A new method for developing an efficient, regional semi-quantitative food frequency questionnaire

doi: 10.6133/apjcn.202505/PP.0005
Published online: May 2025

Running title: New method for developing an efficient SQFFQ

Desheng Liu MB[1,2], Xiaopei Guo MB[3], Hui Chen MB[1], Rui Zheng MB[1], Qingquan Liu PhD[1]

[1]Department of Endocrinology, Tangdu Hospital, Air Force Military Medical University, Baqiao District, Shaanxi, China
[2]Department of Pulmonology, No. 970 Hospital of the PLA Joint Logistics Support Force, Zhifu District, Shandong, China
[3]Department of Rheumatology and Immunology, Xi'an Third Hospital, Weiyang District, Shaanxi, China

Authors' email addresses and contributions:
DL: 366177633@qq.com
Contribution: conceived the study question, and contributed to the study design, supervision of data collection, data analysis and interpretation, and writing the manuscript, undertook data collection and data analysis, and contributed to data interpretation, and writing the manuscript.

XG: 1095698643@qq.com
Contribution: undertook data collection.

HC: chh2277@163.com
Contribution: undertook data collection.

RZ: 2521976153@qq.com
Contribution: undertook data collection.

QL: liuqingq@fmmu.edu.cn
Contribution: conceived the study question, and contributed to the study design, supervision of data collection, and reviewing the manuscript.

Corresponding Author: Dr Liu Qingquan, Department of Endocrinology, Tangdu Hospital, Air Force Military Medical University, No. 1 Xinsi Road, Baqiao District, Xi'an, 710000, Shaanxi, China. Tel: +86 13991380683. Email: liuqingq@fmmu.edu.cn

ABSTRACT

Background and Objectives: To establish a new method for developing an efficient and region-specific semi-quantitative food frequency questionnaire (SQFFQ), and to make it a simple, effective and widely applicable measurement tool, this study takes Xi'an City, Shaanxi Province as a case study. **Methods and Study Design:** For 129 common food items, an initial food frequency questionnaire (FFQ) and a 24-hour dietary recall survey (24hDR) are constructed, and a variety of methods are combined to improve the convenience and credibility of the questionnaire itself in the inspection process, and the reliability and validity test are carried out. **Results:** A total of 367 questionnaires are collected in study, including FFQ and 24hDR at various correction stages, in a three-month interval, and finally a semi-quantitative food frequency questionnaire containing 104 food items is obtained. The Pearson correlation coefficient of SQFFQ1 and SQFFQ2 ranges from 0.60-0.75, and the Spearman and Kendall correlation coefficients range from 0.50-0.89, and the correlation coefficient increases slightly after logarithmic transformation by logarithmic transformation, and the CCC correlation coefficient increases to 0.5-0.95. On average, more than 80% of subjects are classified as identical or adjacent tertiles per SQFFQ result. The Pearson, Spearman, Kendall, CCC correlation coefficients between 24hDR and SQFFQ2 range from 0.50-0.90, and on average more than 75% of participants are in the same or adjacent energy and nutrient tertiles in SQFFQ2 versus 24hDR results. **Conclusions:** The semi-quantitative food frequency questionnaire for dietary pattern surveys in Xi'an, Shaanxi can effectively assess nutrient intake in dietary nutrient correlation studies.

Key Words: Semi-quantitative Food Frequency Questionnaire (SQFFQ), 24-hour dietary recall survey (24hDR), usual intake, dynamic programming, validity test

INTRODUCTION

This study aims to develop a semi-quantitative food frequency questionnaire (SQFFQ) to obtain information about the research subjects' food, nutrient intake, and dietary habits.

Studies have demonstrated that a fifth of global mortalities are linked to people's unwholesome diets, characterized by high sugar, salt, and processed meat consumption. Such dietary habits can precipitate heart diseases, cancer, and diabetes. Furthermore, recent data from disease burden research reveals that the proportion of disabilities and fatalities attributed to a spectrum of conditions, namely cancer,[1-4] cardiovascular diseases,[5-7] metabolic disorders like type 2 diabetes,[1, 7] digestive system ailments such as non-alcoholic fatty liver,[8-10] and neurodegenerative diseases including depression[11-14] and cognitive impairment,[15-17] has been on the rise annually. Diet serves as a prevalent and modifiable factor influencing these diverse diseases, and it plays a pivotal role in both their prevention and treatment.[18-19] Upon the successful development of the scale, the research outcomes can be harnessed to assess dietary patterns, ascertain the objective intake of dietary nutrients, and identify the overall impact of the research subjects on their general health, as well as the potential progression of diseases in the future. This information can also offer guidance and recommendations for the optimization and adjustment of dietary patterns in the time to come.

The Food Frequency Questionnaire (FFQ) is a dietary survey method commonly used in nutrition epidemiology studies that uses forms containing dozens of food items to ask respondents about food consumption habits, including the frequency of intake of various foods or food groups and the size of each intake. In population nutrition epidemiological studies, FFQ is required to be as simple and practical as possible, reliable, and effective. In addition, small changes in the population of different genders, cultural backgrounds, geographical environments, and questionnaire structure can affect the credibility and validity of food frequency questionnaires. Therefore, when the target population or the questionnaire itself changes, the credibility and validity of the food frequency questionnaire must first be tested before it can be applied to the relevant population.[20-22] Therefore, this study aims to construct a SQFFQ for the investigation of dietary patterns in the Shaanxi region and test its effectiveness.[23-24] In recent years, Xi'an, as an emerging first tier city, has seen significant economic growth and a complex population composition and disease situation. This study aims to investigate a dietary survey method and its specific construction method for the Xi'an region of China, enabling researchers to obtain dietary nutrients from a dietary perspective in complex social and cultural environments, among long-term residents with stable dietary habits.

Furthermore, in the existing development of FFQ, it is often difficult to balance the item with the subject's objective conditions (cultural level of the subject) and the subjective will (subject's understanding and enthusiasm for the subject), either the item is too complex can not allow the subject to understand or maintain a better motivation, or the item less can not include the subject's better meal program.

Therefore, in the development of this study, the innovative construction of the objective correction method of the options, and the new proportional calculation method for the intake of food items, after the credibility test has obtained good results, can indicate that the FFQ built in this study in the balance of subjects' subjective enthusiasm, understanding difficulty and reaction subjective will can obtain a better feasibility, to the subject's dietary nutrients obtain an objective response.

If the subjects have a low level of culture or enthusiasm for participation, the 24hDR may be the most reasonable option for the validity assessment of the FFQ (SQFFQ) standard. This study improved the 24hDR as a test standard in the FFQ validity testing, introduced the concepts of "usual intake" and "pseudo data" and combined, can improve the objectivity of the 24h dietary recall response to real diet results and also increase its credibility as the FFQ validity testing standard, so that under the premise of low culture of subjects or enthusiasm of participation, the evaluation criteria reflect the objectiveness of reality.[25-26] The method described in this article can be applied to populations with different cultural and social backgrounds, long-term residency, and stable dietary habits. However, the construction of questionnaires for different dietary and sociocultural backgrounds needs to be redone.

The aim of this study, focusing on dietary surveys with regional traits, is to note that when creating dietary questionnaires for wide-ranging regions and various cultural backgrounds, it's tough to balance the scope and accuracy of food coverage. In Chinese diet research, unlike Western studies on the "Mediterranean diet", an obvious and difficult problem arises. The Chinese diet overemphasizes diverse cooking methods and has a rich variety of food materials.

So, in making dietary questionnaires, if too many kinds of finished food products are included, there will be excessive collinearity in dietary nutrients among different food items, and the questionnaires will be very cumbersome. On the other hand, if only basic food materials are stressed, the questionnaires will be thin and overlook the impact of the cooking process on the diet.

Hence, this study limits the dietary survey scope to an environment with specific regional dietary characteristics. First, collect dietary items in this region (including uncooked food materials and cooked food items, ensuring food materials are minimally involved in cooked food production to avoid double-counting of dietary nutrients) and put them into the questionnaire. Then, test its reliability and validity.

Moreover, using this method to build regional dietary questionnaires can create new ones in any Chinese dietary context. Therefore, this study aims to develop a new and somewhat universal way to construct dietary questionnaires, which can achieve comprehensive and convenient dietary surveys that are easy for the whole population to accept and understand. Due to the construction method, the target population needs to have a relatively stable dietary environment to some extent. So, it may not be applicable to those who have changed their dietary environment within a year.

MATERIALS AND METHODS

Ethics Committee approval was obtained from the Institutional Ethics Committee of Tangdu Hospital to the commencement of the study.

Research methods

First, the preliminary experiment was carried out using the initial FFQ questionnaire. The FFQ results from 65 randomly selected subjects were analyzed, and options were modified based on the results. Subjects were residents of community A in Xi'an and met the following dietary conditions (subjects in subsequent FFQ and 24hDR both met the following requirements).

Then the survey was conducted in the same attribute community B scope, and eventually charging 104 community residents in normal living and eating environments, using the revised FFQ questionnaire and the 24hDR, based on the results to screen food items (subjects in the subsequent FFQ and 24hDR both met the following requirements).

Then, the survey was carried out in the community C range of the same attribute, and eventually counting 50 community residents in normal living and eating environments, SQFFQ (SQFFQ1, SQFFQ2) were screened twice, both carried out at intervals of 3 months. In addition, the same subjects received FFQ2 accompanied by a 24hDR. Finally, the SQFFQ validity and reliability testing is carried out based on the results.

The majority of the data in this study are continuous variables, and each food item in the preliminary experiment has undergone a normality analysis. Use the box plot approach for data items with non-normal distribution and the 3σ criteria for data items with normal or almost normal distribution when finding and differentiating outliers; When dealing with outliers, erasing data might lead to information loss and, if uncontrolled, could compromise the test's efficacy. It is therefore seen as using mean imputation. In the handling of missing data, the mean is also employed for filling. In real operation, the number of missing and outlier data for each item should not be too high, and should not exceed 5%. In addition, for subjects with follow-up losses, remove their corresponding past questionnaires.

Subject requirements: (1) The subject needs to be in a stable diet environment for a long period of time (more than six months), such as a long-term home diet, and have a more accurate understanding of the stable food raw materials and cooking methods. (2) Subject geography must be kept in stable areas. (3) Subject is an adult; exclusion criteria: If the subject is ill, its condition has a short-term effect on the subject's dietary ability, or if the age is too young or too large, the subject cannot eat independently, and other reasons lead to changes in dietary patterns in the short term, the subject should be excluded.

Additional information regarding the subjects is provided below: Age range: 18 to 60 years old. The ethnicity is Han. The subjects have lived in the Xi'an area for the past year and have gotten cultural education and experienced daily life in Xi'an before becoming adults. They have completed the nine-year compulsory education in the local environment of Xi'an, China; The Xi'an area referenced in this article is divided into three administrative districts: Xincheng District, Yanta District, and Baqiao District.Xi'an, the economic and cultural center of northwest China, has a diverse personnel makeup. To increase and clarify the questionnaire's information collecting efficiency and testing efficacy, the screening and recruitment of subjects must be clarified and limited to a long-term, controllable environment. As a result, the purpose of this article is to analyze and develop appropriate adults who have lived in Xi'an for a long time and are familiar with the local culture.

In addition, the number of subjects referred to the number in previous relevant studies, which was more appropriate given the presumption of racial and geographical limitations, usually 50-100 cases.[27-29] In this study, a total of 367 questionnaires were collected by combining all the initial FFQ (FFQ for pre-experimental use), modified FFQ, items-screened SQFFQ, and 24hDR.

The specific strategy for collecting and summarizing data from participants in this study was primarily to train a group of survey personnel to fill out information on tablets, mobile phones, and paper questionnaires under the supervision of survey personnel in a face-to-face setting. The researchers then entered the paper's material, summarized the data, and performed statistical analysis. In terms of participant recruitment, participants actively registered and participated using the contact information provided on the Tangdu Hospital Endocrinology Department poster, and survey personnel provided face-to-face guidance to participants on how to complete the questionnaire. In addition, we mobilized community forces to promote this research project, eventually ensuring that the living environment of participants was limited to the designated community. The location of the survey is determined by the individual's situation, usually in designated meeting rooms in hospital or community activity centers. The first batch of subjects recruited for this study beginning in October 2022, and all data collection completed by January 2024.

We employed the notion of usual intake to simulate and get the subjects' nutritional status within one year based on the 24-hour DR results. All dietary survey work in articles such as SQFFQ and 24-hour DR will be done on workdays, as patients' weekend special dietary conditions may impair their recall and questionnaire completion. The flowchart illustrating SQFFQ development process is shown below.

Step 1: Initial FFQ pre-setting

Questionnaire 1: The 24hDR and initial FFQ questionnaire

Step 2: Community A – Inclusion of local food items, Correction of initial FFQ options

Questionnaire 2: Revised FFQ questionnaire and the 24hDR

Step 3: Community B – Food item screening

Questionnaire 3: SQFFQ (SQFFQ1, SQFFQ2) and the 24hDR

Step 4: Community C – Validity and reliability testing

In accordance with the questionnaire development steps of this study, the general characteristics of the participants in each community is supplemented in Table 6, covering indicators such as age, gender, BMI, education level, and monthly income. Among them, age and BMI are measurement data, while gender, education level, and monthly income are enumeration data. The monthly income survey data is divided into four levels: "4000 and below, 4001-8000, 8001-12000, and above 12001", and the educational level survey data is divided into four levels: "University (junior college and above), Senior high school (including secondary vocational school), Junior high school, Primary school and below".

Initial FFQ pre-setting

Initial FFQ food items setting

The initial FFQ's food standards were based on a compilation of many questionnaires from the body of existing literature. In addition, to achieve the inclusion of local food items, localized supplements were also made to the food items at this stage based on the results of 24hDR. A number of dietary rating criteria are also used in this approach to classify foods.The nutrient calculation of food is based on the data on the 2009 China Food Composition Table.The initial FFQ's food standards shown in Table 1.

Food item weight calculation

Multiple methods for calculating the weights of food items under large food categories

For various refined rice varieties with lower dietary intake, such as millet, corn, corn flour, glutinous rice, brown rice, rye, black rice, purple rice, oats, barley, buckwheat, quinoa, and so on, first conduct a large classification of "other rice" to "frequency" and "quantity" surveys, and then conduct a subjective ratio of various rice meals in the "Other rice." The goal of this strategy is to decrease the time and burden of a single test for participants while still including a variety of food items in the FFQ questionnaire.

The consumption ratio calculation results are compared using three proportion calculation methods: the initial ratio calculation, the adjusted proportion calculation (the denominator is not 1), and the adjusted proportion calculation (the denominator is 1).

Results of the initial ratio calculation: The subject subjectively is likely not to pay attention to whether the sum is 100, and the premise of this method is to simply take the sum of all options as the denominator, and then calculate their respective proportions, so the subject's subjective choice of the sum of the denominators of the consumption ratio is 100 to truly reflect the subject's will.

The following are the results of the adjusted percentage calculation (the denominator is not 1): The modified ratio calculation approach squares the original percentage calculation and compensates for molecules based on the number of options.

The following are the results of the adjusted percentage calculation (the denominator is 1): The proportion calculation here on the basis of the above can truly reflect the subjective intention of the subject to a certain extent.

Proportion calculations under various individuals' subjective situations

The following is a list of various potential objective manifestations of the subjects' subjective choices, and a comparison is made between the original proportion calculation method and the adjusted proportion calculation method in terms of the subjects' subjective expression to show how the adjusted calculation method can more accurately reflect the subjects' subjective wishes.

Comparing the outcomes of the initial ratio calculation with the adjusted proportion calculation (the denominator is 1) will help you understand the differences. The lower row's 0 indicates that the former's initial ratio calculation result is bigger than the latter's adjusted proportion calculation (the denominator is 1), while the lower row's 1 indicates that the former is smaller than the latter.

Scenario 1: The slope and range of the latter will increase appropriately if the change in proportion curve is smoother and the range is small; conversely, if the change in proportion curve is smoother and the mean is close to 100, the slope and range of the latter will increase weakly. The smooth curve may reflect the subject's subjective belief that all types of food intake are similar.

1) If the proportion of options selected in the beginning is small, and the person subjectively believes that their intake is relatively low and the difference is not significant, they may tend to converge in the selection of options in the future. Therefore, it is necessary to try adding some slopes and ranges to widen the gap.

2) If the proportion chosen in the previous options is relatively large at the beginning, the person may subjectively believe that their intake is too high and the difference is not significant. If they choose the proportion too close to the upper limit and do not attempt to choose a smaller proportion, it may reflect the subject's subjective belief that the intake of each food is large and there is no clear difference. Therefore, the adjusted formula and attempts to widen the gap should be weakened.

Scenario 2: If the change curve of the proportion is steeper and the range is large, the latter will further increase the amplitude of the slope and range; If the curve of change in proportion is steep and the mean is close to 100, the latter will increase the slope and range by a weaker extent, and the steep curve can reflect the subject's subjective belief that the intake of various foods is quite different.

1) If the proportion of the previous option selection is small at the beginning, that is, if the subject subjectively believes that the intake in front of them is small, but finds that he may have a larger intake in the later items, and then, in the proportion selection of the latter items, he may gradually try to increase the proportion selection, so that when the upper limit is almost reached, it is found that it can no longer be increased, and then subjectively weakens the increase in the proportion of subsequent items, so it is necessary to increase the slope and range to open the gap.

2) If the proportion of the previous option selection at the beginning is large, that is, if the subject subjectively believes that the intake in front of them is large, but finds that he may have a smaller intake in the back, and then, in the proportion selection of the latter items, he may gradually try to reduce the proportion selection, so that when the lower limit is almost reached, it is found that it can no longer be reduced, and then subjectively weakens the reduction of the proportion of the subsequent items, so it is necessary to increase the slope and range to open the gap.

Understanding the results of different proportion calculations

Due to the underlying concept of subjective selection, which states that the more subdivided the food item is, the more the subject's decision will subjectively tend to converge in regard to the food item, this adjustment is meant to enhance the degree of dispersion of differences across detailed meals.

As an illustration, the respondents' subjective preference for tiny food intake values, such as 1 and 2, differs from their preference for bigger values, such as 99 and 100; the former is around 0.5:1 while the latter is roughly 0.99:1. Therefore, for the unit values within each subdivision of food items, the smaller the overall mean, the difference between the food items that the subjects want to express with smaller intake, is more likely to calculate a larger difference in the initial ratio calculation.

As a result, the necessity to lessen the degree of dispersion increases with the number of unit values that have been acquired per meal intake.The adjusted proportion calculation (the denominator is 1), was shown to be capable of reflecting the genuine purpose underlying the subject's subjective decision.

Correction of initial FFQ options and food item screening

Correction of the initial FFQ option

Design of the option-correction questionnaire

The individuals who were community members in Xi'an and the surrounding regions and who satisfied the following dietary circumstances were first investigated using the initial FFQ (subjects in subsequent FFQ and 24hDR met the following requirements), the requirements and number of participants and their references have been described above.

In previous studies, the division of option intervals was often more subjective, relying on the experimenter's subjective division of options for the habitual daily purchase and intake of something, and there was no more objective way to divide according to the survey results. In this study, the questionnaire was initially adjusted according to the results, and the options were modified according to the various statistical distributions of the results of various questions.

Option correction

The statistical distribution of the findings may be separated into the two following scenarios when the results of the initial FFQ are gathered.

(1) Scenario 1 - The results of bimodal distribution as shown in Supplementary materials Table 1.

The interval allocation of option 4.5.6, as depicted in the image, is irrational, and the subjects have trouble understanding option 5, which eventually falls between options 4 and 6. It can be said that there is a phenomena of cross overlap between the results of option 4.5.6 if the probability of choosing choices 4 and 6 is the same and the results of options 4 and 6 are comparable.

The multiple choice questions in which the survey results fit this situation are (Frequency):

Rice, fresh dark vegetables, fresh light vegetables, fresh fruits, red meat, milk, and milk made tea drinks, salt

The multiple choice questions in which the survey results fit this situation are (Intake):

Tea Beverage Eggs

(2) Scenario 2 - The result of skewed distribution as shown in Supplementary materials Table 2.

The options must be broken down since the statistical findings of the scenario depicted in the image are biased to one side. The outcomes of option 1.2 in the table are comparable, and things with comparable results have to be broken down.

For example, salt appears to be distributed evenly while high-salt foods like pickles and pickles have a skewed distribution. If this is the case, only option 1 needs to be decomposed because the sample size of that option is too large in comparison to other sample sizes. Option with 0% intake is not taken into account.

The multiple choice questions in which the survey results fit this situation are (Frequency):

Animal oil, beer, liquor, fruit wine, liquor, liquor, etc

The multiple choice questions in which the survey results fit this situation are (Intake):

Pickled vegetables, meat, rivers, fresh aquatic products, seafood, aquatic products, animal oil, vegetable oil, salt, beer, wine, fruit wine, wine, liquor, etc

Method of option correction

To mimic the link between subjective cognition of the interval in the option and objective selection of the subject population, the notion of "dynamic programming" is introduced in the option correction approach. Dynamic programming aims to level out the distribution of options.

"n", or the number of samples or individuals, is used in the computation. Ideally, each subject should take up one unit and be placed in a definite sequence (The nature of options supports this supposition.), like Figure 1.

The different methods of segmenting the options are appropriate in the best case scenario, where the test group selection is essentially uniformly distributed. Time series 1 (k=1) can be used to represent the ideal condition， according to the time series in Figure 2.

So when $n = \infty$, there is an ideal situation for any segmentation option in any time series.Accordingly, each state of k will have an ideal state under the assumption that $n = \infty$, and the ideal state of $k = n-1$ will likewise exist when $k = n$. Under the premise of $n \neq \infty$, any state of k will not have an ideal state, and the ideal state at $k = n-1$ may not exist when $k = n$ (as Figure 3 shown).

On the basis of this, we thus infer that each rise in k is in the direction of boosting the degree of deidealization. This leads us to the conclusion that as k increases, idealization is less likely to occur.

An ideal distribution has lengths for each interval that are equal. Ideally, each interval will have the same expected value.

Accordingly, each k in the interval may be thought of as a "guessing opportunity". Because it is easier and quicker to avoid judgment when the subject's subjective will is known as opposed to "the individual subject in each unit chooses and puts his own will on a one-dimensional grid with a base unit of 1 and an interval number of n".

Each grid may be viewed as an opportunity for the purpose of "guessing opportunity" comprehension. Because the subject's choice is arbitrary and subjective. Each "small grid" (unit interval with length 1) that reflects the most fundamental and essential will of the subject is a test or guessing opportunity for the tester to perform the task of analyzing the investigation's findings because it can be assumed that the people in this "big grid" (option interval) share a certain common value, but the tester is unaware of this value.

Therefore, when the tester guesses, he uses each expected value to guess the common value that reflects the situation of the real group.

The selection intents of every subject who filled out a "big grid" (interval of the option) should be taken into account as a whole. Even the choices will be distributed among all little units over the interval.

The feedback on each of the tester's hypotheses is just a yes or no when participants in a shared "big grid" are considered as a whole.

The increase in the number of guesses between k and k-1, from the standpoint of dynamic programming, also results in an increase in the permitted range of predicted value.Here, we suppose that f(k) is the maximum expected value function, and that n/f(k) = g(k) is the cost function.

The larger the g(k), the larger the expected range of our testers' limited guesses, which only accounts for a small range of the total number of units of the subject. There is a high likelihood that our expected range cannot accommodate the extreme subjective intention distribution of subjects, leading to a higher cost of trying.

In other words, the common value of the range 0-f(k) may be verified, and once it is surpassed, it prevents the tester from making a reliable prediction. "k" is the number of "small grid" (unit interval of length 1) in the "big grid" (interval of the option). Additionally, if f(k) > n, an ideal state cannot be created while n is limited.

In order to identify the ideal state in a limited time series k, it is necessary to maintain f(k) =< n.

In conclusion, the issue may be transformed into a discrete sequence with increasing intervals of 1 from 0 to n and the process of estimating the common value of subjects using the tester's predicted value.

Here, f(n) is a function of the number of subjects corresponding to each small unit n in the interval, with the common value n'=argmin(f(n)), as shown in Figure 4.

$$f(k) = (k^2 + k)/2$$

There are several permutations in each interval (i.e., the real condition of each individual subject in the interval) because the interval size of each choice is the same. The common value n' is viewed as a value that reflects the set if it is assumed that there is a set of all conceivable permutations.

As a result, the transformed problem is fundamentally a topic in dynamic programming. In order to explore the size of n 'in the interval, and n' must also be found in k times limited tester guessing, failing which the ideal state cannot be produced. This also increases the cost of attempts, so it is necessary to as much as possible lower the cost of attempts under the presumption that the ideal state can be contained.

Food item screening

The treated initial FFQ and 24hDR were used to create a SQFFQ based on evidence of food items and nutrient consumption for multiple regression analysis (MRA), contribution analysis (CA), and favorable food modification screening food items.

Multiple regression analysis for screening food items (MRA)

The amount of each nutrient provided by all foods was used as the response variable. Based on the nutrients, multiple regression analysis was performed using the stepwise entry method and the optimal subset, and the food items were screened based on the size of the contribution of the corresponding variable of the respective variables, that is, the accumulation of the coef In later research, screening standards are often 85% or 90%.

$Y_i = \beta_0 + \beta_1 X_{i2} + \ldots\ldots\ldots.. + \beta_n X_{in} + \varepsilon_i$

Y_i : Various dietary nutrients

β_0, β_1, β_2 ……. βn: Regression coefficient

X_{i1},X_{i2},X_{i3} .... X_{in}: Nutrients ingested by each food

The progressive growing stage of Cumulative R2 may be expressed using the stepwise regression approach due to the high number of independent variables. The variables are then filtered based on the stepwise findings, and the optimal subset regression analysis is performed.

To avoid the issue of overfitting, the first stepwise regression can take globality into account to obtain a wide range of variable screening. According to the change of the cumulative R2 value, the optimal regression equation that can take globality into account can be obtained. This is because there are many features, and the optimal subset suitable for fewer features cannot be used here directly. The variables are therefore ordered by cumulative R2 value at the moment of inclusion following stepwise regression screening, and the variables are then further screened by the ideal subset.

Cumulative contribution analysis to screen food items (CA)

The food items are ranked from largest to smallest based on the amount of several target nutrients provided by various foods consumed on average per person's daily intake. Each nutrient is still tabulated separately, and the percentage of nutrients provided by each food accounts for the total amount of the nutrient is calculated. The percentages are accumulated in order of the percentage from largest to smallest.

It is determined what proportion of a given nutrient different foods contribute, and the cumulative percentage contribution rate is determined by adding together the biggest to the lowest percentage contributions. In later investigations, the screening criterion were often 85% or 90%, and 85% was chosen as the screening level in this study.

Food Items with dominant food modification screening

The meal at the top of each category in the 24hDR is utilized as a sub-item in the category to make up for any potential FFQ shortfalls. Based on the 24hDR data, each major food category's absolute intake was sorted, the contribution of each food item to the major category's absolute intake was computed, and the food items were filtered based on a >1% threshold.

24hDR treatment

The 24hDR typical intake concept and calculation:

An individual's 24hDR, based on any contextual and behaviorally viable 24hDR average, nonetheless contains a significant amount of intra-individual variance for the majority of food components.

As a result, the idea of normal intake is presented to describe the subject's long-term nutritional intake for efficacy testing.

We often presume that food consumption during a 24-hour period is unbiased estimate of usual intake. Based on this, it is clear that the distribution of the individual average has greater volatility than the actual distribution of usual intake, causing normal intake to be higher or lower than the biased estimate.

A single 24hDR at a certain time may overestimate or underestimate a person's genuine typical consumption, but if the 24hDR is used frequently to the same individual, the estimation error will be eliminated. Nevertheless, the absence of bias does not imply the absence of mistake.

Based on this supposition, the deviation from the usual intake can be quantified as the sum of the deviation in average usual intake between groups grouped based on subject attributes (between-person difference), and the deviation between the individual and the group in which the individual belongs (within-person difference is divided by n, where n is the number of independent 24hDR repeat tests).

The calculation methods of usual intake and energy correction algorithm are presented in the supplementary materials.[30-31]

Testing of the new SQFFQ

The revised SQFFQ questionnaire was received following the food item screening, and the first SQFFQ1, follow-up SQFFQ2, and related 24hDR surveys were done.

The average daily intake of various foods and nutrients for each subject that corresponded to the two surveys (SQFFQ1 and SQFFQ2) was calculated based on the findings of the SQFFQ and 24hDR carried out in the community C range, and the validity and reliability of retest were assessed.

Additionally, the SQFFQ2 results' validity was examined in light of the 24hDR survey's energy and nutrient consumption. One of them uses the usual intake to calculate the nutrients for the 24hDR evaluation.

Statistical analysis

STATA 9.2 is used for statistical analysis in this case. When expressing descriptive data, quantities or percentages are used, the mean and standard deviation of nutrient consumption is used. Pearson, Spearman, Kendall, and CCC assessed the reliability of retesting between two FFQs, and the stronger correlation may to some extent lessen the within-person variance of the two FFQ surveys. Additionally, Pearson, Spearman, Kendall, and CCC may assess the validity of the test between the FFQ and 24hDR that combines the energy adjustment technique with the notion of usual intake under the presumption of two nutrient distribution assumptions.

When the variables of interest are continuous, the CCC indicator, which is one of them, is more suited to monitoring consistency than other indicators since it can assess how reproducible a new set of data is with the original observations. The state that exhibits more efficacy when the sample size is small and outliers are present can be supplemented by the Kendall rank correlation test, which is also a supplement.

The individuals' energy and nutrient intake were categorized into three levels based on the mean of the SQFFQ and a 24hDR. The results of the 24hDR and SQFFQ2 were then compared, along with the results of the nutrient distribution tests for SQFFQ1 and SQFFQ2. The continuous variables are then divided into three equal intervals, and the percentage of participants' energy and nutrients that fall into the same, adjacent, and extreme triple partitions is calculated. The outcome is shown as a percentage of participants

RESULTS

Food item screening and correction results for initial FFQ

Multiple regression analysis for screening food items (MRA)

The optimal subset and cumulative R2 value of food types corresponding to each nutrient are shown in Supplementary Table 3.

Cumulative contribution analysis to screen food items (CA)

The results of the nutrient contribution rate of various foods are presented in Supplementary Table 4.

Food Items with Dominant Food Modification Screening

The proportion of 24hDR consumption in absolute terms for each particular food category are presented in Supplementary Table 5.

Results of the New SQFFQ's Effectiveness Tests

Dietary nutrients mean and standard deviation for SQFFQ1, SQFFQ2, and 24hDR

Before and after data with 24hDR and SQFFQ2 energy correction (Repeat and missing follow-up were removed)

The results of 24hDR and SQFFQ2 energy correction are presented in Supplementary Table 6.

Before and after data with SQFFQ1 and SQFFQ2 energy correction (Repeat and missing follow-up were removed)

The results of SQFFQ1 and SQFFQ2 energy correction are presented in Supplementary Table 7.

Reliability test

The results of tests linked to Pearson, Spearman, Kendall, and CCC using the nutritional intake of SQFFQ1 and SQFFQ2 are displayed in the Table 2.

The daily intake of micronutrients like "Na, K, vitamins" and other nutrients with less daily intake was distributed below 0.60 in the Pearson test of dietary daily intake of nutrients (non-energy-corrected non-logarithmic conversion) between SQFFQ1 and SQFFQ2. However, nutrients with higher daily intake like "energy, protein, fat, saturated fatty acids, monounsaturated fatty acids, polyunsaturated fatty acids, carbohydrates" were distributed between 0.60 and 0.75.

The correlation coefficients of the Pearson, Spearman, Kendall, and CCC tests will all increase significantly in the dietary daily nutrient intake (logarithmic treatment without energy correction) data, despite the fact that the correlation coefficients between the Kendall, CCC, and Pearson tests will be significantly smaller than those between the Pearson and Spearman tests. Additionally, CCC (consistency test) values may be greater than 0.60.

The Pearson or CCC correlation coefficient between SQFFQ1 and SQFFQ2 (with energy adjustment, without logarithmic conversion) for nutrients with higher daily intake of "energy, protein, fat, saturated fatty acids, monounsaturated fatty acids, polyunsaturated fatty acids, carbohydrates" will be found to remain above 0.50 after using energy adjustment;

However, due to energy-adjusted operations, the dietary daily nutrient intake (with energy adjustment, without logarithmic conversion) data between SQFFQ1 and SQFFQ2 with lower daily nutrient intake will be significantly reduced, but the CCC test correlation coefficient will remain at a level that is comparable to dietary daily nutrient intake.

Additionally, the CCC correlation coefficient of nutrients with higher daily intake was increased and significant, in contrast to the pre-energy adjusted data, while the Pearson correlation coefficient of dietary daily nutrient intake (with energy adjustment and logarithmic conversion) was lower than that of dietary daily nutrient intake (with energy adjustment, without logarithmic conversion).

Presentation of the third-equal reliability test findings

The results of the third-equal reliability test are presented in Table 3.

The distribution of nutrients across the same, adjacent, or extreme tertiary partitions is represented in the outcome, which is reported as a percentage of participants after comparing the first and second SQFFQs and categorizing the continuity variable into three equal intervals.

Participants were split into the same third, with daily intakes ranging from 30-40% for low intake to 60-70% for high intake. Over 85% of participants, on average, fall into the same or nearby categories. It is clear that the energy-adjusted data will see a proportional reduction of the same one-third.

Validity test

Calculate the Pearson, Spearman, Kendall, and CCC correlation tests, as well as the Pearson, Spearman, Kendall, CCC correlation tests following the computation of the decay coefficient, to determine the validity of the comparison between the findings of the SQFFQs and the 24hDR. Table 4 presents the result of the validity of the comparison between the findings of the SQFFQs and the 24hDR.

It is clear from the data that most dietary nutrients have high test validity before energy adjustment. And while the correlation coefficient before or after the decay calculation is still high for the data before energy adjustment, whether the usual intake for the non-logarithmic start calculation or the usual intake for the initial logarithmic calculation, the improvement of the data using the decay coefficient is minimal due to the relatively small variance.

It should be observed that the correlation coefficient of the test does not allow the usual intake of either the beginning logarithmic calculation or the non-logarithmic starting calculation after energy adjustment to be statistically significant.

For instance, the CCC test result for the non-logarithmic starting calculation of the usual intake (assumption B) dietary daily nutrient intake (with energy adjustment, without logarithmic conversion) protein is 0.035, while the Pearson test result is 0.015 for the calculation of the usual intake (assumption B) dietary daily nutrient intake (with energy adjustment and logarithmic conversion) protein.

Additionally, the CCC test result for vitamin A (with energy adjustment, without logarithmic conversion) at the beginning of the logarithmic calculation (assumption A), and the CCC test result for vitamin A (with energy adjustment and logarithmic conversion) at the beginning of the logarithmic calculation (assumption A) for nutrients with low daily intake was 0.119 and 0.119, respectively.

It might be that the residual essence is really employed to eliminate the influence on the nutrients owing to the linear relationship between energy and nutrients, and the constant is added to assure visibility. This is because the adjustment given to energy is the correction utilizing residuals. Energy adjustment cannot be carried out since the short-term 24-hour dietary review data only reflect transient nutritional information and do not reflect the linear relationship between energy and other nutrients in the long-term dietary performance of individual participants. As can be seen from the tables above, no obvious link was established in any nutrient to support this after energy adjustment for the usual intake based on the development of two assumptions (assumption B, assumption A).

As a result, in the validation process, the option "2. without energy adjustment, with logarithmic conversion" in Assumption B and Assumption A shows only very slight differences. This is due to the double logarithmic calculations. When we keep only three decimal places, the statistical results of the two assumptions are essentially the same. For the options "3. with energy adjustment, without logarithmic conversion; 4. with energy adjustment and logarithmic conversion" in Assumption B and Assumption A during validation, energy correction is actually unnecessary. The reason is that the introduction of 24hDR usual intake is meant to compensate for the weight of low-intake nutrients. If the original data of low-intake nutrients of 24hDR is adjusted based on energy before calculating the usual intake, the low-intake nutrients, which are already affected by the one-sidedness of the 24hDR investigation, will be pushed even closer to zero, leading to significant distortion. Therefore, energy correction in the calculation of 24hDR usual intake is redundant. Regarding the statistical assumptions related to the two types of usual intakes, whether logarithmized or not, they are mainly focused on ensuring the stability of the distribution variance of low-intake nutrients.

Presentation of the third-equal validity test findings

The distribution of nutrients across the same, adjacent, or extreme tertiary partitions was compared between SQFFQ and 24hDR, and the continuity variable was separated into three equal intervals. The result was represented as a percentage of participants. The results of the comparison of same, adjacent, or extreme tertiary partitions between SQFFQ and 24hDR are in Table 5.

It is clear that the distribution and reliability test results are comparable.

DISCUSSION

In this study, we assessed the SQFFQ's validity and retest reliability in order to examine adult dietary consumption of energy and nutrients in Xi'an, Shaanxi. The ultimate objective is to have a SQFFQ that most people can easily and quickly fill out, and studies have indicated that it takes participants less than half an hour to complete one on average. As a result, it can be said that this SQFFQ may be used to quickly assess dietary patterns, and modify and improve dietary patterns in various research settings.

The efficiency of SQFFQ is comparatively good given that the correlation coefficients found in this study are greater or within the range of numbers reported in other prior studies.

A 1-month gap between two surveys is thought to be sufficient to minimize the unintended effects of seasonal food and yearly changes in personal dietary habits, according to earlier studies on the reliability testing.[27-28, 32]

In terms of validity testing, numerous 24h dietary recalls are often performed to confirm SQFFQ for the 24h dietary recall that was used in testing.

Commonly utilized 24-hour multiple recalls' average values nevertheless reflect substantial internal variances between individuals.[25-26]

This study offered a mix of "usual intake" and "pseudo data" methodologies, avoiding the unpredictability of prior 24-h dietary recalls repeated two to three times." Usual intake" can represent the subject's long-term intake of a specific nutrient.[33]

Based on the findings of the dietary review, the "pseudo data" will create each subject's 365-day daily food intake. It will then compute the "usual intake" under the presumption that the "pseudo data" is an unbiased estimate of the "usual intake".

The 24-hour dietary recall can more accurately reflect actual dietary results when "usual intake" and "pseudo data" are used together. The 24-hour dietary recall for validity testing was often carried out the second day following SQFFQ in earlier research.

Nevertheless, whether the 24-hour dietary recall was performed prior to or following the SQFFQ, the time gap between the two assessments may cause interference. Additionally, the 24-hour dietary recall measurement created by fusing "usual intake" and "pseudo data" may have different cycles than the SQFFQ measurement, which can help to more effectively eliminate interference between two cycles.[33]

According to earlier studies, the 24-hour dietary recall and the second SQFFQ (SQFFQ2) often had better levels of consistency. This could be as a result of the fact that completing the first SQFFQ (SQFFQ1) actually enables individuals to better comprehend their food consumption and communicate it.[27-28, 32]

The test results in this study may have some limitations due to the small sample size.Some claim that the minimal sample size for validity testing is 100 when the survey technique, such as a 24-hour dietary recall, has fewer repetitions. A significant correlation of 0.4 between the 24-hour dietary recall and the FFQ (SQFFQ) needs a minimum sample size of 30 individuals, which is less than the sample size used in this investigation.[28-29]

According to various research, two repeats are needed for 24-hour dietary recalls in order to achieve statistical efficacy that is high enough to completely exclude the influence of within-person variation.[27-28] The combination of "usual intake" and "pseudo data" in this study, in our opinion, can also successfully reduce within-person variation.

FFQ (SQFFQ) and 24-hour dietary recall have a generally accurate association, however there are some discrepancies that can be seen. We hypothesize that the use of a composite nutritional database to determine FFQ (SQFFQ) and food intake for a 24-hour dietary recall may be the cause. When building the relationship between food and nutrients, due to the updating and supplementation of various foods in the Chinese dietary database in different years, appropriate foods should be replaced from the nutrient databases of other years and other countries and regions, and they should be cross checked to ensure that they are within the similar range of corresponding nutrients. Priority should be given to the Chinese dietary database source data if there is a large discrepancy in the related nutrients between two separate databases of foods.

In addition to the previously mentioned factors, the intrinsic characteristics of a 24-hour dietary recall may also affect the outcomes. The findings of the 24-hour dietary recall are still only transitory nutritional data, which cannot accurately reflect the long-term eating patterns of specific people, even after processing using "usual intake and pseudo data." In reality, energy adjustment eliminates the effects of the linear relationship between energy and nutrients on nutrients by utilizing the essence of residuals. Energy adjustment produces incorrect findings if energy data cannot accurately reflect long-term eating patterns. The validity testing after energy adjustment was therefore shown to be invalid, which might account for this point.

Prior to food item and option correction, the initial FFQ for food items was based on the inclusion and classification of food items using dietary survey scales from several areas in China, and it had reference to years of domestic dietary recommendations in China. As a result, the original food item in this study can account for the fact that most foods that may be included in the FFQ are generally considered to be universal, and the modified SQFFQ stays the same. The initial food items can take into account Xi'an's eating pattern because it is a socially and economically developed area of northwest China.

The present study endeavors to pioneer a novel and relatively versatile approach for constructing dietary questionnaires. When this method is employed in other regions, it will give rise to entirely new dietary questionnaires that feature compact scale architectures and enhanced acceptability. It should be noted that there will inevitably be certain disparities in the food items of individuals within diverse dietary contexts. For some particular groups of people, this method can be given a try. If, after the initial FFQ construction and FFQ item screening in the early stage, the quantity of food items is not excessively abundant, it implies that there exists a certain degree of objective "commonality" in the diet among these specific populations. Here, the "commonality" pertains to the resemblance in diet among individuals within a specific regional dietary and cultural milieu. Consequently, the questionnaire construction method of this study can be suitably applied.

Conclusion

This study's semi-quantitative food frequency questionnaire (SQFFQ) was designed to look at dietary habits in the Shaanxi Province region of Xi'an. This is based on statistical data from numerous literature on reliability and validity testing. The results are comparable to the testing outcomes of other earlier FFQ (SQFFQ) studies. It is believed to have good reliability and validity after testing and can continue to be used as an effective tool for further investigating large amounts of nutrient intake in various research context.[27-29]

CONFLICT OF INTEREST AND FUNDING DISCLOSURE

No potential conflict of interest was reported by the authors.

This work was supported by the Special scientific research project for health care of the whole army (Vol.20BJZ12).

REFERENCES

1. Khan J, Gul P and Liu K. Grains in a Modern Time: A Comprehensive Review of Compositions and Understanding Their Role in Type 2 Diabetes and Cancer. Foods. 2024; 13:2112. doi: 10.3390/foods13132112.
2. Bishayee A, Penn A and Bhandari N, et al. Dietary plants for oral cancer prevention and therapy: A review of preclinical and clinical studies. Phytother Res. 2024; 38:5225-5263. doi: 10.1002/ptr.8293.
3. Wyatt M, Choudhury A and Von Dohlen G, et al. Randomized control trial of moderate dose vitamin D alters microbiota stability and metabolite networks in healthy adults. Microbiol Spectr. 2024; 12:e8324. doi: 10.1128/spectrum.00083-24.
4. Ungvari Z, Fekete M and Fekete JT, et al. Adherence to the Mediterranean diet and its protective effects against colorectal cancer: a meta-analysis of 26 studies with 2,217,404 participants. Geroscience. 2024. doi: 10.1007/s11357-024-01296-9.
5. Barron K, Ford CD and Dunlap S. Exploring the cardiovascular health behaviors of college students and the areas that need increased public health efforts. Public Health Nurs. 2024; 41:1331-1341. doi: 10.1111/phn.13407.
6. Hu Y, Sun Q and Han Y, et al. Role of lifestyle factors on the development and long-term prognosis of pneumonia and cardiovascular disease in the Chinese population. Chinese Med J. 2024. doi: 10.1097/CM9.0000000000003160.

7. Pourmontaseri H, Sepehrinia M and Kuchay MS, et al. The association between energy-adjusted dietary inflammatory index and metabolic syndrome and its mediatory role for cardiometabolic diseases: a prospective cohort study. Front Nutr. 2024; 11:1429883. doi: 10.3389/fnut.2024.1429883.
8. Millar SR, Harrington JM and Perry IJ, et al. Ultra-processed food and drink consumption and lipoprotein subclass profiles: A cross-sectional study of a middle-to older-aged population. Clin Nutr. 2024; 43:1972-1980. doi: 10.1016/j.clnu.2024.07.007.
9. Expression of Concern: The effects of DASH diet on weight loss and metabolic status in adults with non-alcoholic fatty liver disease: A randomized clinical trial. Liver Int. 2024; 44:3114. doi: 10.1111/liv.16084.
10. Zhu W. Effective roles of exercise and diet adherence in non-alcoholic fatty liver disease. World J Gastroentero. 2024; 30:3456-3460. doi: 10.3748/wjg.v30.i29.3456.
11. Yang Y, Chen Y and Jia X, et al. Association of dietary flavonoid intake with the prevalence and all-cause mortality of depressive symptoms: Findings from analysis of NHANES. J Affect Disord. 2024; 366:44-58. doi: 10.1016/j.jad.2024.08.150.
12. Jebeile H, Baur LA and Kwok C, et al. Symptoms of Depression, Eating Disorders, and Binge Eating in Adolescents With Obesity: The Fast Track to Health Randomized Clinical Trial. Jama Pediatr. 2024; 178:996-1005. doi: 10.1001/jamapediatrics.2024.2851.
13. Ogden CL, Ansai N and Fryar CD, et al. Depression and diet quality, US adolescents and young adults: National Health and Nutrition Examination Survey, 2015-March 2020. J Acad Nutr Diet. 2024; 125:247-255. doi: 10.1016/j.jand.2024.08.007.
14. Lugon G, Hernaez A and Jacka FN, et al. Association between different diet quality scores and depression risk: the REGICOR population-based cohort study. Eur J Nutr. 2024; 63:2885-2895. doi: 10.1007/s00394-024-03466-z.
15. Seelarbokus BA, Menozzi E and Schapira A, et al. Mediterranean Diet Adherence, Gut Microbiota and Parkinson's Disease: A Systematic Review. Nutrients. 2024; 16:2181. doi: 10.3390/nu16142181.
16. Godos J, Micek A and Currenti W, et al. Fish consumption, cognitive impairment and dementia: an updated dose-response meta-analysis of observational studies. Aging Clin Exp Res. 2024; 36:171. doi: 10.1007/s40520-024-02823-6.
17. Sanz-Martos AB, Roca M and Plaza A, et al. Long-term saturated fat-enriched diets impair hippocampal learning and memory processes in a sex-dependent manner. Neuropharmacology. 2024; 259:110108. doi: 10.1016/j.neuropharm.2024.110108.
18. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019; 393:1958-1972. doi: 10.1016/S0140-6736(19)30041-8.
19. Forouhi NG and Unwin N. Global diet and health: old questions, fresh evidence, and new horizons. Lancet. 2019; 393:1916-1918. doi: 10.1016/S0140-6736(19)30500-8.
20. Margetts BM and Nelson M. Measuring dietary exposure in nutritional epidemiological studies. Nutr Res Rev. 1995; 8:165-178. doi: 10.1079/NRR19950011.

21. Randall E, Marshall J and Graham S, et al. Frequency of food use data and the multidimensionality of diet. J Am Diet Assoc. 1989; 89:1070-1075.
22. Stefanik PA and Trulson MF. Determining the frequency intakes of foods in large group studies. Am J Clin Nutr. 1962; 11:335-343. doi: 10.1093/ajcn/11.5.335.
23. Willett WC, Sampson L and Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol. 1985; 122:51-65. doi: 10.1093/oxfordjournals.aje.a114086.
24. Pietinen P, Hartman AM and Haapa E, et al. Reproducibility and validity of dietary assessment instruments. II. A qualitative food frequency questionnaire. Am J Epidemiol. 1988; 128:667-676. doi: 10.1093/oxfordjournals.aje.a115014.
25. Willett WC. Dietary assessment methods. Br J Nutr. 1995; 74:141-143. doi: 10.1079/bjn19950113.
26. Imaeda N, Fujiwara N and Tokudome Y, et al. Reproducibility of a semi-quantitative food frequency questionnaire in Japanese female dietitians. J Epidemiol. 2002; 12:45-53. doi: 10.2188/jea.12.45.
27. Lee MS, Pan WH and Liu KL, et al. Reproducibility and validity of a Chinese food frequency questionnaire used in Taiwan. Asia Pac J Clin Nutr. 2006; 15:161-169.
28. Marques-Vidal P, Ross A and Wynn E, et al. Reproducibility and relative validity of a food-frequency questionnaire for French-speaking Swiss adults. Food Nutr Res. 2011; 55. doi: 10.3402/fnr.v55i0.5905.
29. Deschamps V, de Lauzon-Guillain B and Lafay L, et al. Reproducibility and relative validity of a food-frequency questionnaire among French adults and adolescents. Eur J Clin Nutr. 2009; 63:282-291. doi: 10.1038/sj.ejcn.1602914.
30. Xia PF, Zhang YB and Liu G, et al. The application of energy adjustment models in nutritional epidemiology. Zhonghua Yu Fang Yi Xue Za Zhi. 2020; 54:228-232. doi: 10.3760/cma.j.issn.0253-9624.2020.02.022.
31. Willett WC, Howe GR and Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr. 1997; 65:1220S-1231S. doi: 10.1093/ajcn/65.4.1220S.
32. Kim J, Kim Y and Ahn YO, et al. Development of a food frequency questionnaire in Koreans. Asia Pac J Clin Nutr. 2003; 12:243-250.
33. Tooze JA, Kipnis V and Buckman DW, et al. A mixed-effects model approach for estimating the distribution of usual intake of nutrients: the NCI method. Stat Med. 2010; 29:2857-2868. doi: 10.1002/sim.4063.

Table 1. Initial FFQ food standards

Category	Food
Cereals	
Rice	Rice (cooked), white porridge (cooked), millet, cornmeal, glutinous rice, brown rice, rye, black rice, purple rice, oats, barley, buckwheat, quinoa
Flour	Noodles, steamed buns, flower rolls, cakes
Other cereals	Churros, bread, biscuits
Potato	Potatoes, yams, konjac, sweet potatoes
Soybean	Vegetarian chicken, dried tofu, tofu
Fresh dark vegetables	Green peppers, carrots, tomatoes, spinach
Fresh light vegetables	Radishes, cucumbers, cabbage
Fresh mushrooms	Shiitake mushrooms, oyster mushrooms, grass mushrooms
Pickled vegetables	Squeezed vegetables, pickled snow mushrooms, salted mustard heads, sauce radish, sauce cucumber, tofu milk
Fresh fruit	Bananas, watermelons, peaches, grapes, apples, pears, oranges, oranges
Red meat	Raw pork, raw lamb, raw beef
Poultry	Raw duck, raw chicken, raw goose
Processed meat products	Luncheon meat, skin sausage, sausage, sauce meat, ham
River fresh	Blue carp, sea bass, silver carp, river shrimp, grass carp
Seafood	Yellow fish, sea shrimp, ribbon fish, pomfrets, crabs
Nuts	Peanuts, almonds, pecans, walnuts, pistachios, sunflower seeds
Legumes	
Whole	Phaseolus vulgaris, lentils, cowpeas (long beans)
Pure beans	Mung beans, red beans, broad beans, sprouts, peas
Onions and garlic	Green onion, garlic, ginger, onion, leek
Fresh algae	Kelp, wakame, seaweed
Dairy products	Milk, yogurt
Fruit juice	Peaches, grapes, apples, pears, oranges, oranges
Sugary drinks	Sugary drinks
Tea drinks	Green tea, black tea, flower tea
Animal oil	Lad, butter, duck fat, sheep fat
Vegetable oil	Peanut oil, soybean oil, rapeseed oil, sesame oil, cottonseed oil, sesame oil, sunflower oil
Salt	Edible salt
Eggs	Eggs, duck eggs
Liquor	Beer, cider, liquor
Cooking sugars	Brown sugar, white sugar, cottony sugar, rock sugar
Liquid soy products	Soy milk

Table 2. Reliability of SQFFQ1 - SQFFQ2 comparison

	Reliability							
	1				2			
	Pearson	Spearman	Kendall	CCC	Pearson	Spearman	Kendall	CCC
Energy (kJ)	0.67	0.86	0.60	0.31	0.88	0.86	0.60	0.69
Protein (g)	0.68	0.73	0.55	0.28	0.74	0.73	0.55	0.81
Fat (g)	0.69	0.62	0.44	0.32	0.67	0.62	0.44	0.85
Saturated fatty acids (g)	0.69	0.55	0.40	0.36	0.59	0.55	0.40	0.90
Monounsaturated fatty acids (g)	0.72	0.53	0.37	0.37	0.72	0.53	0.37	0.88
Polyunsaturated fatty acids (g)	0.57	0.59	0.43	0.25	0.63	0.59	0.43	0.83
Cholesterol (mg)	0.49	0.37	0.25	0.55	0.43	0.37	0.25	0.97
Carbohydrates (g)	0.50	0.68	0.51	0.39	0.70	0.68	0.51	0.73
Total dietary fiber (g)	0.26	0.36	0.25	0.20	0.38	0.36	0.25	0.68
Insoluble dietary fiber (g)	0.27	0.46	0.33	0.08	0.48	0.46	0.33	0.47
Vitamin A (μg)	0.38	0.35	0.25	0.49	0.37	0.35	0.25	0.83
Vitamin K (μg)	0.09	0.05	0.03	0.34	0.08	0.05	0.03	0.71
Vitamin E (mg)	0.36	0.52	0.37	0.29	0.55	0.52	0.37	0.70
Vitamin B1 (mg)	0.63	0.72	0.54	0.19	0.74	0.72	0.54	0.50
Vitamin B2 (mg)	0.51	0.63	0.48	0.31	0.65	0.63	0.48	0.86
Vitamin B6 (mg)	-0.01	0.17	0.12	0.43	0.14	0.17	0.12	0.78
Niacin (mg)	0.67	0.64	0.47	0.34	0.65	0.64	0.47	0.83
Vitamin C (mg)	0.16	0.28	0.19	0.28	0.22	0.28	0.19	0.62
Folic acid (μg)	0.20	0.24	0.17	0.20	0.24	0.24	0.17	0.59
Choline (mg)	0.41	0.09	0.07	0.25	0.25	0.17	0.12	0.83
Sodium (mg)	0.45	0.37	0.25	0.19	0.44	0.37	0.25	0.63
Potassium (mg)	0.51	0.61	0.44	0.21	0.64	0.61	0.44	0.64
Calcium (mg)	0.26	0.48	0.35	0.25	0.53	0.48	0.35	0.78
Zinc (mg)	0.55	0.70	0.54	0.23	0.72	0.70	0.54	0.74
Selenium (μg)	0.68	0.73	0.55	0.30	0.73	0.73	0.55	0.91
Phosphorus (mg)	0.57	0.76	0.59	0.24	0.76	0.76	0.59	0.73
Magnesium (mg)	0.40	0.65	0.48	0.14	0.68	0.65	0.48	0.59
Iron (mg)	0.53	0.74	0.56	0.18	0.75	0.74	0.56	0.66
Iodine (μg)	-0.03	0.24	0.15	0.01	0.10	0.24	0.15	0.66
Copper (mg)	0.57	0.71	0.53	0.19	0.72	0.71	0.53	0.75
Manganese (mg)	0.24	0.53	0.37	0.27	0.54	0.53	0.37	0.79
Moisture(g)	0.05	0.07	0.06	0.14	0.04	0.07	0.06	0.86

[†]Statistical analysis methods: Reliability testing based on Pearson, Spearman, Kendall, and CCC methods using the nutritional intake of SQFFQ1 and SQFFQ2.
[‡]1: without energy adjustment or logarithmic conversion; 2: without energy adjustment, with logarithmic conversion; 3: with energy adjustment, without logarithmic conversion; 4: with energy adjustment and logarithmic conversion.

Table 2. Reliability of SQFFQ1 - SQFFQ2 comparison (cont.)

	Reliability							
	3				4			
	Pearson	Spearman	Kendall	CCC	Pearson	Spearman	Kendall	CCC
Protein (g)	0.56	0.39	0.25	0.37	0.67	0.39	0.25	0.48
Fat (g)	0.26	0.11	0.08	0.34	0.49	0.11	0.08	0.39
Saturated fatty acids (g)	0.27	0.11	0.08	0.68	0.44	0.11	0.07	0.67
Monounsaturated fatty acids (g)	0.11	0.05	0.03	0.51	0.25	0.05	0.03	0.57
Polyunsaturated fatty acids (g)	0.29	0.30	0.20	0.29	0.32	0.29	0.20	0.47
Cholesterol (mg)	0.24	0.26	0.18	0.86	0.34	0.26	0.18	0.53
Carbohydrates (g)	0.25	0.08	0.08	0.35	0.10	0.08	0.08	0.73
Total dietary fiber (g)	-0.03	-0.07	-0.05	0.32	0.07	-0.07	-0.05	0.67
Insoluble dietary fiber (g)	-0.29	-0.16	-0.11	0.14	-0.25	-0.15	-0.11	0.38
Vitamin A (μg)	0.08	0.06	0.03	0.55	0.45	0.06	0.03	0.64
Vitamin K (μg)	0.15	-0.03	-0.03	0.31	-0.06	-0.03	-0.03	0.72
Vitamin E (mg)	0.10	0.10	0.07	0.33	0.06	0.10	0.07	0.70
Vitamin B1 (mg)	0.18	0.06	0.04	0.09	0.04	-0.03	-0.03	0.11
Vitamin B2 (mg)	0.30	0.21	0.15	0.57	0.13	0.09	0.07	0.72
Vitamin B6 (mg)	-0.24	-0.19	-0.13	0.40	-0.04	-0.23	-0.16	0.75
Niacin (mg)	0.42	0.29	0.20	0.48	0.45	0.29	0.20	0.65
Vitamin C (mg)	-0.06	0.02	0.02	0.29	-0.01	0.02	0.02	0.57
Folic acid (μg)	0.12	-0.02	-0.02	0.23	0.04	-0.02	-0.02	0.55
Choline (mg)	0.20	0.09	0.05	0.23	0.21	0.07	0.05	0.93
Sodium (mg)	0.27	0.26	0.18	0.33	0.19	0.26	0.18	0.34
Potassium (mg)	0.17	0.02	0.00	0.27	0.49	0.02	0.00	0.34
Calcium (mg)	0.04	-0.08	-0.06	0.40	0.07	-0.08	-0.06	0.63
Zinc (mg)	0.31	0.18	0.12	0.28	0.41	0.18	0.11	0.43
Selenium (μg)	0.25	0.25	0.18	0.47	0.28	0.25	0.18	0.39
Phosphorus (mg)	0.29	0.15	0.10	0.22	0.57	0.15	0.10	0.28
Magnesium (mg)	-0.12	0.00	0.00	0.17	0.17	0.00	0.00	0.26
Iron (mg)	0.20	0.11	0.07	0.19	0.43	0.11	0.07	0.35
Iodine (μg)	-0.14	-0.01	0.00	0.01	0.01	-0.01	0.00	0.39
Copper (mg)	0.23	0.12	0.07	0.24	0.16	0.11	0.06	0.38
Manganese (mg)	-0.14	-0.11	-0.08	0.59	-0.09	-0.11	-0.09	0.69
Moisture(g)	-0.09	-0.08	-0.06	0.34	-0.14	-0.08	-0.06	0.73

†Statistical analysis methods: Reliability testing based on Pearson, Spearman, Kendall, and CCC methods using the nutritional intake of SQFFQ1 and SQFFQ2.
‡1: without energy adjustment or logarithmic conversion; 2: without energy adjustment, with logarithmic conversion; 3: with energy adjustment, without logarithmic conversion; 4: with energy adjustment and logarithmic conversion.

Table 3. Third-equal reliability test

	Reliability											
	1			2			3			4		
	same tertiles %	adjacent tertiles %	extreme tertiles %	same tertiles %	adjacent tertiles %	extreme tertiles %	same tertiles %	adjacent tertiles %	extreme tertiles %	same tertiles %	adjacent tertiles %	extreme tertiles %
Energy (kJ)	41.5	58.5	0.00	41.5	58.5	0.00						
Protein (g)	55.4	40.0	4.62	55.4	40.0	4.62	43.1	43.1	13.8	43.1	43.1	13.8
Fat (g)	50.8	46.2	3.08	50.8	46.1	3.08	30.8	49.2	20.0	30.8	49.2	20.0
Saturated fatty acids (g)	44.6	49.2	6.15	44.6	49.2	6.15	38.5	40.0	21.5	38.5	40.0	21.5
Monounsaturated fatty acids (g)	52.3	40.0	7.69	52.3	40.0	7.69	38.5	36.9	24.6	38.5	36.9	24.6
Polyunsaturated fatty acids (g)	60.0	30.8	9.23	60.0	30.8	9.23	36.9	49.2	13.9	36.9	49.2	13.9
Cholesterol (mg)	43.1	43.1	13.8	43.1	43.1	13.8	46.2	36.9	16.9	46.2	36.9	16.9
Carbohydrates (g)	64.6	27.7	7.69	64.6	27.7	7.69	35.4	43.1	21.5	35.4	43.1	21.5
Total dietary fiber (g)	56.9	27.7	15.4	56.9	27.7	15.4	30.7	46.2	23.1	30.7	46.2	23.1
Insoluble dietary fiber (g)	46.2	43.1	10.7	46.2	43.1	10.7	33.9	36.9	29.2	33.9	36.9	29.2
Vitamin A (μg)	44.6	43.1	12.3	44.6	43.1	12.3	27.7	52.3	20.0	27.7	52.3	20.0
Vitamin K (μg)	35.4	41.5	23.1	35.4	41.5	23.1	29.2	43.1	27.7	29.2	43.1	27.7
Vitamin E (mg)	47.7	43.1	9.23	47.7	43.1	9.23	47.7	24.6	27.7	47.7	24.6	27.7
Vitamin B1 (mg)	61.5	33.9	4.62	61.5	33.9	4.62	35.4	43.1	21.5	30.8	46.2	23.0
Vitamin B2 (mg)	61.5	33.9	4.62	61.5	33.9	4.62	46.1	33.9	20.0	40.0	40.0	20.0
Vitamin B6 (mg)	38.5	46.2	15.3	38.5	46.2	15.3	24.6	49.2	26.2	24.6	46.2	29.2
Niacin (mg)	56.9	33.9	9.23	56.9	33.9	9.23	49.2	36.9	13.9	49.2	36.9	13.9
Vitamin C (mg)	41.5	46.2	12.3	41.5	46.2	12.3	30.8	46.2	23.0	30.8	46.2	23.0
Folic acid (μg)	38.5	46.2	15.3	38.5	46.2	15.3	27.7	43.1	29.2	27.7	43.1	29.2
Choline (mg)	44.6	30.8	24.6	41.5	36.9	21.6	47.7	32.3	20.0	33.9	53.8	12.3
Sodium (mg)	38.5	43.1	18.4	38.5	43.1	18.4	43.1	36.9	20.0	43.1	36.9	20.0
Potassium (mg)	47.7	46.1	6.15	47.7	46.1	6.15	30.8	43.1	26.1	30.8	43.1	26.1
Calcium (mg)	53.9	36.9	9.23	53.9	36.9	9.23	24.6	46.2	29.2	24.6	46.2	29.2
Zinc (mg)	61.5	33.9	4.62	61.5	33.9	4.62	40.0	40.0	20.0	40.0	40.0	20.0
Selenium (μg)	55.4	40.0	4.62	55.4	40.0	4.62	36.9	46.2	16.9	36.9	46.2	16.9
Phosphorus (mg)	50.8	46.1	3.08	50.8	46.1	3.08	36.9	43.1	20.0	36.9	43.1	20.0
Magnesium (mg)	52.3	40.0	7.69	52.3	40.0	7.69	29.2	46.2	24.6	29.2	46.2	24.6
Iron (mg)	55.4	43.1	1.54	55.4	43.1	1.54	40.0	36.9	23.1	40.0	36.9	23.1
Iodine (μg)	35.4	55.4	9.23	35.4	55.4	9.23	36.9	49.2	13.9	36.9	49.2	13.9
Copper (mg)	55.4	40.0	4.62	55.4	40.0	4.62	32.3	46.2	21.5	32.3	46.2	21.5
Manganese (mg)	63.1	27.7	9.23	63.1	27.7	9.23	26.2	49.2	24.6	26.2	49.2	24.6
Moisture(g)	27.7	50.8	21.5	27.7	50.8	21.5	26.2	49.2	24.6	26.2	49.2	24.6

†Statistical analysis methods: Third-equal reliability test.

‡1: without energy adjustment or logarithmic conversion;2:without energy adjustment, with logarithmic conversion;3:with energy adjustment, without logarithmic conversion;4:with energy adjustment and logarithmic conversion.

Table 4. Validity of SQFFQ2 - 24hDR comparison

	Validity-Assumption B							
	1				2			
	Pearson	Spearman	Kendall	CCC	Pearson	Spearman	Kendall	CCC
Energy (kJ)	0.640	0.450	0.321	0.833	0.588	0.450	0.321	0.772
Calcium (mg)	0.700	0.621	0.442	0.787	0.658	0.621	0.442	0.715
Copper (mg)	0.575	0.498	0.333	0.729	0.570	0.498	0.333	0.702
Iron (mg)	0.660	0.581	0.414	0.844	0.619	0.581	0.414	0.704
Iodine (μg)	0.571	0.436	0.314	0.789	0.417	0.436	0.314	0.885
Potassium (mg)	0.631	0.636	0.460	0.785	0.615	0.636	0.460	0.716
Magnesium (mg)	0.744	0.602	0.426	0.740	0.627	0.602	0.426	0.683
Manganese (mg)	0.649	0.379	0.246	0.795	0.528	0.379	0.246	0.783
Sodium (mg)	0.695	0.437	0.321	0.716	0.558	0.437	0.321	0.782
Phosphorus (mg)	0.683	0.557	0.398	0.776	0.641	0.557	0.398	0.670
Selenium (μg)	0.599	0.457	0.323	0.756	0.531	0.457	0.323	0.676
Zinc (mg)	0.684	0.482	0.342	0.769	0.605	0.482	0.342	0.678
Saturated fatty acids (g)	0.619	0.389	0.280	0.810	0.423	0.389	0.280	0.842
Insoluble dietary fiber (g)	0.571	0.561	0.392	0.720	0.589	0.561	0.392	0.673
Monounsaturated fatty acids (g)	0.638	0.327	0.229	0.848	0.429	0.406	0.285	0.861
Cholesterol (mg)	0.512	0.406	0.285	0.737	0.417	0.413	0.307	0.632
Choline (mg)	0.378	0.413	0.307	0.604	0.389	0.350	0.255	0.758
Protein (g)	0.630	0.408	0.273	0.768	0.539	0.408	0.273	0.671
Polyunsaturated fatty acids (g)	0.571	0.478	0.344	0.717	0.551	0.478	0.344	0.842
Moisture(g)	0.504	0.552	0.397	0.805	0.527	0.552	0.397	0.850
Carbohydrates (g)	0.647	0.414	0.300	0.846	0.609	0.414	0.300	0.756
Vitamin A (μg)	0.499	0.477	0.325	0.885	0.432	0.477	0.325	0.830
Vitamin B1 (mg)	0.502	0.446	0.319	0.788	0.564	0.446	0.319	0.717
Vitamin B2 (mg)	0.639	0.570	0.406	0.814	0.618	0.570	0.406	0.741
Vitamin B6 (mg)	0.332	0.398	0.270	0.906	0.376	0.398	0.270	0.874
Vitamin C (mg)	0.512	0.392	0.266	0.935	0.407	0.392	0.266	0.889
Vitamin E (mg)	0.704	0.525	0.390	0.887	0.572	0.525	0.390	0.794
Vitamin K (μg)	0.308	0.353	0.258	0.726	0.248	0.342	0.247	0.976
Niacin (mg)	0.590	0.336	0.239	0.803	0.448	0.336	0.239	0.666
Folic acid (μg)	0.566	0.450	0.310	0.894	0.369	0.450	0.310	0.891
Fat (g)	0.652	0.397	0.277	0.826	0.523	0.397	0.277	0.790
Total dietary fiber (g)	0.715	0.592	0.415	0.884	0.557	0.592	0.415	0.731

† Statistical analysis methods: Validity testing based on Pearson, Spearman, Kendall, and CCC methods using the nutritional intake of SQFFQ2 and 24hDR.
‡ Assumption B:Non-logarithmic calculation of usual intake; Assumption A: The initial stage begins to calculate the logarithmicly.
§ 1: without energy adjustment or logarithmic conversion; 2: without energy adjustment, with logarithmic conversion; 3: with energy adjustment, without logarithmic conversion; 4: with energy adjustment and logarithmic conversion

Table 4. Validity of SQFFQ2 - 24hDR comparison (cont.)

	Validity-Assumption A							
	1				2			
	Pearson	Spearman	Kendall	CCC	Pearson	Spearman	Kendall	CCC
Energy (kJ)	0.453	0.450	0.321	0.012	0.588	0.450	0.321	0.012
Calcium (mg)	0.492	0.621	0.442	0.149	0.658	0.621	0.442	0.151
Copper (mg)	0.575	0.498	0.333	0.729	0.570	0.498	0.333	0.702
Iron (mg)	0.651	0.581	0.414	0.818	0.619	0.581	0.414	0.693
Iodine (μg)	0.570	0.436	0.314	0.770	0.417	0.436	0.314	0.885
Potassium (mg)	0.434	0.636	0.460	0.027	0.615	0.636	0.460	0.028
Magnesium (mg)	0.498	0.602	0.426	0.148	0.627	0.602	0.426	0.179
Manganese (mg)	0.649	0.379	0.246	0.795	0.528	0.379	0.246	0.783
Sodium (mg)	0.440	0.437	0.321	0.068	0.558	0.437	0.321	0.106
Phosphorus (mg)	0.458	0.557	0.398	0.076	0.641	0.557	0.398	0.072
Selenium (μg)	0.561	0.457	0.323	0.647	0.531	0.457	0.323	0.602
Zinc (mg)	0.680	0.482	0.342	0.761	0.605	0.482	0.342	0.674
Saturated fatty acids (g)	0.619	0.389	0.280	0.809	0.423	0.389	0.280	0.842
Insoluble dietary fiber (g)	0.571	0.561	0.392	0.719	0.589	0.561	0.392	0.673
Monounsaturated fatty acids (g)	0.638	0.327	0.229	0.848	0.429	0.406	0.285	0.861
Cholesterol (mg)	0.409	0.406	0.285	0.094	0.417	0.413	0.307	0.081
Choline (mg)	0.376	0.402	0.305	0.604	0.381	0.341	0.253	0.758
Protein (g)	0.552	0.408	0.273	0.573	0.539	0.408	0.273	0.534
Polyunsaturated fatty acids (g)	0.570	0.478	0.344	0.716	0.551	0.478	0.344	0.842
Moisture(g)	0.461	0.552	0.397	0.627	0.485	0.552	0.397	0.856
Carbohydrates (g)	0.530	0.414	0.300	0.401	0.609	0.414	0.300	0.312
Vitamin A (μg)	0.457	0.477	0.325	0.122	0.432	0.477	0.325	0.103
Vitamin B1 (mg)	0.639	0.570	0.406	0.814	0.618	0.570	0.406	0.741
Vitamin B2 (mg)	0.332	0.398	0.270	0.906	0.376	0.398	0.270	0.874
Vitamin B6 (mg)	0.475	0.392	0.266	0.826	0.407	0.392	0.266	0.769
Vitamin C (mg)	0.704	0.525	0.390	0.886	0.572	0.525	0.390	0.793
Vitamin E (mg)	0.308	0.353	0.258	0.726	0.248	0.342	0.247	0.976
Vitamin K (μg)	0.584	0.336	0.239	0.794	0.448	0.336	0.239	0.658
Niacin (mg)	0.565	0.450	0.310	0.899	0.369	0.450	0.310	0.890
Folic acid (μg)	0.633	0.397	0.277	0.776	0.523	0.397	0.277	0.760
Fat (g)	0.715	0.592	0.415	0.884	0.557	0.592	0.415	0.731
Total dietary fiber (g)	0.715	0.592	0.415	0.884	0.557	0.592	0.415	0.731

†Statistical analysis methods: Validity testing based on Pearson, Spearman, Kendall, and CCC methods using the nutritional intake of SQFFQ2 and 24hDR.
‡Assumption B:Non-logarithmic calculation of usual intake; Assumption A: The initial stage begins to calculate the logarithmicly.
§1: without energy adjustment or logarithmic conversion; 2: without energy adjustment, with logarithmic conversion; 3: with energy adjustment, without logarithmic conversion; 4: with energy adjustment and logarithmic conversion

Table 3. Third-equal validity test

	Validity-Assumption A and B					
	1			2		
	same tertiles %	adjacent tertiles %	extreme tertiles %	same tertiles %	adjacent tertiles %	extreme tertiles %
Energy (kJ)	35.4	52.1	12.5	35.4	52.1	12.5
Calcium (mg)	52.1	45.8	2.08	52.1	45.8	2.08
Copper (mg)	41.7	50.0	8.33	41.7	50.0	8.33
Iron (mg)	54.2	37.5	8.33	54.2	37.5	8.33
Iodine (μg)	43.8	45.8	10.4	43.8	45.8	10.4
Potassium (mg)	60.4	33.3	6.25	60.4	33.3	6.25
Magnesium (mg)	47.9	47.9	4.17	47.9	47.9	4.17
Manganese (mg)	39.6	47.9	12.5	39.6	47.9	12.5
Sodium (mg)	52.1	37.5	10.4	52.1	37.5	10.4
Phosphorus (mg)	54.2	37.5	8.33	54.2	37.5	8.33
Selenium (μg)	37.5	50.0	12.5	37.5	50.0	12.5
Zinc (mg)	50.0	41.7	8.33	50.0	41.7	8.33
Saturated fatty acids (g)	45.8	41.7	12.5	45.8	41.7	12.5
Insoluble dietary fiber (g)	50.0	43.7	6.25	50.0	43.7	6.25
Monounsaturated fatty acids (g)	45.8	37.5	16.7	45.8	37.5	16.7
Cholesterol (mg)	27.1	60.4	12.5	27.1	60.4	12.5
Choline (mg)	52.1	37.5	10.4	37.5	54.2	8.33
Protein (g)	45.8	43.8	10.4	45.8	43.8	10.4
Polyunsaturated fatty acids (g)	54.2	33.3	12.5	54.2	33.3	12.5
Moisture(g)	50.0	41.7	8.33	50.0	41.7	8.33
Carbohydrates (g)	39.6	50.0	10.4	39.6	50.0	10.4
Vitamin A (μg)	50.0	41.7	8.33	50.0	41.7	8.33
Vitamin B1 (mg)	45.8	41.7	12.5	45.8	41.7	12.5
Vitamin B2 (mg)	56.2	37.5	6.25	56.2	37.5	6.25
Vitamin B6 (mg)	47.9	41.7	10.4	47.9	41.7	10.4
Vitamin C (mg)	50.0	37.5	12.5	50.0	37.5	12.5
Vitamin E (mg)	54.2	37.5	8.33	54.2	37.5	8.33
Vitamin K (μg)	43.8	45.8	10.4	43.8	45.8	10.4
Niacin (mg)	37.5	50.0	12.5	37.5	50.0	12.5
Folic acid (μg)	45.8	50.0	4.17	45.8	50.0	4.17
Fat (g)	41.7	41.7	16.6	41.7	41.7	16.6
Total dietary fiber (g)	54.1	41.7	4.17	54.1	41.7	4.17

†Statistical analysis methods: Third-equal validity test

‡Assumption B:Non-logarithmic calculation of usual intake; Assumption A: The initial stage begins to calculate the logarithmicly

§1: without energy adjustment or logarithmic conversion; 2: without energy adjustment, with logarithmic conversion; 3: with energy adjustment, without logarithmic conversion; 4: with energy adjustment and logarithmic conversion

Table 6. Characteristics of participants

	Questionnaire 1	Questionnaire 2	Questionnaire 3	
	Community A	Community B	Community C	Community C
	24hDR/ Initial FFQ	Revised FFQ/ 24hDr	SQFFQ1	SQFFQ2/ 24hDR
Age (year)	38.2±12.2	39.2±12.9	41.7±12.9	41.7±12.9
Gender-male	35 (53.9)	59 (56.7)	24 (48.0)	24 (48.0)
Gender-female	30 (46.1)	45 (43.3)	26 (52.0)	26 (52.0)
BMI (kg/m^2)	21.2±1.67	21.2±1.48	21.0±1.69	23.0±2.19
Education				
University (junior college and above)	21 (32.3)	24 (23.1)	10 (20.0)	10 (20.0)
Senior high school (including secondary vocational school)	11 (16.9)	26 (25.0)	10 (20.0)	10 (20.0)
Junior high school	11 (16.9)	33 (31.7)	12 (24.0)	12 (24.0)
Primary school and below	22 (33.9)	21 (20.2)	18 (36.0)	18 (36.0)
Monthly income				
RMB 4000 and below	15 (23.1)	29 (27.9)	17 (34.0)	15 (30.0)
RMB 4001-8000	14 (21.5)	18 (17.3)	6 (12.0)	6 (12.0)
RMB 8001-12000	13 (20.0)	35 (33.7)	16 (32.0)	17 (34.0)
RMB 12001 and above	23 (35.4)	22 (21.1)	11 (22.0)	12 (24.0)

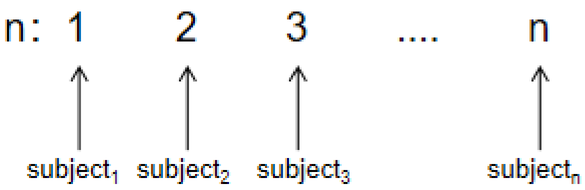

Figure 1. Ideal situation

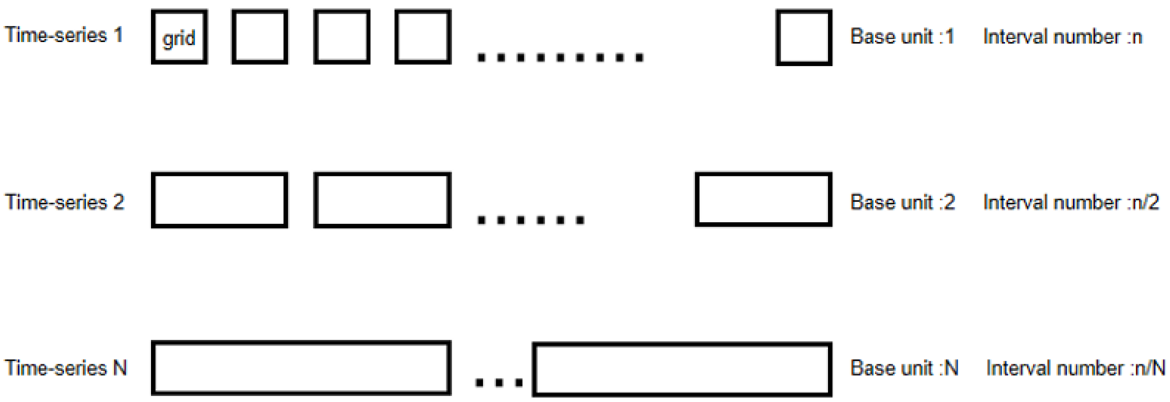

Figure 2. Time series

Figure 3. Taking k=N as an example

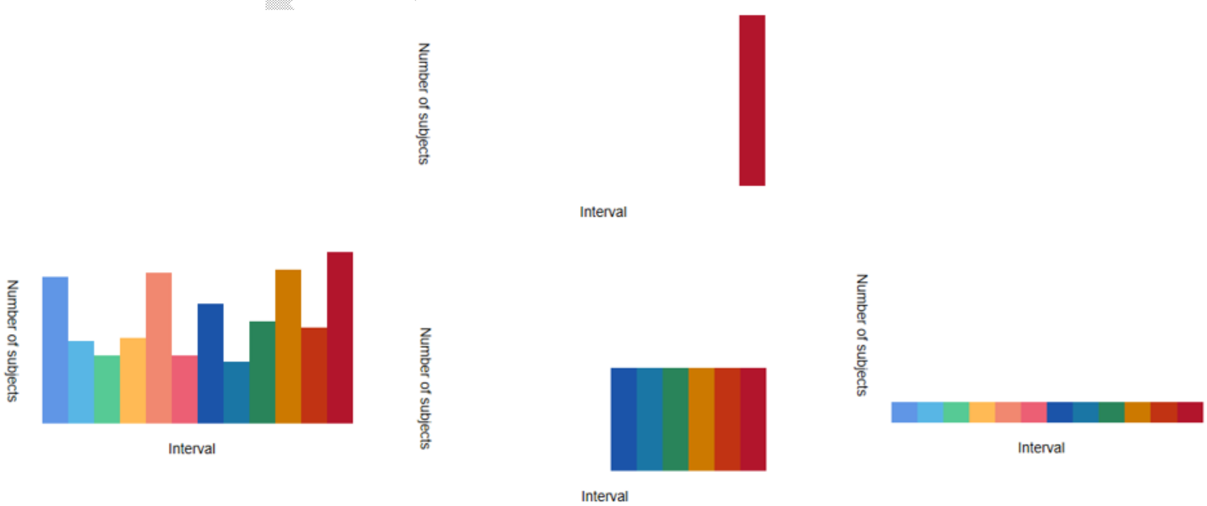

Figure 4. xxxxxx