

# The Korea National Health and Nutrition Examination Survey as a Primary Data Source

Editorial

**Hyun Ah Park**

Department of Family Medicine, Inje University Seoul Paik Hospital, Inje University College of Medicine, Seoul, Korea

The Korea National Health and Nutrition Examination Survey (KNHANES) is a population-based cross-sectional survey designed to assess the health related behavior, health condition, and nutritional state of Koreans (<http://knhanes.cdc.go.kr/>). It was conducted as a tri-annual survey for the first 3 cycles, implemented in 1998, 2001, and 2005. Beginning with the 4th cycle of 2007, it was converted to an annual survey. The 2011 data was recently opened to the public, and is free for all researchers who want to use it.

The KNHANES provides a rich source of data which are easy to access and can be quickly obtained. Its ethical problems are minimal and there are no adverse effects in conducting the study. Using the KNHANES saves time, money, and personnel that would otherwise be spent collecting data, and provides a larger and higher-quality database beyond the capacity of any one individual researcher. Further, because the study participants are representative of the Korean population, the results have external validity.

Therefore, this data is helpful to young investigators. In fact, the portion of submitted articles to the Korean Journal of Family Medicine (KJFM) using the KNHANES as the primary source of data is significant. Specifically, one article in 2007, three articles in 2008, four articles in 2009, three articles in 2010, five articles in 2011, and five articles in 2012 used the KNHANES data as their primary data source, totaling 21 articles (6.2%) among 338 original articles during the previous five years.

However, there are several points investigators should keep in mind when using the KNHANES data.

First, the KNHANES data are open to the public and there is no systematic process to control the research topic assignment. Many researchers might investigate the same topic simultaneously, especially popular topics like obesity and metabolic syndrome. Before deciding on the study topic, a thorough search of the database like Medline, EMBASE, and the abstracts of related conferences are needed. Once the topic is decided upon, the

article should be submitted as soon as possible. Otherwise someone else may publish a similar study from the same dataset before you do.

Second, investigators must be familiar with the huge and complex data structure and study design of the KNHANES. Non-response and multi-stage probability sampling should be taken into account by using survey statistics to estimate the data of the whole Korean population. Despite this, lots of studies submitted to KJFM using the KNHANES do not apply sampling weight in their analysis. In such a case the estimated association might be biased and revision is required.

Third, the KNHANES is a secondary data source like the Korea Youth Risk Behavior Web-based Survey and the Community Health Survey, and has the same limitations that they have.<sup>1)</sup> The main purpose of the KNHANES is to produce national statistics, not to answer a specific research question. Therefore, the particular information the investigator wants may not be collected. For these reasons, the dataset should be examined carefully to confirm that it includes the necessary data. There also exists a potential for errors or mistakes in the data, even with its rigorous documents control system. Missing data is also a problem in the KNHANES.<sup>2)</sup>

## CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

## REFERENCES

1. Coyer SM, Gallo AM. Secondary analysis of data. *J Pediatr Health Care* 2005;19:60-3.
2. Park HA. Rate of missing socioeconomic factors in the 4th KNHANES. *Korean J Fam Med* 2012;33:406-9.