Original Article

Development of food frequency questionnaires and a nutrient database for the Prospective Urban and Rural Epidemiological (PURE) pilot study in South India: Methodological issues

Ankalmadagu Venkatasubbareddy Bharathi MSc[1], Anura Vishwanath Kurpad PhD[1], Tinku Thomas PhD[1], Salim Yusuf PhD[2], Govindachar Saraswathi PhD[3] and Mario Vaz MD[1]

[1]*St. John's Research Institute, St. John's National Academy of Health Sciences, Bangalore, India.*
[2]*Population Health Research Institute, McMaster University, Hamilton, Canada.*
[3]*Department of Studies in Food Science and Nutrition, University of Mysore, Mysore, India.*

Purpose: To develop Food Frequency Questionnaires (FFQs) and nutrient databases for urban and rural Indian populations with diverse dietary habits for the PURE (Prospective Urban and Rural Epidemiological) pilot study. **Procedure:** 24 hour dietary recalls were obtained from 84 rural and 60 urban subjects. From a comprehensive food list, separate FFQs were developed for the two groups. Nutrient analysis of the FFQ required the selection of foods, development of recipes and application of these to cooked foods to develop a nutrient database. The FFQs were piloted in 80 urban and 77 rural subjects. Separately for each group, a stepwise regression method was used to identify foods contributing to a cumulative 90 % of variance to total energy intake. Nutrient and food group intakes were compared using an independent t-test. **Main Findings:** The urban and the rural FFQs contained 129 and 102 foods respectively, of which 82 foods were common to both. Fourteen urban foods and eight rural foods explained a cumulative 90% of variance for total energy intake. Daily intakes for most nutrients and food groups were two to three fold higher in the urban than in the rural group. **Conclusions:** In Indian populations with diverse dietary habits, using standard methods to develop separate FFQs can capture dietary intakes adequately. To develop nutrient databases, substitution of local food composition tables with data from other sources using standard methods to match foods can be adopted.

Key Words: Databases, nutrient assessment, India, epidemiologic methods, questionnaires

INTRODUCTION

Various methods such as single or multiple 24 hour dietary recalls, weighed diet records, self reports of diet history and Food Frequency Questionnaires (FFQ) have been used to assess dietary intakes in populations.[1-3] None of these methods has particularly high accuracy in determining food intake.[4,5] Studies, may have measurement errors from random misreporting or systemic reporting bias.[6-8] Nevertheless, FFQs have been used as the method for long term diet assessment in many epidemiological studies.[9-12] Benefits of using FFQs over other dietary methods are that they are relatively simple in construct and easy to administer. They also have a lower responder burden when capturing long term habitual intakes in epidemiological studies.[13]

The FFQ structure has as its backbone, a food list containing foods that are eaten reasonably often by an appreciable part of the population and which contain substantial amounts of the nutrients of interest, the intake of which can vary from individual to individual. In India, evaluation of the food lists of published FFQs developed for regions such as Kerala, Gujarat and Lucknow show that less than 20% of foods are similar across the FFQs due to regional variations in food habits.[14-16] Hence in India, a single FFQ is not likely to capture the variations in dietary intakes in populations with different dietary habits, unless a very long food list is used which becomes impractical. This paper describes the issues involved in developing FFQs and the nutrient database for an urban and a rural population with diverse dietary habits for the PURE (Prospective Urban and Rural Epidemiological) Pilot Study being conducted in South India.

The PURE Pilot Study seeks to prospectively evaluate the impact of societal and individual level risk factors, including diet and physical activity, in the genesis of coronary heart disease through the development of a paired urban and rural cohort. The urban cohort is located

Corresponding Author: Bharathi AV, St. John's Research Institute, St. John's National Academy of Health Sciences, Opp BDA Complex, Koramangala, Bangalore – 560034, Karnataka, India.
Tel/Fax: +91 080 25532037
Email: bharathi@iphcr.res.in; bharathiiphcr@yahoo.com
Manuscript received 19 May 2007. Initial review completed 29 May 2007. Revision accepted 5 July 2007.

at Bangalore, a city in South India and the rural cohort comprises a cluster of villages in Palamner Taluk in Chittor district, Andhra Pradesh about 140 km away from Bangalore.

MATERIALS AND METHODS

Subjects

To generate food lists required to develop FFQs, dietary habits were obtained from the subjects from one village in Palamner Taluk, Chittor District, which falls within the area where the FFQ will be applied for the PURE study. Majority of the subjects were Hindus, the predominant community in this population, and also comprised of Muslims. Adults of varying age groups that included both males and females were contacted to obtain their dietary habits. All the subjects contacted participated in this study. Subjects from the urban areas were participants of other research studies, majority of whom were Hindus and from various socio economic groups. This was deemed necessary since the PURE study in its urban setting aimed to evaluate individuals across a wide socio-economic stratum.

The FFQs subsequently developed was piloted in the population that was part of the larger cohort study. Thus, the urban subjects were recruited from the housing colony of the Hindustan Aeronautics Limited factory, a public sector undertaking with over 13,000 employees in Bangalore. The housing colony provided accommodation to a wide range of employees (8 'workmen' categories and 9 'officer' categories that included clerks and general managers). The rural subjects were recruited from two villages in Palamner Taluk, in Andhra Pradesh State, the same geographical location of the rural part of the PURE study. The participants were from villages that were accessible by mud roads, without a permanent health clinic, but with access to a primary school and within 2 km from a sealed road. 95% of urban and rural subjects who were initially contacted participated in the study. The rural sample was 80% Hindu, 20% Muslim while the urban sample was 88% Hindu, 10% Christian and 1% Muslim. All subjects gave informed consent. The study was approved by the Institute Ethical Review Board of St. John's Medical College and Hospital in compliance with the recommendations of the Indian Council of Medical Research.

Preparation of the food lists:

The food list for the rural FFQ was obtained from 24 hour dietary recalls in 84 individuals. The urban food list was developed using 180 days of diet recalls from a three day diary obtained from 60 urban subjects. All the subjects reported portion sizes based on standard household measures which included a glass, ladle, bowl and teaspoon. The dietary habits thus obtained from rural and urban populations were used to develop the food lists.

Methods

Height and weight were recorded in the subjects who participated in a pilot assessment of the FFQ. Height was recorded to the nearest 0.1 cm with the subjects standing erect, head held in the Frankfurt plane, chest at full inspiration, back touching the wall and shoes removed as recommended.[17] Weight was measured to the nearest 0.1 kg on an electronic weighing scale, calibrated with standard weights, with subjects without footwear and with standard indoor clothing. Basic details including age and qualitative information on the diet were provided by the respondent.

Interviewers with at least 12 years of schooling who were familiar with the local languages (Kannada and Telugu) were trained to obtain consent, describe procedures, administer questionnaires and obtain anthropometry. The questionnaires were translated into local languages. Mock sessions were conducted in local languages prior to the actual administration of the questionnaires to ensure that the interviewers were familiar with the food items in the local dialect and to ensure uniformity in the interviewer techniques. FFQs were administered by the same team in the urban and the rural locations.

Development of the Food Frequency Questionnaire

To develop the food list, as dietary diaries and the 24 hour recall methods alone may not capture intake of all foods, other methods listed below were used to obtain a more comprehensive food list:

a) An open question format was used to elicit information on the most common foods consumed during festivals and social gatherings.
b) Nutritionists familiar with local food habits checked the food list for face validity and indicated if any local foods were missed from the 24 hour dietary recalls.
c) Nutritionists also reviewed the Indian Food Composition Tables[18] to include unreported foods for the region that were high sources of energy, protein, fat, carbohydrate, various fatty acids, cholesterol, Vitamin C, thiamin, riboflavin, niacin, β-carotene, calcium, iron, magnesium, sodium, potassium or amino acids.

The food lists obtained were separately compiled for the urban and the rural FFQs into 9 food groups: a) cereals b) legumes / pulses / gravies c) chutneys / salad/ papad d) non-vegetarian foods e) sweets and snacks f) beverages e) miscellaneous additions f) fruits and g) vegetables. Inclusion of foods in the FFQ food list was done separately for the urban and the rural FFQs using the standard approach of:

a) Retaining the independent identity of staple foods and items reported by a large proportion of the people. (e.g. ragi ball with rice and plain ragi ball (a cooked millet, one with rice and the other without) were retained as separate foods in the rural FFQ).
b) Combining similar foods consumed by few people that had similar portion sizes, consumption patterns and nutrient density (e.g. rice preparations with different flavourings were combined as a single food item in the urban and the rural FFQs).
c) Retaining the independent identity of foods where the presence of a particular nutrient of interest was high (e.g. masala dosa and plain dosa (rice pancakes, one with a potato stuffing and the other without) were retained as separate foods as masala dosa has a higher content of carbohydrate than plain dosa).
d) Eliminating foods that did not contribute any of the nutrients of interest (e.g. coconut water).

After application of these methods and constraining it to the food items that were specific to each location, there were 129 items in the urban FFQ and 102 items in the rural FFQ of which 82 items were common to both. The major differences in the urban and the rural food list were related to cereals and non-vegetarian foods. The urban FFQ had 10 additional cereal food items and 11 additional non-vegetarian food items in comparison to the rural FFQ (Table 1). The FFQs were constructed for interviewer based administration in both locations as some of the rural people are illiterate. To obtain an estimate of portion size, subjects were shown common household portions, namely a glass, cup, small teaspoon, large tablespoon and a ladle. These were the same portion sizes used at both urban and rural sites. For each food item the average portion size and the frequency of consumption (per 'day', 'week', 'month', or 'year') were documented. Seasonal availability of foods was estimated through a market survey from five vendors in the urban and the rural areas separately so that nutrient intake of these foods could be estimated. Since the seasonal availability of the fruits reported by the vendors varied, the median value for the number of months the seasonal foods were available was estimated. The FFQs were piloted to assess if all the foods in the FFQ were clear and easily understood by both men and women in various local languages.

Table 1. Number of food items in the urban and the rural Food Frequency Questionnaires.

	Total food items		Common food items†
	Urban	Rural	
Cereals	21	13	11
Lentils / dhals / gravies	7	11	5
Chutney /salads /pickle/papad	12	7	5
Non-vegetarian	19	10	8
Snacks	11	8	7
Desserts / sweets	11	7	5
Beverages	11	13	10
Miscellaneous additions‡	7	5	4
Fruits	16	16	15
Vegetables	14	12	12
Total number of food items	129	102	82

† Number of foods common to both the urban and the rural Food Frequency Questionnaires.; ‡ Includes food additions such as butter, clarified butter, jam, sugar, cheese, ketchup and additional salt.

Table 2. The number of foods obtained from the various food composition tables for individual nutrients in the PURE nutrient database.

	Indian Food Composition Tables	USDA†	McCance and Widdowsons composition of foods	Similar foods‡
All nutrients	1501 (18)	5759 (66)	108 (1)	202 (2)
Energy	102	26	-	-
Calcium	93	33	1	-
Iron	89	36	1	1
Sodium	50	72	1	3
Vitamin C	87	40	1	-
β-carotene	86	-	33	2
Saturated fat	-	119	1	4

† United States Department of Agriculture nutrient database, USDA, Release No.14; ‡ For nutrients that were unavailable from any of the food composition tables, a value from a single or an average from multiple foods that were similar to the food with a missing nutrient was obtained.

DEVELOPMENT OF A NUTRIENT DATABASE

Recipe collection process

In the Indian Food Composition Tables, data on the nutrient content for raw foods for only select nutrients are available. Hence for this study a database of recipes for the foods included in the FFQs was obtained separately from the urban and the rural groups. In the rural areas, only women provided recipes. In the urban areas 90 % of the informants were women. In all instances people who provided the recipes were individuals who routinely cooked these recipes in their households. Recipes were collected for 93 mixed foods for the urban FFQ and 68 mixed foods for the rural FFQ. Although separate recipes were also obtained for the foods common to both FFQs, for 26 of these (32%) the same recipe was used for both locations since the recipes collected were similar in composition and preparation. The recipes that were similar in both locations included: simple cereal preparations, sweets, beverages and raw food preparations such as chutneys. All the recipes were accepted after checking for face validity by consulting other housewives / nutritionists. The food items were cooked in the metabolic kitchen at the research facility according to the recipes provided or in the houses of individuals who provided the recipes. Standard instructions were provided to research assistants and the individuals preparing the recipes. A weighing scale (sensitive to 2g) was used to weigh all the ingredients. Weights were obtained for the edible portion of the foods used in the recipe.

Use of food composition tables

The Indian Food Composition Tables were used to estimate the nutrient content of the raw ingredients reported in the recipes. Except for protein, carbohydrate and energy, all other nutrients were analyzed directly using standard methods. The Indian Food Composition Tables provide nutrient data for raw foods and provide nutrient data on 25 nutrients for all foods. For a few select foods, information on 26 additional nutrients are given including amino acids, fatty acids and Vitamin B_{12}. In addition, the following problems were observed with the Indian Food Composition Tables:

a) Some nutrients that were of interest to the study such as cholesterol, various other fatty acids, Vitamin E etc. were unavailable for most foods or in some instances for all the foods.
b) Some of the raw foods used in some recipes were not listed in the Indian Food Composition Tables.
c) For the nutrients available in the Indian Food Compopostion Tables, nutrient information was 'missing' for some foods.

In these cases, data from the United States Department of Agriculture nutrient database (USDA, Release No.14)[19] and McCance and Widdowsons Composition of Foods[20] were obtained (Table 2). For some unique Indian foods (eg; puffed rice, khoa (a milk product), garam masala etc.) none of the above food composition tables provided a value for some nutrients. For such foods, a value was estimated from either a single similar food or the average of multiple foods if more than one single food item in this group was found to be similar or comparable. Finally, for 11% of the 'missing' nutrient data, a zero value was imputed as the nutrient in these foods was known to be negligible or absent. In the absence of information about the micronutrient losses during food preparation and cooking for Indian foods, losses during food preparation were not taken into account while developing the nutrient database.

Issues associated with the use of multiple food composition tables

Foods were often described differently across different food composition tables. In order to overcome this problem a standard method was adopted for the selection of foods where multiple options were available for a single food. The foods were matched by comparing the scientific names provided in the Indian Food Composition Tables and the USDA tables. As scientific names were unavailable in McCance and Widdowsons Food composition tables, foods were matched in terms of food descriptions and if the differences in energy, fat, protein and carbohydrate differed by less than 10% from the Indian Food Composition Tables, a method similarly adopted elsewhere.[21,22] Where appropriate, the weighed values of fatty acids and amino acids, corrected for total fat and total protein content of the Indian Food Composition Tables, were also calculated and substituted in the nutrient database.

Issues of developing a nutrient database

Nutrient content for cooked recipes was obtained by applying a conversion factor accounting for the weight/ volume change on cooking. The same recipes were also used to estimate the contribution of the ingredients to various food groups such as cereals, pulses, vegetables, fruits, salt, sugar, milk products, eggs or meats. Calculation sheets were developed to convert individual reporting of the foods to obtain daily intakes of the nutrients and food groups using the nutrient database. The daily nutrient or food group intake was calculated by multiplying the intake recomputed for a day with the serving size and the nutrient or food group content per portion of the food item. The nutrients and food groups were estimated for all the foods listed in the FFQ and summed to obtain the total nutrient or food group intake per day for an individual.

From the pilot data, details of the major fats used in cooking obtained from each subject while administering the FFQs showed that the urban and rural households used different cooking oils to prepare the mixed dishes. Four different types of cooking oils were reported in the urban areas with majority of the households reporting sunflower oil (80%), followed by groundnut oil (12%) palm oil (5%) and coconut oil (3%). In the rural areas three different cooking oils were reported with majority reporting groundnut oil (79%) followed by sunflower (16%) and palm oil (5%). Hence separate databases had to be developed for the urban and the rural FFQs as the majority of the mixed dishes contained oils as one of the ingredients. The nutritive value of the cooking oil were recalculated for all the mixed recipes for different types of cooking oils and separate databases were created for sunflower, groundnut, coconut and palm oils, the major oils reported in these populations. Based on the major cooking oil reported by the subject, the appropriate database was used to calculate nutrients.

Statistics

Preliminary analysis using Kolmogorov-Smirnov tests indicated the data to be normally distributed, hence parametric tests were applied. In addition to assessing means and standard deviations of daily intake of nutrients and food groups in the urban and the rural group, other statistical methods were applied to the urban and rural pilot sample separately. First, a stepwise linear regression model was applied to obtain a list of the foods that explained a cumulative 90% variance for the total energy intake in the urban and the rural group. In this model, the total daily intakes of the various foods listed in the FFQ were included as the independent variables and the total daily energy intake of the group as the dependent variable. Second, a contributory analysis was performed by estimating the contribution of each food item to total energy intake in the urban and the rural group. Finally an independent sample t-test was used to compare nutrient and food group intake between the urban and rural groups. Data of 4 subjects were excluded from analysis as their energy intakes estimated from the FFQ were either < 500 kcal or >3500 kcal, a cutoff as suggested in other studies.[23] All statistical analyses were done in SPSS (Version 13.0, SPSS Inc, Ill, Ch) and the level of significance was set at 5 %.

RESULTS

Basic characteristics of the subjects who participated in the pilot study are shown in Table 3. 71% of urban subjects and 55% of rural subjects were females. The weight and body mass index of the urban group was significantly higher as compared to that of the rural group ($p < 0.01$), although the mean heights of the two groups was not significantly different. Ninety three percent of the rural group were either underweight or of normal BMI, while the urban group consisted of 46% normal BMI and 56% overweight or obese subjects.

Multiple linear regressions using the stepwise method

Table 3. Socio demographic characteristics of the urban and the rural pilot study group.(Mean ± SD)

	Urban	Rural
Males / females (no.)	23 / 55	34 / 41
Age (years)	44 ± 10	44 ± 12
Weight (kg)	64 ± 11	50 ± 10**
Height (m)	1.56 ± 0.1	1.58 ± 0.1
BMI (kg / m^{-2})	26 ± 4	20 ± 4**

** Significant difference between urban and rural group $p < 0.01$

Table 4. Contribution of the foods to total daily energy intake in the urban and the rural pilot study group.

Rank	Foods	Cumulative $r^{2\dagger}$	% contribution[‡]	Ranking of food[§]
		Urban (N=78)		
1	Veg & non-veg puff	0.19	0.1	90
2	Idlis	0.35	2.2	10
3	Sambar	0.45	6.0	3
4	Tea	0.55	4.9	6
5	Lime rice, veg pulao, puliyogre etc.	0.62	1.6	13
6	Pizza, Burgers	0.19	0.1	100
7	Lamb, beef – cutlet	0.72	0.0	120
8	Plain cooked polished rice	0.75	18.6	1
9	Ragi ball	0.81	5.7	4
10	Cauliflower	0.84	0.1	101
11	Chaat	0.86	0.1	92
12	Mysore pak, ladoo, etc.	0.87	0.1	84
13	Roti, pulkhas	0.89	0.7	29
14	Chapathi, parathas etc.	0.90	7.6	2
		Rural (N=75)		
1	Mysore pak, Ladoo etc	0.24	0.2	44
2	Ragi balls (mixed with rice)	0.38	33.4	1
3	Plain Polished Rice	0.51	23.4	2
4	Ragi balls	0.66	2.2	5
5	Idli	0.77	0.7	22
6	Nuts	0.83	4.5	3
7	Banana	0.87	1.0	17
8	Vegetable Sambar	0.91	0.6	27

[†] Cumulative r^2 of the step wise regression method; [‡] % contribution to total energy intake; [§]Ranking of the foods based on the contribution to total energy intake. Food descriptions: Veg & non veg puff: A baked snack with either a vegetable or meat filling. Idli: Steamed rice pancakes. Sambar: Cooked lentil preparation. Lime rice, veg pulao, puliyogre etc.: Cooked flavoured rice prepartions. Pizza, Burgers: Fast foods. Ragiball / Ragi ball with Rice: Steamed millet ball with or without rice. Chaat: Snack prepared from puffed rice, potato, masala powder etc. Mysore pak, ladoo, etc.: A Sweet made from chickpea powder. Roti, pulkhas / Chapathi, parathas etc.: Roasted cereal pancake with or without addition of oil. Vegetable Sambar: Cooked mixed vegetable gravy.

showed that 14 urban foods and 8 rural foods explained a cumulative 90 % of the variance for total energy intake (Table 4). The type of the foods identified from the multiple linear regressions varied between the two groups. The foods identified from the regression method included those foods that were reported in the diet recalls as well as those obtained from the additional methods that were used while developing the FFQ. While 14 foods identified in the stepwise regression contributed 48% of the total energy intake in the urban group, eight foods alone contributed 66% of the total energy intake in the rural group

Assessment of daily energy intake in the urban and the rural groups (Table 5) ranged from 782 kcal to 2842 kcal (3272 kJ to 11891 kJ). While the daily total energy intakes were similar in the urban and the rural group, intakes of most other nutrients were higher in the urban group. Carbohydrate intake was the major source of energy in both groups with the rural group reporting 13 % higher intake in comparison to the urban group ($p < 0.01$). Fat contributed to 27 % of the total energy intake in the urban group in comparison to 15 % in the rural group ($p < 0.01$). Among fatty acids, intakes of most fatty acids were higher in the urban group as compared to the rural group ($p < 0.01$). However the urban group reported a high intake of total saturated fat while the rural group reported a higher intake of total monounsaturated fat. Dietary intakes of the urban group for nutrients such as β-carotene, vitamin E, sodium, vitamin C, cholesterol, saturated fat, polyunsaturated fat, Vitamin B_{12} and folate, was two to three fold higher in comparison to the rural group. While the urban group reported consumption of twice the amount of vegetables, milk, sugar, meat and egg in comparison to the rural group, intake of cereals were however higher in the rural group. (Table 6) ($p < 0.01$).

DISCUSSION

In this study, we describe the process used for the development of separate FFQs for an urban and a rural population in South India with very diverse food habits. We recognized that there were certain limitations in the methods that were used to develop the FFQs. Among them was whether the food lists that were developed would reflect the dietary habits of the entire population that we aimed to study. It is extremely difficult to obtain a representative sample of the population given the heterogeneity of the populations in terms of socio-demographic groups. For example, the Hindu community in the rural area alone consists of 12 castes and many more sub-castes.[24] In order to address these limitations while developing the FFQs, we used additional methods that would ensure that we developed a food list that would be inclusive rather than exclusive. This included the use of an open question format to obtain the foods consumed during festivals and social gatherings, consulting experts familiar with the local food habits and use of local food composition tables. In the pilot sample, nutrient and food group intakes

Table 5. Comparison of the daily nutrient intakes of the urban and the rural pilot study group.(Mean± SD)

	Urban N=78	Rural N=75
Energy (kcal)	2068 ± 541	1968 ± 539
Energy (kJ)	8653 ± 2264	8234 ± 2255
Protein (g)	63 ± 17	52 ± 16**
Total Fat (g)	61 ± 19	33 ± 19**
Carbohydrate (g)	316 ± 87	365 ± 99**
Total Saturated fat (g)	21 ± 8	10 ± 5**
Total Monounsaturated fat (g)	16 ± 5	12 ± 8**
Total Polyunsaturated fat (g)	19 ± 8	8 ± 5**
Cholesterol (mg)	149 ± 92	62 ± 38**
Fiber (g)	10 ± 4	8 ± 3 **
Vitamin E (mg)	15 ± 7	6 ± 4**
Vitamin C (mg)	150 ± 77	64 ± 40**
Thiamin (mg)	1.4 ± 0.4	2.6 ± 2.4**
Riboflavin (mg)	1.6 ± 0.5	1.0 ± 0.4**
Niacin (mg)	15 ± 5	15 ± 7
Pantothenic acid (mg)	6.3 ± 2	7.1 ± 2**
Vitamin B6 (mg)	2.0 ± 0.6	1.6 ± 0.5**
Folate (mcg)	334 ± 110	177 ± 62**
Vitamin B12 (mcg)	4.0 ± 5.6	1.9 ± 2.2**
β-carotene (mcg)	4719 ± 2600	1531 ± 1187**
Calcium (mg)	1040 ± 347	603 ± 297**
Phosphorous (mg)	1481 ± 393	1163 ± 340**
Iron (mg)	17 ± 6	12 ± 5**
Magnesium (mg)	578 ± 174	545 ± 186
Sodium (mg)	3580± 1212	1377 ± 779**
Potassium (mg)	2685 ± 868	1692 ± 620**
Copper (mg)	2.3 ± 1.0	3.1 ± 2.6**
Manganese (mg)	7.4 ± 3.8	6.8 ± 3.7
Zinc (mg)	9.4 ± 2.6	9.0 ± 2.7

** Independent t-test between the urban and the rural group significant at $p < 0.01$

Table 6. Comparison of daily food-group intakes in the urban and the rural pilot study group.(Mean ± SD)

	Urban N=78	Rural N=75
Cereals (g)	258 ± 81	375 ± 118**
Legumes (g)	51 ± 26	30 ± 13**
Vegetables (g)	146 ± 70	62 ± 36**
Fruits (g)	144 ± 120	97 ± 81**
Milk (g)	356 ± 180	170 ± 129**
Meat (g)	28 ± 30	14 ± 18**
Fat (g)	30 ± 12	11 ± 7**
Sugar (g)	24 ± 20	7 ± 7**
Eggs (g)	16 ± 17	5 ±5**

** Independent t-test between the urban and the rural group significant at $p < 0.01$

obtained from the pilot sample were within reasonable ranges, this suggest that the FFQs were able to capture dietary intakes in the sample fairly adequately and were also able to discriminate intakes between the urban and the rural groups, the main objective of using FFQ's in the PURE study.

A decision to maintain separate food lists for the urban and rural populations was based on observations made while collecting and collating dietary habits to develop the food lists for the FFQs. Inclusion of all foods would have increased the food list and would have 40 % of foods not being consumed by the rural group as against 15% for the urban group. As studies have shown larger food lists lead to overestimation of frequency of intake[25] and administering a larger food list in relation to their dietary patterns would have introduced a disproportionate measurement error in the rural group as compared to the urban group. Hence a decision to maintain separate FFQs for the urban and the rural group were made.

While developing the nutrient database for this study, the Indian Food Composition Tables have been used as the backbone. For 'missing' nutrients and other nutrients that were of interest to the PURE study and unavailable in the Indian Food Composition Tables, data was substituted from other food composition tables. We did recognize that there were alternative Asian food composition tables that could also have provided some nutrients. However rather than collate from many Asian food composition tables we took the approach of using the most comprehensive food composition tables that were available to us, such that we could limit the derivation of nutrients to a few tables. We did this because we recognized that there are variations between nutrient composition tables for the same food linked to the genetic make up of the foods, the analytical procedures used to estimate the nutrients or due to the differences in the description of the foods.[26] With substitution of missing data from the above food composition tables, a zero value was imputed for about 11% of the data where the nutrient was known to be absent in the foods. However, errors associated with the use of multiple food composition tables and the impact of revisions of food composition tables in the estimation of dietary intakes in Indians will need to be reviewed periodically. The underestimation of nutrient contents would have been considerable if the Indian Food Composition Tables alone were to be used, in part due to a large body of missing data.[27]

This study highlights the methodological issues related to the development of FFQs for the PURE pilot study in South India. To minimize the errors associated with the questionnaire development and structure, a standard approach was set up to develop separate FFQs in populations with diverse dietary habits. Problems with using the local food composition tables and the methods adopted to improve the quality of nutrient database for this study have been discussed.

ACKNOWLEDGEMENTS

The Authors would like to thank Dr. A Jacob, Director of Emmaus Swiss Leprosy Project, Palamaner, who facilitated this study, Smitha JT, Anamika K and Jayachitra who supervised field operations for this study. Sumathi Swaminathan provided input and edited the final manuscript.

AUTHOR DISCLOSURES

Ankalmadagu Venkatasubbareddy Bharathi, Anura Vishwanath Kurpad, Tinku Thomas, Salim Yusuf, Govindachar Saraswathi and Mario Vaz, no conflicts of interest.

REFERENCES

1. Xu J, Eilat-Adar S, Loria C, Goldbourt U, Howard BV, Fabsitz RR, Zephier EM, Mattil C, Lee ET. Dietary fat intake and risk of coronary heart disease: the Strong Heart Study. Am J Clin Nutr. 2006;84:894-902.

2. Jakobsen MU, Overvad K, Dyerberg J, Schroll M, Heitmann BL. Dietary fat and risk of coronary heart disease: Possible effect modification by gender and age. Am J Epidemiol. 2004;160:141-149.
3. Soinio M, Laakso M, Lehto S, Hakala P, Ronnemaa T. Dietary fat predicts coronary heart disease events in subjects with type 2 diabetes. Diabetes care. 2003;26:619-624.
4. Livingstone ME, Black AE. Markers of the validity of reported energy intakes. J Nutr. 2003;133:895S -920S.
5. Porrini M, Gentile MG, Fidanza F. Biochemical validation of a self-administered semi-quantitative food-frequency questionnaire. Br J Nutr. 1995;74:323-333.
6. Paul RD, Rhodes DG, Kramer M, Baer DJ, Rumpler WV. Validation of a food frequency questionnaire by direct measurement of habitual ad libitum food intake. Am J Epidemiol. 2005;162:806-814.
7. Bingham SA, Luben R, Welch A, Wareham N, Khwa KT, Day N. Are imprecise methods obscuring a relation between fat and breast cancer? Lancet. 2003;362:212-214.
8. Schatzkin A, Kipnis V, Carroll RJ, Midthune D, Subar AF, Bingham S, Schoeller DA, Troiano RP, Freedman LS.A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. Int J Epidemiol. 2003;32: 1054-1062.
9. Song Y, Sesso HD, Manson JE, Cook NR, Buring JE, Liu S. Dietary magnesium intake and risk of incident hypertension among middle-aged and older US women in a 10-year follow-up study. Am J Cardiol. 2006; 98:1161-21.
10. Liese AD, Gilliard T, Schulz M, D'Agostino RB Jr, Wolever TM. Carbohydrate nutrition, glycemic load, and plasma lipids: the Insulin Resistance Atherosclerosis Study. Eur Heart J. 2007; 28:80-7.
11. Drogan D, Klipstein-Grobusch K, Dierkes J, Weikert C, Boeing H. Dietary intake of folate equivalents and risk of myocardiaol infarction in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Postdam study. Public Health Nutr. 2006; 9:465-71.
12. Tucker KL, Morita K, Qiao N, Hannan MT, Cupples LA, Kiel DP. Colas, but not other carbonated beverages, are associated with low bone mineral density in older women: The Framingham Osteoporosis Study. Am J Clin Nutr. 2006; 84:936-42.
13. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilization of food-frequency questionnaires-a review. Public Health Nutr. 2002; 5:567-87.
14. Herbert JR, Gupta PC, Bhonsle RB, Murti PR, Mehta H, Verghese F, Aghi M, Krishnaswamy K, Mehta FS. Development and testing of a quantitative food frequency questionnaire for use in Kerala, India. Public Health Nutr. 1998; 1:123-130.
15. Hebert JR, Gupta PC, Bhonsle RB, Sinor PN, Mehta H, Mehta FS. Development and testing of a quantitative food frequency questionnaire for use in Gujarat, India. Public Health Nutr. 1999; 2: 39-50.
16. Pandey D, Bhatia V, Boddula R, Singh HK, Bhatia E. Validation and reproducibility of a food frequency questionnaire to assess energy and fat intake in affluent north Indians. Natl Med J India. 2005; 18: 230-235.
17. Lohman TG, Roche AF, Martorell.R. Anthropometric Standardization Reference Manual. Champaign: Human Kinetics Books, 1988.
18. Gopalan C, Ramshastri BV, Balasubramanian SC. Nutritive value of Indian foods. Hyderabad: National Institute of Nutrition, 1989.
19. United States Department of Agriculture. Nutrient data laboratory. http://www.ars.usda.gov/nutrientdata (accessed on 5 January 2006).
20. Welch AA, Unwin ID, Buss DH, Paul AA, Southgate DAT. McCance and Widdowson's The Composition of Foods, 5th edn. Cambridge: Royal Society of Chemistry, 1995.
21. Merchant AT, Dehghan M, Chifamba J, Terera G, Yusuf S. Nutrient estimation from an FFQ developed for a Black Zimbabwean population. Nutr J. 2005; 4: 37.
22. Dehghan M, Al Hamad N, Yusufali A, Nusrath F, Yusuf S, Merchant AT. Development of a semi-quantitative food frequency questionnaire for use in United Arab Emirates and Kuwait based on local foods. Nutr J. 2005; 4: 18.
23. Willet W. Nutritional Epidemiology. Oxford: Oxford University Press, 1990.
24. Ryan JG, Bidinger PD, Prahlad Rao N, Pushpamma P. The determinants of individual diets and nutritional status in six villages of Southern India. Research bulletin No.7. Patancheru, International Crops Research Institute for the Semi-Arid Tropics, 1984.
25. Krebs-Smith SM, Heimendinger J, Subar AF, Patterson BH, Piyonka E. Using food frequency questionnaires to estimate fruit and vegetable intake: association between the number of questions and total intake. J Nutr Educ. 1995; 27: 80-85.
26. Bharathi AV, Vaz M. Dilemmas in planning diets using existing food composition tables. Natl Med J India. 2004; 17: 280.
27. Chen Y, Ahasan H, Parvez F, Howe GR. Validity of a food frequency questionnaire for a large prospective cohort study in Bangladesh. Br J Nutr. 2004; 92: 851-859.

Original Article

Development of food frequency questionnaires and a nutrient database for the Prospective Urban and Rural Epidemiological (PURE) pilot study in South India: Methodological issues

Ankalmadagu Venkatasubbareddy Bharathi MSc[1], Anura Vishwanath Kurpad PhD[1], Tinku Thomas PhD[1], Salim Yusuf PhD[2], Govindachar Saraswathi PhD[3] and Mario Vaz MD[1]

[1]St. John's Research Institute, St. John's National Academy of Health Sciences, Bangalore, India.
[2]Population Health Research Institute, McMaster University, Hamilton, Canada.
[3]Department of Studies in Food Science and Nutrition, University of Mysore, Mysore, India.

印度南部前瞻性城鄉流行病學(PURE)先驅研究飲食頻率問卷之發展及營養素資料庫：方法學議題

目的：為印度飲食習慣不同的都會及鄉村族群的前瞻性城鄉流行病學 (PURE; Prospective Urban and Rural Epidemiological) 先驅研究，發展飲食頻率問卷(FFQs)及營養素資料庫。**程序**：收集 84 名鄉村及 60 名城市研究對象之 24 小時飲食回憶紀錄。從一個完整的食物清單，針對這兩群人，各自發展出一份 FFQs。為了 FFQ 營養素分析的需要，建立了一個營養素資料庫，過程包括食品選擇、食譜研發及應用這些於熟食。以 80 名城市及 77 名鄉村的研究對象，進行這兩份 FFQs 的先驅研究。在兩群人中，各採用逐步迴歸，找出對總熱量攝取的貢獻累積 90%變異數的食物。使用獨立樣本 t 檢定，比較營養素及食物組攝取。**主要結果**：城市及鄉村的 FFQs 分別包括 129 項及 102 項食物，其中有 82 種食物兩者都有。14 項城市食物及 8 項鄉村食物可解釋總熱量攝取 90%累積變異數。城市組的大部份營養素及食物組的每日攝取量，均較鄉村組高出 2 至 3 倍。**結論**：在不同飲食習慣的印度族群中，使用標準方法分別發展出來的 FFQs 可以充分攫取飲食攝取訊息。可採用標準方法找出其他來源相配食物的數據來代替當地的食物成分表，來發展營養素資料庫。

關鍵字：資料庫、營養素評估、印度、流行病學方法、問卷。