

Development of a dish-based, semi-quantitative FFQ for the Korean diet and cancer research using a database approach

Min Kyung Park^{1*}, Dong Woo Kim¹, Jeongseon Kim², Sohee Park³, Hyojee Joung⁴, Won O. Song⁵ and Hee Young Paik^{1,6}

¹Department of Food and Nutrition, Seoul National University, Seoul, Republic of Korea

²Cancer Epidemiology Branch, Research Institute, National Cancer Center, Goyang-si, Gyeonggi-do, Republic of Korea

³Cancer Biostatistics Branch, Division of Cancer Registration and Surveillance, National Cancer Center, Goyang-si, Gyeonggi-do, Republic of Korea

⁴Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea

⁵Department of Food Science and Human Nutrition, Michigan State University, East Lansing, MI, USA

⁶Research Institute of Human Ecology, Seoul National University, Seoul, Republic of Korea

(Received 1 March 2010 – Revised 8 September 2010 – Accepted 13 September 2010 – First published online 24 November 2010)

Abstract

We used a database approach in developing a dish-based, semi-quantitative FFQ for Korean diet and cancer research. Cancer-related dietary factors (CRDF) recognised in the scientific community and dietary intake data from the 2001 Korean National Health and Nutrition Examination Survey and the 2002 Korean National Nutrition Survey by Season were used. The list of dishes (n 993) was those reported to be consumed by individuals over 30 years of age during all four seasons. The resulting 112-dish list was selected using contribution analyses and variability analyses to detect between-person variation for CRDF and non-CRDF nutrients. Variations of each dish were grouped into one dish for the final list of 112 dishes, which were then linked to the nutrient database. The final 112 dish items consisted of nine Korean staple dishes, including rice and noodles, twenty-five soups and stews, fifty-four side dishes, nine beverages, nine fruit dishes and six alcoholic beverages. The percentage coverages of energy, protein, fat, carbohydrate and alcohol intake in the selected 112 dishes were 82.4, 76.4, 68.9, 86.0 and 99.8%, respectively. Dietary exposure to cancer-related Korean dietary factors can be assessed by this new dish-based, semi-quantitative FFQ. This new instrument can calculate the intake of CRDF along with non-CRDF nutrient intake for cancer research.

Key words: Cancer: Database approach: Dish-based FFQ

In Korea, mortality due to cancer has been among the highest causes of death since 1983⁽¹⁾ and has been increasing steadily⁽²⁾. The medical cost associated with each cancer case is higher than that of other types of diseases, and a significant proportion of cancer incidence is associated with dietary and lifestyle risk factors. Consequently, cancer research has become critically important for the whole nation.

Doll & Peto⁽³⁾ estimated that changes in dietary habits alone could reduce cancer incidence by a third. The World Cancer Research Fund/American Institute for Cancer Research reported that 30–40% of carcinogenesis could be prevented by optimal nutrition, regular physical activity and the prevention of obesity⁽⁴⁾. In 2007, ten dietary and physical activity recommendations for cancer prevention were established⁽⁵⁾.

Cancer and dietary risk factors have extensively been studied with various diet survey tools. A multiethnic cohort study in Hawaii and Los Angeles⁽⁶⁾ used a newly developed quantitative FFQ considering traditional dishes for each ethnic group⁽⁷⁾. The FFQ was calibrated for each ethnic and sex group. In the Women's Health Initiative, the relationship between low-fat diets and colorectal and breast cancer^(8,9) was investigated by a semi-quantitative FFQ including 122 food items⁽¹⁰⁾ and the minority population FFQ was used in the Women's Health Trial Feasibility Study⁽¹¹⁾. In the European Prospective Investigation into Cancer and Nutrition project, dietary survey data and nutrient databases of each country, in cohorts of ten European nations, were integrated into one dataset for an accurate dietary survey for the entire study. Then, FFQ appropriate for the dietary culture of each country were developed,

Abbreviations: CRDF, cancer-related dietary factors; KNHANES, Korean National Health and Nutrition Examination Survey.

* **Corresponding author:** M. K. Park, fax +82 2 880 8775, email anapmk30@gmail.com

validated and calibrated with each country's research results^(12–16). The Japan Collaborative Cohort Study used an FFQ consisting of thirty-five food items and a questionnaire that assessed the intake of alcohol and the preference for fatty and salty foods⁽¹⁷⁾, with portion size determined by a validation study⁽¹⁸⁾.

Recently, Korean cohorts have been established to study the relationship between diet and cancer^(19,20). Several small-scale surveys or laboratory studies have reported that traditional Korean diets are cancer-protective as a result of their emphasis on allium-containing vegetables such as garlic^(21,22). However, Korean diets have several undesirable attributes such as high salt and a high level of barbecued meat consumption. Furthermore, Korean cooking is characterised by various similar ingredients prepared by different approaches and mixed with many similar seasonings. Because cancer-related dietary factors (CRDF) are relevant to culture-specific cooking methods and ingredients, focusing on the consumption of prepared dishes, instead of food ingredients, was deemed appropriate for Korean diet-related cancer research. A few Korean FFQ for studying chronic diseases have been developed^(23,24), but no FFQ for cancer research is currently available. In the present study, we aimed to develop a dish-based, semi-quantitative FFQ for diet and cancer research using a database approach.

Methods

Selection of cancer-related dietary factors

Of the CRDF that have been widely accepted in the scientific community⁽⁵⁾, sixteen factors that are extractable from the Korean National Health and Nutrition Examination Survey (KNHANES) were selected for the present study. Specifically, those factors varied from the 'convincing' to the 'limited suggestive' categories from the World Cancer Research Fund/American Institute for Cancer Research's report⁽⁵⁾. Several suggested factors, such as β -carotene supplements, lycopene, folate, Se and vitamin D, were not included in the present study because these data were unavailable in the Korean food composition database.

Both allium-containing vegetables (sum of vegetables such as garlic, onion, leeks, scallions and chives) and garlic are noted separately as cancer-protective factors in the World Cancer Research Fund/American Institute for Cancer Research publication. Garlic, but not allium-containing vegetables, was included in developing the present instrument. When all dishes consumed by the subjects in the KNHANES were included in multivariate regression for variability analysis, the accumulated correlation coefficient (r^2) value of allium-containing vegetables did not reach 0.9. This indicates that if all of the dishes eaten by the subjects in the KNHANES are included in the final FFQ, allium vegetable intake cannot adequately

Table 1. Descriptions of subjects whose dietary intake data were included in the present study

(Numbers, percentages, mean values and standard deviations)

	All		Men		Women	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Total	6490	100	2989	46.1	3501	53.9
Distribution by age (years)						
30–49	3760	57.9	1796	60.0	1964	56.1
50–64	1658	25.6	774	25.9	884	25.2
65 +	1072	16.5	419	14.0	653	18.6
Age (years)						
Mean	49.0		48.1		49.8	
SD	13.8		13.1		14.4	

explain between-person variations. This phenomenon occurs because allium vegetables are added in very small amounts as flavourings in almost every Korean dish.

Sugar-containing foods, a dietary risk factor of colorectal cancer⁽⁵⁾, were not included in the present study. Sugar-containing foods in the Korean diet include snacks, cookies, cakes, breads, coffee from vending machines, ice cream and yogurt, in addition to added sugar and syrup

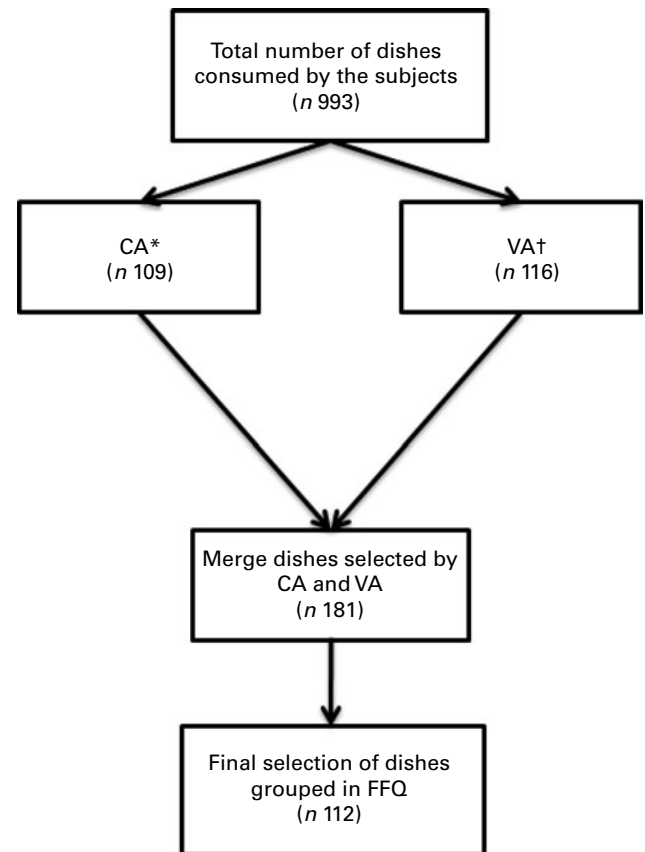


Fig. 1. Steps used to develop the 112-dish list from 993 dishes consumed by the subjects. * Contribution analysis (CA) is a statistical approach to identify dishes that explained each cancer-related dietary factor (CRDF) significantly. The dishes used in the final study are those contributed collectively over 50% for each risk factor. † Variability analysis (VA) is conducted by multiple regressions to identify dishes that differentiate individuals whose intakes of CRDF differ. The analysis was conducted until the cumulative sum of the square of correlation coefficient reached 0.9 for two or more CRDF.

in prepared foods. However, there was no publicly released database of sugar content in these dishes; therefore, sugar content could not be calculated. Subsequently, the sixteen CRDF selected for the present study were β -carotene, Ca, vitamin C, Na, retinol, red meat, alcohol, processed meats, fruits, vegetables (including garlic and carrot), garlic, carrots, milk, dairy food (including milk), beans and fish.

Determination of the dish list and portion size

Using a database approach, we determined the dish list and portion sizes using large-scale Korean dietary intake survey datasets: the KNHANES (2001) and Korean National Nutrition Survey by Season (2002), which are the most current seasonal survey datasets available in Korea. Both surveys utilised one 24 h recall. The purpose of the Korean National Nutrition Survey by Season was to assess the variations in dietary intake in terms of dishes, foods and nutrients among the four seasons. Subjects between the two surveys were intentionally overlapped. Subjects over 30 years old in the 2002 Korean National Nutrition Survey by Season numbered 1644, 1665 and 1558 people for the spring, summer and autumn surveys, respectively, with a mean of 1623 people. Therefore, 1623 people were selected randomly from a total of 5722 people of the same age group in the 2001 KNHANES winter survey. Those who participated in the survey over two seasons were regarded as different persons. Our working dataset was proved to be representative in terms of nutrient intake of the full dataset of the 2001 KNHANES using Student's *t* test (data not shown).

The total number of subjects (Table 1) was 6490, and this value was the total number of subjects for all the four seasons. These subjects had a mean age of 49 years, with men and women accounting for 46.1 and 53.9%, respectively, and the highest frequency age group was 30–49 years of age (57.9%), followed by 50–64 years of age (25.6%), and an over 65-year-old group (16.5%).

Dish list

Of the 993 dishes eaten by the subjects, the final dish list was selected from the results of the contribution analysis and from the results of the variability analysis of the effects of between-person variation (Fig. 1). Both analyses were performed for the intake of CRDF and non-CRDF nutrients, i.e. energy, carbohydrate, protein, fat, etc., which were also included for nutritional assessment purposes.

First, we tried to select the number of dishes to ensure that the total contribution to each CRDF and non-CRDF nutrient would be more than 90%. To fulfil these criteria, we had to choose over 400 dishes; therefore, we chose dishes collectively contributing more than 50% to each risk factor used in the final study. Variability analyses were then performed using multiple regression analyses to select dishes that explained between-person variations in the intake of two or more CRDF. All dishes eaten by subjects (993 dishes) were included in one model as independent variables for each risk factor, and independent variables were selected by a stepwise method for each model. Through the processes, dishes with over 90% accumulated square of r^2 for each CRDF and non-CRDF nutrient were selected and then merged with those

Table 2. Dishes selected from both the contribution analysis (CA) and variability analysis (VA) for the highest number of cancer-related dietary factors (CRDF)

Type of analysis	Number of CRDF*	Selected dishes
CA	11	Kimchi stew
	10	Soyabean paste soup
	9	Rolled rice (Gimbap), cooked rice with assorted vegetables (Bibimbap), beef soup, Alaska Pollack stew and loach soup
	8	Fried rice, noodles for special occasion, black bean sauce noodles, Chinese-style noodles with vegetables and seafood, chopped noodles, sea mustard soup, dried Pollack soup, beef soup with seasoned red pepper sauce, pork-on-the-bone soup, soyabean paste stew, dog soup, steamed pork, chives pancake, stir-fried pork, fried vegetables with potato noodles, hairtail boiled with seasonings and mackerel boiled with seasonings
VA	8	Soyabean paste stew
	6	Rolled rice (Gimbap), soyabean paste soup with radish leaves and kimchi stew
	5	Cooked rice with assorted vegetables (Bibimbap), noodles, noodles for special occasion, soyabean paste soup, plain, soyabean paste soup with mallow, beef soup with seasoned red pepper sauce, kimchi stew with tuna, hot Alaska Pollack stew, dog soup, loach soup, steamed pork, watery kimchi made of sliced radishes, Chinese cabbage kimchi, radish leaves kimchi, radish root and leaves kimchi
	4	Black-bean-sauce noodles, Chinese-style noodles with vegetables and seafood, soyabean paste soup with Chinese cabbage, soyabean paste soup with spinach, sea mustard soup with beef, beef soup, Alaska Pollack stew, hot pacific cod stew, radish leaves stew, fast-fermented soyabean paste stew, steamed bean sprouts with seafood, roasted dried laver, kimchi pancake, chives pancake, zucchini pancake, stir-fried pork, fried vegetables with potato noodles, seasoned spinach, cooked, Chinese cabbage salad, mustard leaved kimchi, cubed radish (cubed radish kimchi), rapeseed kimchi, pickled sesame leaves, lettuce

* Total number of CRDF is sixteen.

over 50% accumulated sum of the contribution to reach a manageable number of dishes. The resulting 181 dishes were then regrouped by similarity in main ingredients and/or serving unit, and four alcoholic beverages, fruits and yogurts were added to reach the final 112 dishes. For example, dishes such as ‘soyabean paste soup, plain’, ‘soyabean paste soup, chard’, ‘soyabean paste soup, Chinese cabbage’, ‘soyabean paste soup, spinach’, ‘soyabean paste soup, radish leaves’, ‘soyabean paste soup, mugwort’ and ‘soyabean paste soup, mallow’ were grouped into one dish item, ‘soyabean paste soup’, as presented in Table 3.

Determination of frequency response formats

The frequency response formats of consumption were classified into four types: nine categories for rice, soups, stews and side dishes (i.e. never to three times/d); nine categories for beverages (i.e. never to over six times/d); eight categories for fruits (i.e. never to over four times/d)⁽²⁵⁾; eight categories for alcoholic beverages (i.e. never to twice/d). The estimated intake of CRDF and non-CRDF nutrients was then converted to the total daily intake according to the frequency and serving size categories.

Decision of portion sizes

The portion sizes in the present study were determined based on the reported amount consumed in the KNHANES data. The portion sizes of small, medium and large for each dish represented the amount consumed in g by the KNHANES subjects at the 25th, 50th and 75th percentiles, respectively. A few exceptions were those dishes with a super central tendency with no or little variation in serving sizes among the population. The examples are steamed rice, rice with mixed grains, fruit juices and fried eggs, and their portion sizes corresponded to the 10th, 50th and 90th percentiles for small, medium and large portion sizes, respectively. In actual surveys, photographs of the representative dishes were presented on the questionnaire to differentiate the three portion sizes.

Development of a nutrient database of the 112 dish items

For each of the 112 dish items on the questionnaire, a nutrient database (per medium portion) was created from the analytical data from the 2001 KNHANES and the 2002 Korean National Nutrition Survey by Season, the recipes of each dish and the frequency weights. The final 112 dish items were made from 3260 ingredients. The ingredients appearing at a frequency of less than 1% were removed from each dish after an examination of the recipes. Consequently, the number of ingredients found in the 112 dish items was reduced to 1769. The nutrient composition of each dish item was calculated according to the median value of the nutrients of each dish. Groups of dishes were assigned weights of intake frequency from

Table 3. Development process for portion sizes of each dish item and nutrient database*

Example of a dish item	Variation	Frequency	Frequency wt	Energy (kJ)	Protein (g)	Vegetable (g)	Garlic (g)	Portion sizes (g)†			
								Small	Medium	Large	
								Magnification value	Weight	Magnification value	Weight
Soyabean paste soup	Soyabean paste soup, plain	657	0.30	183.9	3	6.2	0.0				
	Soyabean paste soup, chard	36	0.02	154.7	3.7	70.7	0.7				
	Soyabean paste soup, Chinese cabbage	344	0.16	129.6	2.5	58.0	0.7				
	Soyabean paste soup, spinach	134	0.06	142.1	3.2	47.5	1.0	0.7	43.2	1.00	61
	Soyabean paste soup, radish leaves	726	0.33	171.4	3.7	46.3	0.9				
	Soyabean paste soup, mugwort	135	0.06	167.2	3.3	21.9	0.0				
	Soyabean paste soup, mallow	151	0.07	179.7	5.2	71.6	1.2				
Nutrient composition‡			1	167.2	3.3	36.7	0.6				

* Nutrient database was calculated per medium portion size of dish.

† Small, medium and large portion sizes represent 25th, 50th and 75th percentiles, respectively, of the weighted portion sizes consumed by the subjects.

‡ Nutrient composition of the dish item = $\sum(\text{frequency weight} \times \text{nutrient composition})/\text{number of variations}$.

the analysis data. The nutrient database for alcoholic beverages was added after calculations of alcohols using the nutrient database from the appendix of the seventh Korean Recommended Dietary Allowance⁽²⁶⁾, with the measured weight of one unit for each type of alcoholic beverage as purchased from markets.

Results

The rankings of dishes with respect to the CRDF in the contribution analysis and the variability analysis are summarised in Table 2. The most frequently selected dish in the contribution analysis was 'kimchi stew', which contributed significantly to the intake of a total of eleven CRDF: red meat, processed meat, vegetables, garlic, beans, fish, Na, β -carotene, Ca, vitamin C and retinol. The next dish was 'soyabean paste soup', which contributed to a total of ten CRDF. The most frequently selected dish in the variability analysis was 'soyabean paste stew', which contributed significantly to the intake of a total of eight CRDF: vegetables, garlic, beans, fish, Na, Ca, β -carotene and vitamin C. The next dish items were 'rolled rice', 'soyabean paste soup with radish leaves' and 'kimchi stew', which contributed to a total of six CRDF. In these rankings, seven different types of kimchi and five different types of soyabean paste soup were included, indicating that these

two items are important contributors for both CRDF and non-CRDF nutrients in the Korean diet.

These 112 dish items consisted of nine types of staple foods, including rice and noodles, twenty-five types of soups and stews, fifty-four types of side dishes, nine types of beverages, nine types of fruits and six types of alcoholic beverages. A database of dish items was constructed, considering the median serving size of the actual distribution and frequency weight of each dish. An example of the frequency weights for 'soyabean paste soup' is shown in Table 3. The database was assigned a weighted mean value, which multiplied the intake frequency with the nutrients and food content of each dish. The energy content of 'soyabean paste soup, Chinese cabbage' showed 129.6 kJ/61 g (medium portion) as the lowest, whereas the energy content of 'soyabean paste soup, plain' showed 183.9 kJ/61 g as the highest. However, the frequency weight of 'soyabean paste soup, plain' was higher (0.30) than 'soyabean paste soup, Chinese cabbage' (0.16). Therefore, the weighted mean energy of the dish item 'soyabean paste soup' was calculated as 167.2 kJ/61 g, which was close to the energy of 'soyabean paste soup, plain'. Weights of vegetables and garlic have variation and ranges from 6.2 to 71.6 and from 0 to 1.2 g, respectively; however, their weighted means were calculated as 36.7 and 0.6 g, respectively. The protein content

Table 4. Number of dishes selected from the contribution analysis (CA) and variability analysis (VA) and the percentage coverage by the FFQ dishes for each cancer-related dietary factor (CRDF) and non-CRDF nutrient intake

Nutrients and foods	Number of dishes by CA	Number of dishes by VA	Number of total dishes in the FFQ*	Coverage (%)
CRDF				
β -Carotene	16	60	62	82.6
Ca	18	102	103	78.8
Vitamin C	10	49	49	84.1
Na	14	82	83	80.2
Retinol	9	1	9	67.2
Red meat	8	21	21	78.1
Alcohol	2	2	6	99.8
Processed meat	4	9	9	62.1
Fruit	4	8	11	90.9
Vegetables	10	95	95	83.2
Garlic	12	11	17	77.6
Carrots	6	24	24	67.5
Milk	1	1	1	95.6
Dairy food	1	1	3	95.7
Beans	4	21	21	82.6
Fish	28	76	79	56.7
Non-CRDF				
Energy	7	139	139	82.4
Protein	26	175	175	76.4
Fat	27	88	93	68.9
Carbohydrate	2	89	89	86.0
P	15	203	203	80.1
Fe	32	137	140	73.9
K	21	163	163	80.1
Vitamin A	19	44	48	80.3
Vitamin B ₁	14	86	87	80.0
Vitamin B ₂	27	149	151	77.1
Niacin	21	132	133	78.0

*The number of total dishes in the FFQ is not the sum of dishes from the CA and VA because of overlap. Four alcoholic beverages and two yogurts, not selected from the CA or VA, were added for alcohol and dairy food.

of the 'soyabean paste soup' was also calculated by the frequency weights. The magnification values for portion sizes of 'soyabean paste soup' were 0.7 for 'small' and 1.48 for 'large' compared with the median weight of the 'soyabean paste soup' (the magnification value of median portion size is always '1'). The magnification values for the three portion sizes of all dishes ranged from 0.3 to 0.98 for the 'small' portion and from 1.1 to 6 for the 'large' portion. They were greater than 2 in the 'large' portion for 'chicken gruel', 'steamed seafood with sprout/fried seafood', 'roasted duck' and others, and less than 0.5 for the 'small' portion of 'fried pepper', 'flatfish sushi', 'kale' and others (data not shown).

In Table 4, the number of dishes selected for each CRDF and non-CRDF nutrient and the values of the percentage coverage of CRDF and non-CRDF nutrient intake according to the dish items in the FFQ are presented. The number of dishes selected ranged from 1 (milk) to 103 (Ca) in CRDF and from 48 (vitamin A) to 203 (P) in non-CRDF nutrients. The percentage coverage for energy was 82.4%, whereas the values for protein, fat and carbohydrates were 76.4, 68.9 and 86.0%, respectively. In minerals, the percentage coverage of Fe was the lowest at 73.9%. In vitamins, retinol was the lowest at 67.2%, and vitamin C was the highest at 84.1%. In foods included in the CRDF, alcoholic beverages were the highest at 99.8% and fish was the lowest at 56.7%.

Discussion

Development of this dish-based, semi-quantitative FFQ to evaluate Korean intake of CRDF and non-CRDF nutrients is timely and significant. In using a database approach, the dish items of this newly developed FFQ were identified by using the most current report of the World Cancer Research Fund/American Institute for Cancer Research (2007) and the large-scale, nationally representative KNHANES dataset. This approach might be similar to the development of the Block 2005 FFQ that used dietary data from the National Health and Nutrition Examination Survey 1999–2002⁽²⁷⁾ and of the Diet History Questionnaire of the National Cancer Institute that used the representative national survey data of the continuing survey of food intake by individuals⁽²⁸⁾. These collections are representative FFQ that were formulated using a database approach. Block *et al.*⁽²⁹⁾ reported that a database approach improves an FFQ because it takes advantages of the large-scale national survey data of the representative population. The present study is the first attempt to develop Korean diet FFQ for cancer research using a database approach based on the national dietary survey data. The resulting dietary assessment instrument could be used to facilitate diet-related cancer cohort studies in Korea.

A nutrient composition database for each dish item was made using its recipe from the KNHANES rather than using standard recipes, which do not exist for all dishes selected. The KNHANES data were considered suitable for recipes

because the data were from a large population that represented the population of Korea. In many studies in Korea^(30–34), nutrient databases were formulated using various recipes instead of data pertaining to the particular food or dish. The Block Adult Questionnaire and the National Cancer Institute Diet History Questionnaire also use databases made up of the mean or median values calculated from the analysed data in the creation of a dish and food list⁽³⁵⁾.

The three portion sizes in our FFQ were determined from the rank orders of serving sizes reported by the Korean population in the KNHANES: small, medium and large, representing the 25th, 50th and 75th percentiles, respectively. This approach mirrors the dietary practices of our target population, in contrast to the commonly used arbitrary estimation of 0.5 and 1.5 times the medium size being the small and large serving sizes, respectively^(33,34,36). Similar approaches were taken in determining the portion sizes of the Block Adult Questionnaire and the National Cancer Institute Diet History Questionnaire by utilising percentile cut-offs⁽³⁵⁾.

The number of dishes selected from analyses for each CRDF and non-CRDF nutrient, and the percentage coverage of CRDF and non-CRDF nutrients by the 112 dish items in the FFQ, had large levels of variation (Table 4). In the CRDF, only one dish (plain milk) was selected by the contribution analysis and variability analysis for milk and for Ca, eighteen dishes were selected by the contribution analysis and 102 dishes were selected by the variability analysis, for a total of 103 dishes selected for Ca. In spite of that, the percentage coverage of milk was 95.6%, whereas that of Ca was 78.8%. This means that plain milk could explain milk intake on its own, but 103 dishes could not explain every part of Ca intake. The percentage coverage of alcoholic beverages, fruits, milk and dairy food exceeded 90%, whereas that of fish was less than 60%, partly because the selected fish dishes (*n* 79) could not represent the diverse fish dishes consumed by Koreans in the KNHANES (*n* 256). The percentage coverage of the six alcoholic beverages in our list approached 100% because four more frequently consumed alcoholic beverages were added, even though only two dishes were selected in each analysis for alcohol. Among cancer-related vitamins, the percentage coverages of retinol, β -carotene and vitamin C were 67.2, 82.6 and 84.1%, respectively. This is probably due to the incomplete nature of the KNHANES nutrient database for retinol. Nine dishes were selected for retinol because the KNHANES nutrient database for retinol has many missing values. The percentage coverage among macronutrients was highest for carbohydrates with 86% and lowest for fat with 68.9%. Kim *et al.*⁽³⁷⁾ also reported that the percentage coverage of fat was lowest among the macronutrients in their study of the Korean diet, that of Ca was 80%, which is similar to the results of the present study, and that of Na was 61.6%, which is much lower than that in

the present results. In the study of Block *et al.*⁽²⁹⁾, the percentage coverage of the FFQ exceeded 90% for all nutrients compared with the dietary survey data of the National Health and Nutrition Examination Survey. The differences in the percentage coverage of macronutrients may be explained by the differences in the FFQ and also by the differences in the study population, i.e. Koreans in the study of Kim *et al.*⁽³⁷⁾ v. Americans in the study of Block *et al.*⁽²⁹⁾. The percentage coverage of Fe in these three studies was low, possibly because of a large between-person variation as reported in the study of Bingham⁽³⁸⁾. In the study of Shahar *et al.*⁽³⁹⁾ in Israel, the number of foods explaining 80% of the between-person variability based on the regression model (variability analysis in the present study) for selected nutrients were all under forty. Dishes selected in the variability analysis of the present study explain 90% of the between-person variability, but over 100 dishes explained 80% of the between-person variability for some nutrients in the present study. In another study, Shahar *et al.*⁽⁴⁰⁾ compared the number of foods explaining 90% of the between-person variability among people who were born in different countries. The number of selected foods for nutrients differs considerably among these groups. This means that the number of foods or dishes explaining the between-person variability is different among populations that have a different dietary culture, and we need more dishes to explain enough between-person variability in the Korean population.

In general, if a small number of dishes represent diverse dishes that contribute to the selected CRDF, the percentage coverage is low. Conversely, if a small number of Korean dishes contribute to a selected dietary factor (such as plain milk for milk and dairy products), and that dish is selected for the final dish list, the percentage coverage becomes high. This underscores the importance of selection and exclusion processes for the final dish list in developing dish-based FFQ. Equally important is the development of cultural-specific FFQ that are developed from representative dietary survey data that reflect new trends of dietary lifestyles⁽⁴¹⁾.

The present study used the 2001–2 KNHANES dataset, which represents the sole and latest collection of data with seasonal dietary intake in Korea. The dataset had only 1 d of dietary intake data collection, which was extrapolated to estimate the usual intake. However, future surveys are planned to include 2 d for dietary surveys. As a result of the present study, this newly developed dish-based, semi-quantitative FFQ that calculates the intake of CRDF in the Korean diet is available for future cancer and nutrition research.

Acknowledgements

The present work was supported by the National Cancer Center in the Republic of Korea (grant no. 0720660). We thank Jae Eun Shim, PhD, for helpful suggestions during

developing the FFQ and manuscript revisions. M. K. P. designed the study, performed the statistical analysis, interpreted the data and wrote the manuscript. D. W. K. designed the FFQ checklist and performed the statistical analysis. J. K., S. P. and H. J. designed the study, contributed to the statistical analysis and data interpretation, and revised the manuscript. W. O. S. contributed to the data interpretation and revised the manuscript. H. Y. P. conceived of the study, participated in its design and coordination, and revised the manuscript critically for important intellectual content. All authors read and approved the final version of the manuscript. All authors declare no personal or financial interest in the content of the paper.

References

1. Korea National Statistical Office (2009) Annual Report on the Cause of Death Statistics. Daejeon, Republic of Korea: Korea National Statistical Office.
2. Ministry for Health Welfare and Family Affairs (2009) Annual report of the central cancer registry in Korea. Seoul, Republic of Korea: Ministry for Health Welfare and Family Affairs.
3. Doll R & Peto R (1981) The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst* **66**, 1191–1308.
4. World Cancer Research Fund/American Institute for Cancer Research (1997) *Food, Nutrition and the Prevention of Cancer: A Global Perspective*. Washington, DC: American Institute for Cancer Research.
5. World Cancer Research Fund/American Institute for Cancer Research (2007) *Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective*. Washington, DC: American Institute for Cancer Research.
6. Kolonel LN, Henderson BE, Hankin JH, *et al.* (2000) A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am J Epidemiol* **151**, 346–357.
7. Stram DO, Hankin JH, Wilkens LR, *et al.* (2000) Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. *Am J Epidemiol* **151**, 358–370.
8. Chlebowski RT, Pettinger M, Stefanick ML, *et al.* (2004) Insulin, physical activity, and caloric intake in postmenopausal women: breast cancer implications. *J Clin Oncol* **22**, 4507–4513.
9. The Women's Health Initiative Study Group (1998) Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials* **19**, 61–109.
10. Patterson RE, Kristal AR, Tinker LF, *et al.* (1999) Measurement characteristics of the Women's Health Initiative food frequency questionnaire. *Ann Epidemiol* **9**, 178–187.
11. Kristal AR, Shattuck AL & Patterson RE (1999) Differences in fat-related dietary patterns between black, Hispanic and White women: results from the Women's Health Trial Feasibility Study in minority populations. *Public Health Nutr* **2**, 253–262.
12. Deharveng G, Charrondiere UR, Slimani N, *et al.* (1999) Comparison of nutrients in the food composition tables available in the nine European countries participating in EPIC. European Prospective Investigation into Cancer and Nutrition. *Eur J Clin Nutr* **53**, 60–79.
13. Hoffmann K, Kroke A, Klipstein-Grobusch K, *et al.* (2002) Standardization of dietary intake measurements by nonlinear calibration using short-term reference data. *Am J Epidemiol* **156**, 862–870.

14. Michels KB, Welch AA, Luben R, *et al.* (2005) Measurement of fruit and vegetable consumption with diet questionnaires and implications for analyses and interpretation. *Am J Epidemiol* **161**, 987–994.
15. Slimani N, Kaaks R, Ferrari P, *et al.* (2002) European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study: rationale, design and population characteristics. *Public Health Nutr* **5**, 1125–1145.
16. Voss S, Charrondiere UR, Slimani N, *et al.* (1998) EPIC-SOFT a European computer program for 24-hour dietary protocols. *Z Ernahrungswiss* **37**, 227–233.
17. Iso H, Date C, Noda H, *et al.* (2005) Frequency of food intake and estimated nutrient intake among men and women: The JACC Study. *J Epidemiol* **15**, S24–S42.
18. Date C, Fukui M, Yamamoto A, *et al.* (2005) Reproducibility and validity of a self-administered food frequency questionnaire used in the JACC study. *J Epidemiol* **15**, S9–S23.
19. Shin A, Lim SY, Sung J, *et al.* (2009) Dietary intake, eating habits, and metabolic syndrome in Korean men. *J Am Diet Assoc* **109**, 633–640.
20. Shin A, Lim S, Sung J, *et al.* (2009) Dietary habit and bone mineral density in Korean postmenopausal women. *Osteoporos Int* **21**, 947–955.
21. Kim DJ, Shin DH, Ahn B, *et al.* (2003) Chemoprevention of colon cancer by Korean food plant components. *Mutat Res* **523–524**, 99–107.
22. Park K, Kim S, Suh M, *et al.* (1991) Inhibitory effects of garlic on the mutagenicity in salmonella assay system and on the growth of HT-29 human colon carcinoma cells. *Korean J Food Sci Technol* **23**, 370–374.
23. Ahn Y, Lee JE, Cho NH, *et al.* (2004) Validation and calibration of semi-quantitative food frequency questionnaire – with participants of the Korean Health and Genome Study. *Korean J Community Nutr* **9**, 173–182.
24. Kim MK, Yun YM & Kim Y (2008) Developing dish-based food frequency questionnaire for the epidemiology study of hypertension among Korean. *Korean J Community Nutr* **13**, 701–712.
25. Loma Linda University Adventist Health Studies (2002) Food Frequency Questionnaire of Adventist Health Study-2 (AHS-2). <http://www.llu.edu/pages/health/documents/ahs-2.pdf> (accessed 20 October 2009).
26. The Korean Nutrition Society (2000) *Recommended Dietary Allowances for Koreans*. Seoul, Republic of Korea: The Korean Nutrition Society.
27. Block G (2005) Block Adult Questionnaire. http://www.nutritionquest.com/products/questionnaire_screensers.htm (accessed 20 October 2009).
28. Subar AF (1998) National Cancer Institute Diet History Questionnaire. <http://riskfactor.cancer.gov/DHQ/> (accessed 20 October 2009).
29. Block G, Hartman AM, Dresser CM, *et al.* (1986) A data-based approach to diet questionnaire design and testing. *Am J Epidemiol* **124**, 453–469.
30. Ji SK, Kim HS & Choi HM (2008) A study on development and validation of food frequency questionnaire for estimating energy intake of women in child-bearing age. *Korean J Community Nutr* **13**, 111–124.
31. Kim WY & Yang EJ (1998) A study on development and validation of food frequency questionnaire for Koreans. *Korean J Nutr* **31**, 220–230.
32. Lee HJ, Park SJ, Kim HJ, *et al.* (2002) Development and validation of a computerized semi-quantitative food frequency questionnaire program for evaluating the nutritional status of the Korean elderly. *Korean J Community Nutr* **7**, 277–285.
33. Won HS & Kim WY (2000) Development and validation of a semi-quantitative food frequency questionnaire to evaluate nutritional status of Korean elderly. *Korean J Nutr* **33**, 314–323.
34. Yim KS, Lee TY & Park HS (2003) The development and validation of a food frequency questionnaire to assess diets of Korean adolescents. *Korean J Community Nutr* **8**, 277–285.
35. McNutt S, Zimmerman TP & Hull SG (2008) Development of food composition databases for food frequency questionnaires (FFQ). *J Food Compos Anal* **21**, S20–S26.
36. Kim MY, Suh I, Nam CM, *et al.* (2002) The development and evaluation of a simple semi-quantitative food frequency questionnaire using the contribution of specific foods to absolute intake and between-person variation of nutrient consumption. *Korean J Nutr* **35**, 250–262.
37. Kim J, Kim YJ, Ahn YO, *et al.* (2003) Development of a food frequency questionnaire in Koreans. *Asia Pac J Clin Nutr* **12**, 243–250.
38. Bingham S (1987) The dietary assessment of individuals: methods, accuracy, new technique and recommendations. *Nutr Abstr Rev Ser Hum Exp* **57**, 705–742.
39. Shahar D, Fraser D, Shai I, *et al.* (2003) Development of a food frequency questionnaire (FFQ) for an elderly population based on a population survey. *J Nutr* **133**, 3625–3629.
40. Shahar D, Shai I, Vardi H, *et al.* (2003) Development of a semi-quantitative food frequency questionnaire (FFQ) to assess dietary intake of multiethnic populations. *Eur J Epidemiol* **18**, 855–861.
41. Cade J, Thompson R, Burley V, *et al.* (2001) Development, validation and utilisation of food-frequency questionnaires – a review. *Public Health Nutr* **5**, 567–587.