of information on statistical methods provided by the reviewed studies. Instead, we classified both the first and second methods as ordinary statistical analysis, and the last one as design-based analysis. This is a limitation of the present study.

The emerging new era of data analysis using the KNHANES promises a bright future for health and nutrition-related scientific research in Korea, which will eventually increase our knowledge and contribute to a better understanding of human health. As the survey data accumulate annually, the richness of the data for more in-depth analysis becomes increasingly apparent. The demand for taking fuller advantage of such a rich data source is increasing, but it is vital to ensure sound scientific justification and awareness of the limitations of the survey design data. Thus, appropriate statistical analysis that in-

corporates the sample survey design needs to be mandatory.

## CONFLICT OF INTEREST

The authors have no conflicts of interest with the material presented in this paper.

## REFERENCES

1. Korea Centers for Disease Control and Prevention. Second workshop for the data analysis of the Korea National Health & Nutrition Examination Survey and online survey of adolescent health behavior. Cheongwon: Korea Centers for Disease Control and Prevention; 2012, p. 1-60 (Korean).
2. Jacobson AF, Hamilton P, Galloway J. Obtaining and evaluating data sets for secondary analysis in nursing research. *West J Nurs Res* 1993;15(4):483-494.
3. Korn EL, Graubard BI. Analysis of large health surveys: accounting for the sampling design. *J R Stat Soc Ser A Stat Soc* 1995;158(2):263-295.
4. Magee T, Lee SM, Giuliano KK, Munro B. Generating new knowledge from existing data: the use of large data sets for nursing research. *Nurs Res* 2006;55(2 Suppl):S50-S56.
5. Levy PS, Lemeshow S. Sampling of populations: methods and applications. 3rd ed. New York: Wiley-Interscience Publication; 1999, p. 333-363.
6. Lee ES, Forthofer RN. Analyzing complex survey data. 2nd ed. London: SAGA Publications; 2006, p. 11-20.
7. Wheeler DC, VanHorn JE, Paskett ED. A comparison of design-based and model based analysis of sample survey in geography. Atlanta: Rollings School of Public Health, Emory University; 2007, p. 1-34.
8. Lohr SL. Sampling: design and analysis. 2nd ed. Boston: Brooks/Cole, Cengage Learning; 2010, p. 281-320.
9. Korn EL, Graubard BI. Analysis of health surveys. New York: Wiley-Interscience Publication; 1999, p. 8-50.
10. National Center for Health Statistics. The National Health and Nutrition Examination Survey (NHANES): analytic and reporting guidelines. Atlanta: Centers for Disease Control and Prevention; 2012, p. 1-14.
11. Korea Centers for Disease Control and Prevention. Guideline for the evaluation of the fourth Korea National Health and Nutrition Examination Survey. Cheonwon: Korea Centers for Disease Control Prevention; 2009, p. 19-34 (Korean)
12. Lee BK, Kim Y. Iron deficiency is associated with increased levels of blood cadmium in the Korean general population: analysis of 2008-2009 Korean National Health and Nutrition Examination Survey data. *Environ Res* 2012;112:155-163.
13. Park S, Lee BK. Vitamin D deficiency is an independent risk factor for cardiovascular disease in Koreans aged  $\geq 50$  years: results from the Korean National Health and Nutrition Examination Survey. *Nutr Res Pract* 2012;6(2):162-168.
14. Hong S, Oh HJ, Choi H, Kim JG, Lim SK, Kim EK, et al. Characteristics of body fat, body fat percentage and other body composition for Koreans from KNHANES IV. *J Korean Med Sci* 2011;26(12):1599-1605.
15. Hahs-Vaughn DL. A primer for using and understanding weights with national datasets. *J Exp Educ* 2005;73(3):221-248.
16. National Research Council. Methods for analysis of complex surveys. In: National Research Council. Understanding American agriculture: challenges for the agricultural resource management survey. Washington, DC: National Academies Press; 2007, p. 127-148.
17. Burden S, Probst Y, Steel DG, Tapsell LC. The impact of complex survey design on prevalence estimates of intakes of food groups in the Australian National Children's Nutrition and Physical Activity Survey. New South Wales: Centre for Statistical and Survey Methodology, University of Wollongong; 2010, p. 1-21.
18. Kalton G. Models in the practice of survey sampling (revisited). *J Off Stat* 2002;18(2):129-154.
19. Rodgers-Farmer AY, Davis D. Analyzing complex survey data. *Soc Work Res* 2001;25(3):185-192.
20. Brogan D. Software for sample survey data, misuse of standard packages. In: Armitage P, Colton T, editors. Encyclopedia of biostatistics. 2nd ed. Estados Unidos: John Wiley & Sons; 2005, p. 5057-5064.
21. West BT. Statistical and methodological issues in the analysis of complex sample survey data: practical guidance for trauma researchers. *J Trauma Stress* 2008;21(5):440-447.
22. Lemeshow S, Letenneur L, Dartigues JF, Lafont S, Orgogozo JM, Commenges D. Illustration of analysis taking into account complex survey considerations: the association between wine consumption and dementia in the PAQUID study. *Am J Epidemiol* 1998;148(3):298-306.
23. Korn EL, Graubard BI. Examples of differing weighted and unweighted estimates from a sample survey. *Am Stat* 1995;49(3): 291-295.
24. Cho GJ, Park HT, Shin JH, Hur JY, Kim SH, Lee KW, et al. The relationship between blood mercury level and osteoporosis in



- postmenopausal women. *Menopause* 2012;19(5):576-581.
25. Heeringa SG, West BT, Berglund PA. *Applied survey data analysis*. London: CRC Press; 2005, p. 13-50.
  26. Park I, Lee H. Design effects for the weighted mean and total estimators under complex survey sampling. *Surv Methodol* 2004;30(2):183-193.
  27. Centers for Disease Control and Prevention. Specifying weighting parameters; 2011 [cited 2013 Mar 1]. Available from: <http://www.cdc.gov/nchs/tutorials/nhanes/SurveyDesign/Weighting/intro.htm>.
  28. Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey, 2007-2008; 2012 [cited 2013 Mar 1]. Available from: [http://www.cdc.gov/nchs/data/nhanes/nhanes\\_07\\_08/overviewbrochure\\_0708.pdf](http://www.cdc.gov/nchs/data/nhanes/nhanes_07_08/overviewbrochure_0708.pdf).
  29. Lee BK, Kim Y. Proper sample-weighted data analysis is required to confirm the association between blood cadmium concentration and lipid profile in Korean adults: re: Kim K. Blood cadmium concentration and lipid profile in Korean adults. *Environ. Res.* 2011. doi:10.1016/j.envres.2011.12.008. *Environ Res* 2012;115:66-67.
  30. Lee BK. Without the application of appropriate statistical methods, it is difficult to reach the proper conclusion. *Menopause* 2012;19(6):710.
  31. Platt RW, Harper SB. Survey data with sampling weights: is there a "best" approach? *Environ Res* 2013;120:143-144.
  32. Hahs-Vaughn DL, McWayne CM, Bulotsky-Shearer RJ, Wen X, Faria AM. Methodological considerations in using complex survey data: an applied example with the Head Start Family and Child Experiences Survey. *Eval Rev* 2011;35(3):269-303.