Original Article

Reproducibility and construct validity of a food frequency questionnaire for assessing dietary intake in rural and urban Asian Indian adults

Vasudevan Sudha MSc[1], Ranjit Mohan Anjana MD, PhD[2], Parthasarathy Vijayalakshmi MSc[1], Nagarajan Lakshmipriya MSc[1], Natarajan Kalpana MSc[1], Rajagopal Gayathri MSc[1], Rahavan Durga Priyadarshini MSc[1], Hutgikar Madhav Malini MSc[1], Chandrasekaran Anitha MCA[1], Mohan Deepa PhD[3], Sekar Sathish Raj BSc[3], Kumar Parthiban MSc[3], Rajappan Ramakrishnan BSc[3], Gunasekaran Geetha MSc[1], Kamala Krishnaswamy MD[1], Ranjit Unnikrishnan MD, FRCP[2], Viswanathan Mohan DSc[2]

[1]*Department of Foods, Nutrition & Dietetics Research, Madras Diabetes Research Foundation, Chennai, Tamil Nadu, India*
[2]*Department of Diabetology, Madras Diabetes Research Foundation, Chennai, Tamil Nadu, India*
[3]*Department of Epidemiology, Madras Diabetes Research Foundation, Chennai, Tamil Nadu, India*

Background and Objectives: To evaluate the reproducibility and construct validity of the Madras Diabetes Research Foundation FFQ (MDRF-FFQ) with biomarkers for its use in epidemiological settings in India. **Methods and Study Design:** The MDRF-FFQ was administered to 500 participants representing rural and urban areas of 10 Indian states, twice at an interval of 12 months. Reproducibility was assessed using intra cluster correlation coefficients (ICC). Construct validity of carbohydrate and fat intake was assessed using baseline serum lipids by regression analysis. **Results:** Reproducibility as measured by ICC was 0.50-0.77 for saturated fatty acids (SFA) and energy in urban and 0.61-0.72 for protein and SFA in rural areas. The ICC for food groups was 0.53-0.77 for whole grains, fruits and vegetables in urban and 0.50-0.89 for animal foods and whole grains in rural areas. After adjusting for potential confounders, carbohydrate intake was positively associated with serum triglycerides (TG) (β [SE]: +2.3 [0.72] mg/dL; p=0.002) and inversely with high density lipoprotein cholesterol (HDL) (β [SE]:-0.48 [0.12], p<0.001), while dietary fat and SFA (% Energy) were positively associated with HDL, low density lipoprotein (LDL) and total cholesterol and inversely with TG. **Conclusions:** The MDRF-FFQ can be considered as a reliable and valid tool to measure the long-term dietary exposure in respect of macronutrient intakes in Indian populations despite diverse dietary practices.

Key Words: food frequency questionnaire, reproducibility, validity, biomarkers, serum lipid

INTRODUCTION

The increasing burden of non-communicable diseases (NCD) globally as well as in India[1] can be largely explained by the adoption of unhealthy dietary practices[2] consequent on the so-called "nutrition transition". Assessment of long-term dietary habits in the population is an essential first step in devising meaningful nutrition strategies for prevention and management of NCDs. The Food Frequency Questionnaire (FFQ) is the most widely used tool in large nutrition epidemiological studies.[3] The use of validated FFQs enhances elucidation of the relationship of the diet to disease risk.

Dietary intake assessments are challenging, especially in India where there exist diverse cuisines and a wide variety of regional and cultural food habits and practices.[4] However, Indian diets are comprised, in general, of a cereal staple and are meal-based (breakfast, lunch and dinner) in most regions of India. Cereal based food choices form the main courses of the daily meals irrespective of the population's main dietary habits (vegetarian or non-vegetarian).[5-7] Most of the National Surveys conducted in India have used dietary records, dietary recalls or semi-structured interviews to assess the nutritional status and time trends of food and nutrient intake, in rural and urban areas.[8,9] Other noteworthy studies like the

Corresponding Author: Dr RM Anjana, Madras Diabetes Research Foundation & Dr. Mohan's Diabetes Specialities Centre, WHO Collaborating Centre for Non-Communicable Diseases, ICMR Centre for Advanced Research on Diabetes, IDF Centre of Excellence in Diabetes Care, 4, Conran Smith Road, Gopalapuram, Chennai - 600 086, India.
Tel: (9144)4396 8888; Fax: (9144)2835 0935
Email: dranjana@drmohans.com
Manuscript received 05 September 2019. Initial review completed 08 October 2019. Revision accepted 16 December 9
doi: 10.6133/apjcn.202003_29(1).0025

Figure 1. Regions across India where the FFQ data was collected.

Indian Migration study (IMS)[10] Chennai Urban Rural Epidemiology Study (CURES)[7,11] and Prospective Urban Rural Epidemiology (PURE) Study[12] have used validated FFQs for diet-disease risk studies in India.[7,10,12] These studies, however, have their own limitations; for instance, the IMS FFQ was restricted to a selected industrial and rural migrant population while the CURES and PURE developed FFQs separately for urban and rural populations in India.[7,10-12] The lack of a single common FFQ poses challenges in assessing rural and urban diets and their association with chronic disease prevalence. Considering the common availability of many regional food choices in both rural and urban areas[10] today, it was thought worthwhile to develop a single common FFQ for both rural and urban areas of India covering north, south, east, west and northeastern regions.

Reproducibility of an FFQ over longer periods of time (e.g. 12 months) points to the stability of food and nutrient estimates and is a much-needed feature to assess the diet and chronic diseases risk in large epidemiological studies. Biomarkers provide objective assessments, albeit of an intermediate kind, and may minimize subjective dietary measurement error (recall from memory) which often occurs with self -reported dietary intake. The use of 'gold standard' recovery biomarkers such as doubly-labeled water for energy intake and 24 hr urinary sodium measurement for sodium intake are impractical in the epidemiological setting due to the constraints of feasibility and expense.[3] Studies from the West have attempted to validate macronutrient (carbohydrate and fats) intake (as assessed by FFQs) with biochemical markers such as plasma lipids.[13,14] It is known that obese and overweight individuals with higher intake of energy from either fat or carbohydrate or both are likely to have higher blood lipids.[13,14] However, such biomarker-based validation studies have not been carried out in India where carbohydrate and fat together contribute almost 80-90% to the daily energy intake.[11,15]

In view of this, a comprehensive, structured, quantitative national FFQ covering a wide range of commonly consumed rural and urban Indian foods was developed and pre-tested with the help of visual aids by the Madras Diabetes Research Foundation (MDRF). The present study aims to evaluate the reproducibility of the Madras Diabetes Research Foundation FFQ (MDRF-FFQ) over a 12- month period and also to estimate the construct validity for the measurement of carbohydrate and total fat intake using serum lipids as a biomarker in Asian Indian adults from rural and urban areas of all the regions (north, south, east, west and northeast) of India.

METHODS

Study participants

The study was conducted during 7th May 2011 to 5th June 2012. Urban and rural areas of 9 States and 1 Union Territory representing north, south, east, west and north-eastern regions of India were randomly chosen for the study. These were: Karnataka and Tamil Nadu (southern region), Punjab and Chandigarh (Union Territory) from the northern region, Bihar and Jharkhand (eastern region), Gujarat and Maharashtra (western region) and Arunachal Pradesh and Tripura (northeastern region) (Figure 1). The

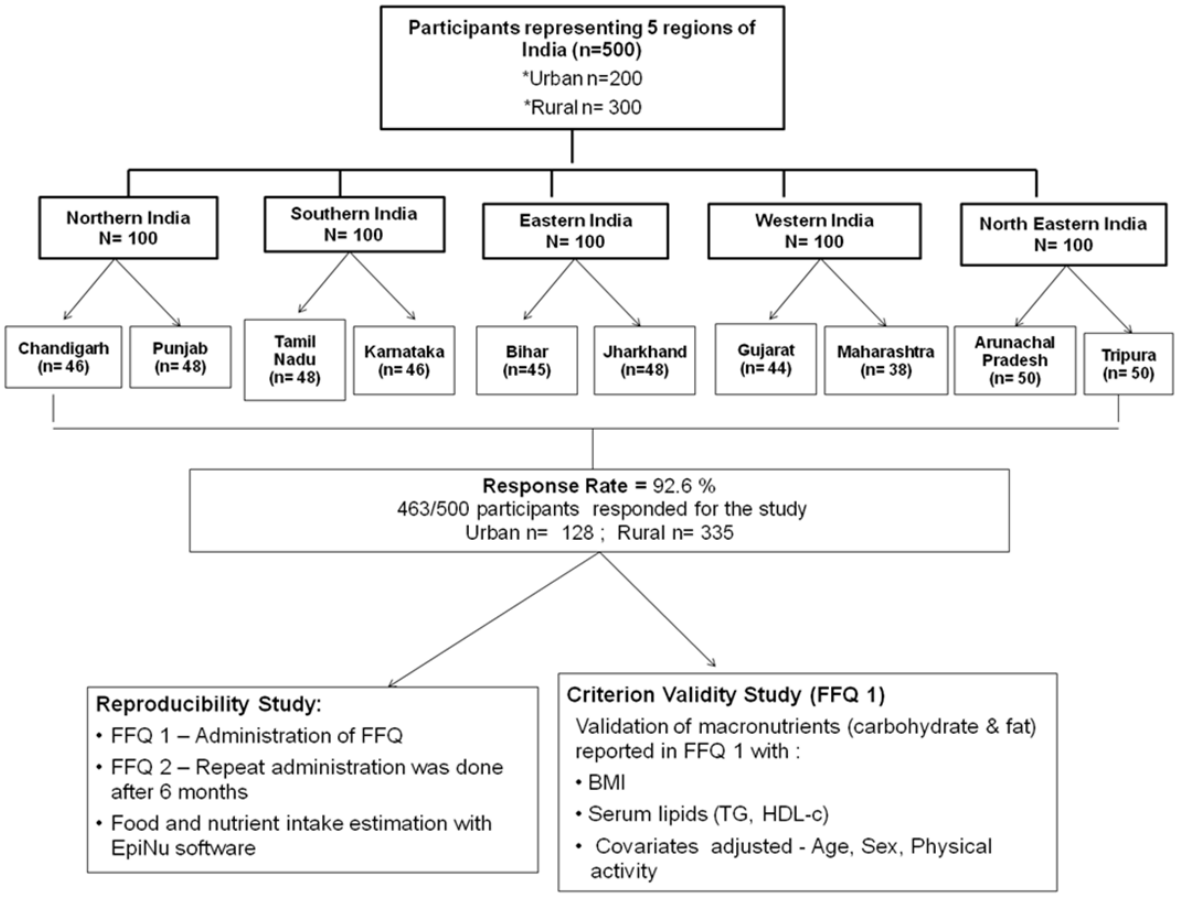

Figure 2. Reproducibility and Validity of MFFQ – Study design.

participants for this study were chosen using a stratified multistage sampling design (similar to the one employed in the India National Family Health Survey-3).[16] From one district from each State/Union Territory, two census enumeration blocks (CEB) in urban areas and three villages in rural areas were randomly selected. From 10 randomly selected households from each selected CEB or village, 1 adult participant (of either sex) was randomly selected (2 CEBs x 10 households = 20; 3 villages x 10 households = 30). Therefore, 50 participants (≥20 years) from each State were selected, making for a total of 500 participants, of whom 463 participated in the study (response rate = 92.6%) (Figure 2). The Institutional Ethics Committee at MDRF approved the study and written informed consent was obtained from each participant before commencing the study. The study is registered in the Clinical Trials Registry of India-(REFCTRI/2008/000174).

Reproducibility and construct validity of MDRF-FFQ

Information on the most frequently consumed foods was separately collected by a data driven approach using 24 h dietary recalls among randomly chosen individuals (not those individuals selected for the present study) from all the selected states. Single 24-hour recall was collected by face-to-face interview during the household visit (December 2010 to February, 2011). The 24-hour recall included either a weekday or a weekend day. The participants were requested to recall all the food and drinks consumed over the last 24 hours in a systematic way (from morning till night) with the help of visual aids. In addition, nutritionists and dietitians from each region were contacted to provide missing regional foods, if any, in the 24 hour recall data. Based on these, an extensive list of foods/food preparation methods was complied. Food items with similar ingredients and method of preparation were grouped together to reduce the length of the questionnaire. Thus, the MDRF-FFQ has a food list containing 236 food items capturing both rural and urban food choices of India.

The food items listed in the MDRF-FFQ were categorized into the usually prepared portion sizes (small / medium / large) and portion utensils (ladle / cup / tsp / tbsp etc.) based on the average weight of the food samples collected from various commercial and non-commercial sources. Frequency of intake was measured using 5 categories ("never," "daily", "weekly", "monthly" and "yearly"). Study participants were asked to report the usual frequency (open ended) as the number of times that best represented their dietary habits over the past one year [for instance a food item consumed 3 times weekly was marked as "3" under the weekly column of the FFQ (using a visual Food Atlas as an aid)].

MDRF-FFQ is an interviewer administered self- reported paper-based questionnaire.[7] Trained interviewers (n=6) with good inter-rater agreement (ICC 0.84) administered the MDRF-FFQ during the period of May 2011 to June 2012. To test reproducibility, the MDRF-FFQs were administered 2 times to the same participants (n=463) at an interval of 12 months (MDRF-FFQ 1 and MDRF-FFQ 2) with the visual aid of photographic Atlas of Indian Foods containing the pictures of various portion sizes.[17]

In this study, the reported carbohydrate and fat intake expressed as percentage of energy were compared with serum lipids-triglycerides; HDL cholesterol, total and low density lipoprotein (LDL) cholesterol respectively for estimating construct validity.

Anthropometric assessments

Anthropometric measurements and blood pressure were assessed using standardized methods. [15] Height (in centimeters) was measured using a stadiometer (SECA Model 214, Seca Gmbh Co, and Hamburg, Germany) and weight (in kilograms) by an electronic weighing scale (SECA Model 807, SecaGmbh Co). Waist circumference was measured with a non-stretchable measuring tape. Individuals were asked to keep both feet together and look straight ahead. The smallest horizontal girth between the iliac crest and the coastal margins at the end of expiration was measured as the waist circumference.[18] Body mass index was calculated using the formula weight in kilograms/height in meters squared. Blood pressure was recorded in the sitting position in the right arm to the nearest 1 mmHg using the electronic OMRON machine (Omron Corporation, Tokyo, Japan). Two readings were taken 5 minutes apart and their mean was taken as the blood pressure.

Biochemical assessments

Blood samples were obtained after 8-10 hours of fasting for biochemical assessments only during the FFQ1 dietary data collection time point. The fasting venous sample was centrifuged within 1 hour of collection at the survey site, and serum was transferred to separate labeled vials and temporarily stored in cold boxes until they were transferred to minus 80◦C freezers in the central laboratory of the Madras Diabetes Research Foundation at Chennai. All the analyses for the study were performed at the National Accreditation Board for Testing and Calibration Laboratories (NABL) and College of American Pathologists (CAP)-accredited central laboratory at Dr Mohan's Diabetes Specialities Centre in Chennai. Two percent of the fasting plasma samples were analyzed for quality control. Accurate coding system was followed to ensure anonymity of samples and also to facilitate tracking of specific samples if the need arose.

Serum cholesterol (cholesterol esterase oxidase-peroxidase-amidopyrine method), serum triglycerides (glycerol phosphate oxidase-peroxidase-amidopyrine method) and HDL cholesterol (direct method polyethylene-glycol-pretreated enzymes) were measured using the Beckman Coulter AU 2700/480 Autoanalyser [Beckman AU (Olympus), Ireland]. LDL cholesterol was calculated using the Friedewald formula.[19] The coefficients of variation for the biochemical assays ranged from 3.1 to 7.6%.[20]

Other assessments

Demographic data and detailed information on smoking and alcohol consumption were collected by trained interviewers. The validated MDRF Physical Activity Questionnaire (MPAQ) was used to assess physical activity levels [PAL (as covariate)] of the participants.[21] The physical activity level category of the participants was determined using the cut-off PAL value.[22] Basal metabolic rate (BMR) of the participants was calculated using age- and sex-specific equations for Indians[23] to test the extent of under-reporters of total energy using <1.2 ratio of EI/BMR as cut off.[24]

Statistical analysis

Analyses were carried out using the statistical analysis software (SAS version 9.0; SAS Institute Inc. Cary, NC). The individual's average daily nutrient intake of the listed food items in the FFQ was computed by multiplying the reported frequency with serving size and per-portion nutrient content using the in-house EpiNu [food and nutrient] database (Version 1 India: Madras Diabetes Research Foundation; 2006). The EpiNU database consists of a collection of a wide range of recipes gathered from different sources. In addition to available data for Indian foods, other nutrient composition tables like United State Department of Agriculture (USDA) database and Malaysian food composition tables were used to ensure the best possible assessment of nutrient data. Similarly, glycaemic index (GI) values of Indian foods were derived from those available in the International GI table in addition to published literature on GI of Indian foods (Epinu 2006).[25] All food groups and nutrients were adjusted for total energy to reduce the measurement error and between-person variation in the food and nutrient intake, using the residual method with total energy intake as the independent variable and the absolute nutrient and food group intake as the dependent variable. Significance of differences between regions was tested using Kruskal Wallis test as the data was not normally distributed and Chi square test was used to test categorical variables. As the data was not normally distributed it was further log transformed. For evaluating the reproducibility of the FFQ, intra-cluster associated correlation coefficients (ICC) were calculated for both nutrients and food groups using FFQ1 and FFQ2 collected at an interval of 12 months and are presented for both rural and urban participants. Nutrients and food group intakes were categorized into quartiles, with the cut-off points for the FFQ1 (original survey) variables also applied to the repeat FFQ2 variables. A 'proc mixed' model was used to compute the adjusted ICC using the SAS. Agreement was tested with Bland Altman analysis[26]for the total energy intake reported in FFQ1 and FFQ2. Multivariate regression analysis was used to assess the construct validity of FFQ using serum lipid (triglycerides and total, LDL and HDL cholesterol) biomarkers collected during the FFQ 1 time point as the dependent variables and carbohydrate (%E); glycemic load, dietary fat (%E), dietary SFA%E as continuous independent variables. Potential confounders like age, sex, BMI, blood pressure, fasting blood glucose, blood cholesterol, LDL-C, physical activity levels, cooking oil, energy intake, refined grains, milk and its products and fruits and vegetables were adjusted in the model. The ratio of energy intake (EI) to basal metabolic rate (BMR)[24] was ascertained as a measure to identify the proportion of under-reporters (EI/BMR ≤1.2) for total energy intake. All tests of significance were two-tailed and a p value of <0.05 was considered significant.

RESULTS

Region-wise demographic, anthropometric and biochemical assessments of the participants are presented in Table 1. Significant differences between regions were seen with respect to body weight, BMI, diastolic blood pressure, serum total cholesterol, LDL-C and physical activity levels.

The region wise intake of nutrients and food groups by the study participants is given in supplementary table. The intake of energy was found to be highest in the North followed by the East, while Northeastern and Southern regions reported the highest intake of carbohydrate and protein. The Southern region reported the highest intake of fat (mainly as saturated fatty acid and poly unsaturated fatty acids while the intake of mono-unsaturated fatty acid was highest in the North. With regard to food groups, the intake of refined cereal and animal foods was reported to be highest in the Northeast, while intake of whole grains, milk and milk products was highest in the North. The intake of fruits, fats and edible oils was found to be highest in the West.

Reproducibility between the two FFQs (FFQ1 and FFQ2) collected from the same participant at a 12 months' interval, assessed using intra-cluster correlation coefficients (ICC) for energy-adjusted nutrients and food groups, and stratified as rural /urban is presented in Table 2.

The energy adjusted ICCs ranged from 0.61 for protein (g/day) to 0.72 for SFA (g/day) for nutrients and from 0.50 for animal foods and fruits to 0.89 for whole grains among food groups in the rural population. The adjusted ICCs between the two FFQs in the urban population ranged from 0.50 for SFA (g/day) to 0.77 for total energy in nutrients and 0.53 for whole grains to 0.77 for fruits and leafy vegetables (g/day) (among food groups). The overall adjusted ICC for nutrients ranged from 0.54 for dietary fiber (g/day) to 0.87 for energy (kcal/day) respectively, while for food groups, it ranged from 0.50 for animal foods (g/day) to 0.75 for pulses and legumes (g/day).

The average energy-adjusted ICCs with respect to sex was 0.69 for nutrients and 0.60 for food groups among males and 0.68 for nutrients and 0.59 for food groups in females (Table 3).

Figure 3 shows the Bland-Altman plot for agreement between FFQ1 and FFQ2 for reported energy intake. The agreement plot revealed heteroscedasticity by visual inspection using residual plot and Kendall's Tau test.30 Accordingly, the data was log transformed and the mean bias and the limits were -0.0037, +0.29 and -0.29.

In the present study, the ratio of self -reported energy intake (EI) from FFQ1 to the basal metabolic rate (BMR) was used to measure the extent of under-reporting of energy intake 3 and the details are presented in Table 4. About 12% of rural adults, 9% of urban adults, 12% male and 10% of the female population were found to be under-reporters in the present study.

Table 5 and 6 show the multivariate adjusted association of Carbohydrates (%E), glycaemic load (GL), dietary fat (%E) and SFA (%E) with the lipid profile after adjusting for potential confounders such as age, sex, literacy, BMI, blood pressure, fasting blood glucose, blood cholesterol, LDL-C, physical activity levels, cooking oil, energy intake, refined grains, milk and milk products and fruits and vegetables. For every unit increase in carbohydrates (%E), there was a significant rise in triglycerides [β (SE): +2.29 (0.72), p=0.002] while HDL cholesterol levels de-

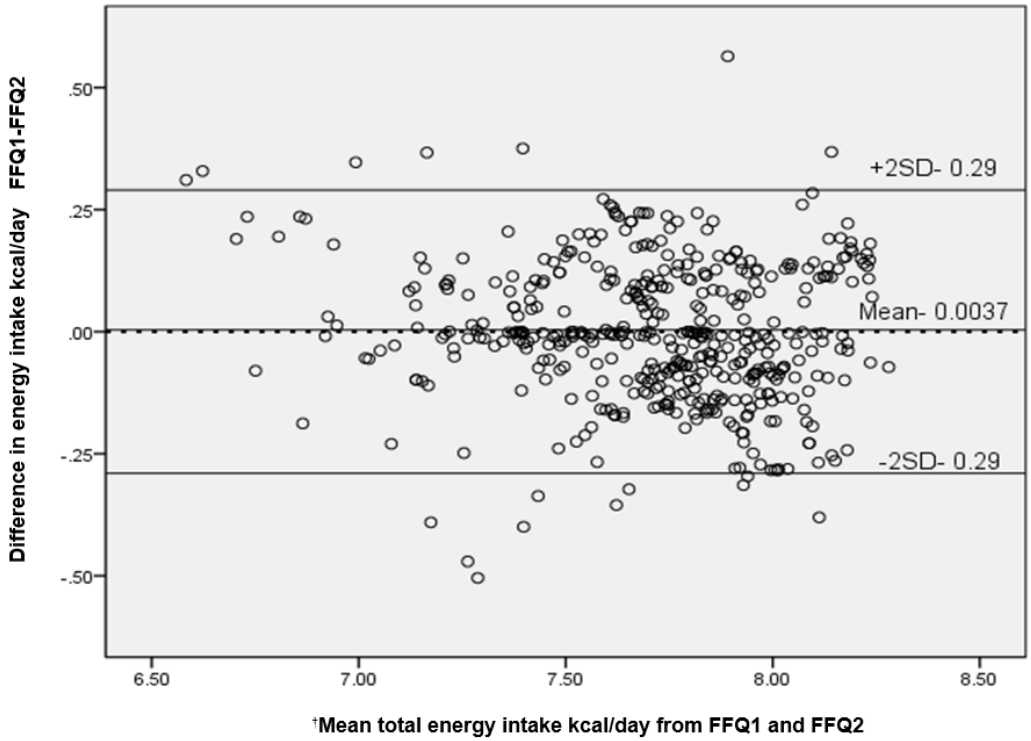

Figure 3. Bland-Altman plot showing the difference between total energy intake from the FFQ 1 and FFQ 2 versus the mean of these two measures. †Mean energy intake from FFQ1 and FFQ2 collected at an interval of 12 months from both urban and rural subjects. Data was log transformed due to heteroscedasticity by visual inspection using residual plot and Kendall's Tau test[31]

Table 1. Region wise demographic, anthropometric and biochemical characteristics of 463 rural and urban participants from 10 Indian states

Description	East		West		North		South		Northeast	
	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Age (yrs)	35.0	25.3	42.0	20.5	36.0	18.0	42.0	24.3	40.0	23.5
Weight (kg)**	50.0	13.3	54.0	18.5	61.5	20.0	54.0	18.0	54.9	13.9
Height (cm)	156	13.3	156	13.0	160	14.3	157	13.0	158	11.4
Body mass index (kg/m^2)**	20.3	5.0	21.9	7.0	23.9	6.0	21.3	6.2	21.4	4.7
Systolic blood pressure (mmHg)	122	21.5	126	19.9	127	19.3	127	20.9	127	24.5
Diastolic blood pressure (mmHg)**	73.5	15.4	78.5	11.4	77.0	13.3	79.0	14.5	81.0	14.0
Serum cholesterol (mg/dL†)**	140	44.8	161	46.0	165	62.0	155	57.5	149	50.5
Serum triglyceride (mg/dL†)*	115	74.0	101	61.5	102	86.5	118	79.8	132	93.5
Serum high density lipoprotein (mg/dL†)	37.0	12.0	40.0	21.0	42.0	14.0	37.0	13.5	37.0	11.0
Serum low density lipoprotein (mg/dL†)**	74.2	34.7	97.8	37.1	93.7	41.8	95.5	41.3	84.8	41.7
Physical activity level (PAL)‡§**, n (%)										
Sedentary (PAL value 1.40-1.69)	36 (38.3)		39 (60.0)		65 (72.2)		71 (62.3)		42 (42.0)	
Moderate (PAL value 1.70-1.99)	35 (37.2)		19 (29.2)		17 (18.9)		30 (26.3)		46 (46.0)	
Vigorous (PAL value 2.00-2.40)	23 (24.5)		7 (10.8)		8 (8.9)		13 (11.4)		12 (12.0)	

PAL: Physical Activity Level

†To convert mg/dL cholesterol to mmol/L, multiply mg/dL by 0.0259. To convert mmol/L cholesterol to mg/dL, multiply mmol/L by 38.7. Cholesterol of 193 mg/dL¼5.00 mmol/L.

‡*p* value for categorical variable test using chi-square

§Based on FAO/WHO cut off for physical activity[26]

*$p<0.05$, **p value <0.001

Table 2. Reproducibility study: Energy adjusted intra cluster correlation coefficients (ICC) of nutrients and food groups from the self-reported food frequency questionnaires collected twice at the interval of 12 months in urban and rural Asian Indian adults (n=463)

	Urban (n= 128)			Rural (n= 335)		
	Median (95% CI)			Median (95% CI)		
	MDRF-FFQ 1 (May 2011)	MDRF-FFQ 2 (June 2012)	ICC (95% CI)†	MDRF-FFQ 1 (May 2011)	MDRF-FFQ 2 (June 2012)	ICC (95% CI)†
Energy (kcal)	2426 (2406-2650)	2478 (2374-2594)	0.77 (0.72-0.84)	2302 (2189-2333)	2321(2224-2379)	0.63 (0.61 - 0.66)
Carbohydrate (g/d)	364 (361-374)	366 (362-377)	0.76 (0.71 - 0.83)	369 (367-375)	372 (371.2-381)	0.69 (0.57 - 0.83)
Protein (g/d)	62.8 (62.7-66.7)	65.4 (65.1-70.7)	0.67 (0.60 - 0.77)	61.9 (62.3-64.4)	63.1 (63.2-66.1)	0.61 (0.59 - 0.64)
Total fat (g/d)	66.8 (62.9-68.2)	72.7 (69.4-75.7)	0.69 (0.62 - 0.77)	65.9 (62.8-66)	69 (67.7-71.5)	0.63 (0.61 - 0.66)
Total SFA (g/d)	22.1 (20.5-23.2)	23.9 (23.3-26.5)	0.5 (0.46 - 0.59)	20.6 (20.2-21.9)	20.5 (20.6-22.7)	0.72 (0.7 - 0.75)
Total MUFA(g/d)	21.8 (20.9-23.8)	22.8 (21.1-23.9)	0.63 (0.54 - 0.76)	23.5 (22.4-24.1)	22.3 (21.6-23.4)	0.63 (0.60 - 0.66)
Total PUFA(g/d)	16 (16.9-19.7)	16.8 (16.7-19.8)	0.7 (0.67 - 0.72)	15 (16.3-17.7)	16 (16.7-18.3)	0.69 (0.62 - 0.79)
Dietary fibre (g/d)	33.6 (32.6-36.2)	29.9 (29.8-33.9)	0.76 (0.71 - 0.83)	34.9 (33.2-35.2)	31.4 (31.4-33.7)	0.71 (0.69 - 0.73)
Glycemic Index	60.7 (60.2-61.3)	61.4 (61.1-62.7)	0.6 (0.57 - 0.66)	61.1 (60.4-61.3)	62 (61.8-62.8)	0.58 (0.56 - 0.61)
Glycemic Load (g/d)	195 (193-204)	194 (186-203)	0.62 (0.58 - 0.67)	199 (200-207)	203 (204-213)	0.63 (0.61 - 0.65)
Refined cereals (g/d)	200 (195-231)	198 (183-223)	0.59 (0.54 - 0.65)	199 (203-227)	203 (199-223)	0.51 (0.48 - 0.54)
Whole grains (g/d)	60.5 (68-95.4)	56.9 (64.7-90.7)	0.53 (0.48 - 0.61)	65.4 (70.1-86)	71.4 (77.7-95.2)	0.61 (0.59 - 0.63)
Pulses and legume (g/d)	44.2 (41.3-50)	44.6 (43.1-52.8)	0.69 (0.66 - 0.72)	40.2 (41.3-46.8)	42.7 (42.2-48.2)	0.52 (0.50 - 0.55)
Milk and milk products (g/d)	244 (246-320)	260 (257-336)	0.61 (0.53 - 0.72)	225 (235-280)	224(237-283)	0.52 (0.48 - 0.57)
Fats and edible oils (g/d)	36.7 (35.8-39.8)	36.2 (35.3-39.4)	0.56 (0.52 - 0.63)	38.9 (36.1-38.7)	36.9 (37.1-39.8)	0.56 (0.54 - 0.59)
Fruits(g/d)	134 (135-173)	133 (137-176)	0.77 (0.72 - 0.84)	129 (132-149)	124 (128-146)	0.89 (0.89 - 0.9)
Leafy vegetables (g/d)	16.1 (21.3-33.3)	18.2 (22.5-35.5)	0.77 (0.72 - 0.84)	13.2 (18.1-24.4)	13.9 (18.5-24.6)	0.69 (0.67 - 0.72)
Other vegetables (g/d)	46.8 (55.8-82.5)	45 (55.5-84.4)	0.74 (0.68 - 0.81)	42.1 (59.3-76.8)	31.3 (49.7-65.4)	0.72 (0.70 - 0.74)
Roots and tuber s(g/d)	224(193-263)	194 (198-256)	0.73 (0.67 - 0.79)	248 (224.0-285)	213 (210-266)	0.71 (0.70 - 0.74)
Animal foods (g/d)	28.1 (33.3-53.4)	27.9 (29.7-60)	0.58 (0.47 - 0.76)	25.7 (35.5-46)	26 (31.5-41.8)	0.5 (0.46 - 0.56)

MDRF-FFQ: Madras Diabetes Research Foundation Food Frequency Questionnaire; ICC: intra cluster correlation coefficients; SFA: Saturated fatty acid; MUFA: Mono-unsaturated fatty acid; PUFA: Poly-unsaturated fatty acid

†ICC measures agreement between FFQ1 and FFQ

Table 3. Reproducibility study: Energy adjusted intra class correlation coefficients (ICC) of nutrients and food groups from the self-reported food frequency questionnaires Asian Indian adults based on gender (n=463)

Nutrient and food group In-take	Male (n=216)			Female (n= 247)		
	Median (95% CI)			Median (95% CI)		
	MDRF-FFQ 1 (May 2011)	MDRF-FFQ 2 (June 2012)	ICC† (95% CI)‡	MDRF-FFQ 1 (May 2011)	MDRF-FFQ 2 (June 2012)	ICC† (95% CI)‡
Energy (kcal)	2410 (2349-2539)	2443 (2357-2543)	0.89 (0.89-0.90)	2267 (2157-2321)	2210 (2179-2354)	0.88 (0.88-0.89)
Carbohydrate (g/d)	371 (366-377)	373 (368-381)	0.72 (0.70-0.74)	365 (365-373)	369 (368-379)	0.74 (0.73-0.76)
Protein (g/d)	62.9 (63.3-66.4)	63.9 (64.7-68.9)	0.75 (0.74-0.77)	61.5 (61.5-63.9)	63.7 (62.8-66)	0.75 (0.73-0.76)
Total fat (g/d)	64.7 (61.2-65.4)	66.8 (66.8-71.7)	0.65 (0.62-0.67)	67.3 (64.1-67.8)	72.3 (69.3-73.6)	0.74 (0.73-0.75)
Total SFA (g/d)	19.8 (19.5-21.6)	20.2 (19.9-22.6)	0.61 (0.58-0.64)	22 (21-22.9)	23 (22.6-24.7)	0.78 (0.77-0.79)
Total MUFA (g/d)	23.1 (21.3-23.5)	22.3 (21.3-23.6)	0. 70 (0.68-0.72)	23.9 (22.5-24.4)	22.6 (21.5-23.5)	0.72 (0.71-0.74)
Total PUFA (g/d)	14.9 (16.1-18)	16.7 (16.8-19)	0.70 (0.69-0.72)	15.7 (16.7-18.5)	16 (16.6-18.4)	0.71 (0.69-0.73)
Dietary fibre (g/d)	33.3 (31.8-34.6)	30.8 (30.4-33.4)	0.64 (0.61-0.67)	35.1 (34-36.3)	31.4 (31.3-34.1)	0.64 (0.62-0.66)
Glycemic Index	61.2 (60.5-61.7)	62.3 (61.9-63.3)	0.66 (0.64-0.69)	60.8 (60.2-61)	61.5 (61.3-62.4)	0.71 (0.70-0.73)
Glycemic Load (g/d)	200(199-209)	201 (199-210)	0.70 (0.68-0.72)	197 (196-204)	197 (199-210)	0.70 (0.69-0.72)
Refined cereals (g/d)	204 (210-240)	215 (203-236)	0.60 (0.57-0.63)	192 (193-219)	197 (187-213)	0.64 (0.62-0.66)
Whole grains (g/d)	55.4 (64.4-85)	59.8 (67.7-90.5)	0.67 (0.65-0.70)	72.8 (73.7-92.1)	72.2 (78.7-97.5)	0.66 (0.64-0.69)
Pulses and legume (g/d)	38 (39.3-46.7)	39.8 (40.5-48.7)	0.71 (0.70-0.73)	42.3 (42.9-48.6)	44.6 (43.9-50.3)	0.68 (0.67-0.71)
Milk and-Milk products (g/d)	210.5 (228-289)	216 (231-292)	0.51 (0.47-0.55)	246 (245-295)	256 (252-305)	0.57 (0.54-0.60)
Fats and edible oils (g/d)	37.4 (34.7-37.9)	36.1 (35.5-39.1)	0.60 (0.57-0.63)	39.3 (37.2-40)	37.3 (37.4-40.3)	0.61 (0.58-0.63)
Fruits(g/d)	131 (133-158)	129 (134-161)	0.55 (0.51-0.59)	130 (132-155)	124 (126-148)	0.50 (0.45-0.57)
Leafy vegetables (g/d)	12 (16-22.1)	13.1 (17.1-24.2)	0.60 (0.58-0.63)	16.3 (21.8-30.9)	15.6 (21.9-30.5)	0.50 (0.47-0.55)
Other vegetables (g/d)	46.4 (61.3-83.4)	36.5 (57.9-81.8)	0.60 (0.57-0.63)	41.8 (55.1-74.7)	32.3 (45.6-61.2)	0.54 (0.51-0.57)
Roots and tubers (g/d)	219 (227-266)	202 (213-242)	0.52 (0.47-0.60)	215 (227-262)	195(204-230)	0.52 (0.47-0.58)
Animal foods (g/d)	34.8 (41.8-57.7)	30.9 (36.1-56)	0.51 (0.47-0.55)	19.6 (29-39.5)	22.2 (27-38.4)	0.62 (0.59-0.65)

MDRF-FFQ: Madras Diabetes Research Foundation Food Frequency Questionnaire; ICC: Intra class correlation; SFA: Saturated fatty acid; MUFA: Mono-unsaturated fatty acid; PUFA: Poly-unsaturated fatty acid

†ICC measures agreement between FFQ1 and FFQ.

‡ICC- Nutrients and food groups were log transformed and further adjusted for covariates age (in years), sex (male/female), BMI (kg/m^2), regions (North, South, East, west and North east), income (INR) (>2000, 2000-5000, 5000-10000, >10000) and education (illiterate, primary, higher secondary and college education).

Table 4. Percent under reporters from the ratio of energy intake and BMR in rural and urban India (n=450)

FFQ 1	n	EI/BMR ratio (mean±SD)	Percentage below the EI/BMR (<1.2)† n (%)
Rural	330	1.75 (0.52)	38 (11.5)
Urban	120	1.83 (0.48)	11 (9.2)
Male	216	1.73 (0.53)	26 (12.0)
Female	234	1.81 (0.49)	23 (9.8)
Overall	450	1.77 (0.51)	49 (10.9)

EI/BMR: energy intake/basal metabolic rate.
†EI/BMR <1.2 are considered as under-reporters and the BMR was calculated from the age- and gender- specific prediction equation[23]

Table 5. Multivariate adjusted regression coefficients for association of carbohydrates (%E), glycemic load with TG and HDL as a measure of construct validity from MDRF FFQ 1 (May 2011))

Description	TG (mg/dL†)			HDL (mg/dL†)		
	β	SE	*p* value	β	SE	*p value
Carbohydrates (% Energy)	2.29	0.72	0.002	-0.48-	0.12	<0.001
Energy adjusted glycemic load	0.38	0.15	0.01	-0.11	0.02	<0.001

TG: triglyceride; HDL: High density lipoprotein; MDRF FFQ: Madras Diabetes Research Foundation Food Frequency Questionnaire
†To convert mg/dL cholesterol to mmol/L, multiply mg/dL by 0.0259. To convert mmol/L cholesterol to mg/dL, multiply mmol/L by 38.7. Cholesterol of 193 mg/dL¼5.00 mmol/L.
*p<0.001 considered to be statistically significant

creased [β (SE): -0.48 (0.12), p<0.001]. Energy adjusted glycemic load (a measure of the carbohydrate quantity and quality) also showed a trend similar to that of carbohydrates (%E). However, for a unit increase in dietary fat, total cholesterol [β (SE): 1.10 (0.51), p=0.032], LDL [β (SE): 0.96 (0.43), p=0.025] and HDL cholesterol [β (SE): 0.57 (0.12), p<0001] significantly increased while triglycerides decreased [β (SE): -1.88 (0.76), p=0.014]. A similar trend was observed for SFA (%E).

DISCUSSION

The present study evaluates, for the first time in India, the reproducibility and construct validity of an interviewer administered comprehensive quantitative national FFQ for adults of both sex residing in rural and urban areas of all regions of India including the Northeast. The MDRF-FFQ is a reliable tool to assess the dietary measures of macronutrients and food groups reported by Asian Indian urban and rural adults. Moderate to good correlation coefficients were found between FFQ1 and the repeat FFQ2 collected at an interval of 12 months for both nutrients and food groups even when stratified by sex, suggesting consistent performance. Construct validity was assessed only with limited biomarkers of blood lipids with the reported intakes of macronutrients such as carbohydrates (%E), glycemic load, total dietary fat (%E) and SFA (%E).

The results of the present study by and large agree with other studies that have reported a moderate to substantial agreement (>0.40 to <0.80) for both nutrients and food. However, SFA, milk and milk products and roots and tubers in this study showed a lower agreement than that cited elsewhere[27,28] especially among urban participants. Several national studies have reported similar ICCs for majority of nutrients (0.40-78) and food groups (0.86-0.99).[7,11,29] Similarly, international studies had shown moderate to high ICC for majority of nutrients (0.42-0.91) and food groups (0.28-0.91).[30,31] The results of the present study are in agreement with the above stated studies (Table 2), thereby reiterating the reliability of the MDRF-FFQ.

Validation of food and nutrient intake by FFQ against biomarkers has enormous value in nutritional epidemiological studies. Biomarkers could reduce subjectivity compared to validation studies with diet records or multiple 24 hour recalls as reference method as the latter is prone to subjectivity like the FFQ. While biomarkers are also prone to errors of estimation and physiological variations, these are unrelated to errors with self-reported dietary assessments[3] and can thereby reflect relationships with nutrient intake as they are based on the biological processes in the body (construct validity). Several studies have shown the effect of intake of carbohydrates and fats on blood lipids.[13,14,32] Hence, in the present study, the energy intake from carbohydrates, fat and SFA reported in FFQ1 were tested against lipid parameters such as triglycerides and HDL, total and LDL cholesterol to evaluate construct validity, which is a first for any FFQ in India.

Nettleton et al. 2009[14] and Ma et al. 2006[33] reported that an increase in carbohydrate calories was associated with a significant increase in triglyceride concentration and decrease in HDL cholesterol concentration. The present study also reported similar findings, after adjusting for potential non-dietary and dietary factors that may affect these lipids. Furthermore, an increase in triglyceride levels and decrease in HDL-c was observed with an increase in GL evaluated from the FFQ (Table 5), similar to reports from the West[34,35] Ours is the first study to assess dietary glycemic load (GL) from the MDRF-FFQ across rural and urban areas of different regions of India. The GL is relevant in diet-disease relationship as it has been well associated with risk of diabetes both in Indian and western populations.[36,37]

Willett et al. 2001[13] reported an increase in HDL-C and a decrease in triglyceride when carbohydrate energy

Table 6. Multivariate adjusted regression coefficients for association of total dietary Fat, SFA with lipid profile as a measure of construct validity from MDRF FFQ 1 (May 2011)

Description	Cholesterol (mg/dL†)			TG (mg/dL†)			HDL (mg/dL†)			LDL (mg/dL†)		
	β	SE	*p* value	β	SE	*p* value	β	SE	*p* value	β	SE	*p* value
Total Fat (% Energy)	1.10	0.51	0.03	-1.88	0.76	0.01	0.57	0.121	<0.001	0.96	0.43	0.03
Total SFA (% Energy)	0.68	0.30	0.02	-1.34	0.59	0.02	0.28	0.094	0.003	0.60	0.25	0.026

TG: triglyceride; HDL: high density lipoprotein cholesterol, LDL: Low density lipoprotein cholesterol; MDRF FFQ: Madras Diabetes Research Foundation Food Frequency Questionnaire

†To convert mg/dL cholesterol to mmol/L, multiply mg/dL by 0.0259. To convert mmol/L cholesterol to mg/dL, multiply mmol/L by 38.7. Cholesterol of 193 mg/dL¼5.00 mmol/L.

‡Adjusted for: age (yrs), sex (M/F), BMI (kg/m^2), region, rural/urban, physical activity category (sedentary 1.40-1.69; moderate: 1.70-1.99; vigorous: 2.0-2.40), systolic and diastolic blood pressure (mmHg), fasting blood glucose (mg/dL), literacy (yes/ no), main cooking oil, total energy (kcal/d) intake

*p<0.001 considered to be statistically.

is replaced by fat in a Western population. Recent findings from 18 countries in the Prospective Urban Rural Epidemiology study (PURE) have shown that a higher percentage energy from fat and SFA was associated with a higher total cholesterol, HDL and LDL-C and lower triglyceride.[38] These agree with the present study findings. Similar findings were also reported in studies done elsewhere.[39-41] This reiterates the MDRF-FFQ's ability to provide a valid measure of dietary fat that could further show the physiologically relevant associations with total cholesterol, TG, LDL and HDL-C (Table 6).

Generally, misreporting of dietary intakes affects the construct validity of the assessment tool and is a barrier in understanding diet-disease relationship.[42] The underestimation of nutrient intakes may be associated with under-reporting of total energy intake. Livingstone and Black 2003 and Black et al. 1991,[43,44] reported widespread prevalence of under-reporting in various nutritional studies. However, the proportion of under-reporters in this study (11%) is much lower than that reported by Bedard et al (43%).[45]

In general, however, studies on FFQ validity are challenging to undertake in a sufficiently large and representative sample of the population for which they have been developed. In addition, there is no gold standard reference method to validate FFQ. One of the important strengths of the MDRF-FFQ is its ability to assess dietary habits of the population in both rural and urban settings of all the regions of India. One may argue that regional diets are diverse and hence separate FFQs are needed for each region. In fact, diets in all regions of India including both rural and urban areas are high in carbohydrates and bulk of the carbohydrate calories is derived from cereal staples (though the choice of grain could differ from region to region) (NSSO 2011-2012). MDRF-FFQ questionnaire is unique as it has been validated using serum triglycerides and HDL-C as construct validity for assessing the dietary GL of the population. However, the FFQ has a few limitations as well. The test-retest for reproducibility assumption that true intake did not change between the 2 administrations of FFQ1 and FFQ2 cannot be confirmed with certainty. The accuracy of responses was dependent on the memory of each individual and would be subject to recall bias.[46] Micronutrients were not evaluated for reproducibility and construct validity owing to budgetary constraints and incomplete micronutrient composition provided in food composition tables. Besides, another limitation was that the serum lipids were measured at only one time, which might not account for intra-individual variability in these parameters. Further, due to budget constraints and feasibility challenges, assessment of recovery biomarkers like doubly-labelled water for energy intake and 24 –hour urine for sodium intake could not be carried out.

Conclusion

MDRF-FFQ can be considered a realistic, practical and economical tool for assessing usual dietary habits of Asian Indian populations, based on the evaluation of the major macronutrients, carbohydrates and fat, that provides 3/4th of daily energy. MDRF-FFQ has validity for the measurement of physiologically important variations in macronutrient intake and presumptively, in the assessment of long-term dietary exposure in studies of chronic disease in India.

ACKNOWLEDGEMENTS

We would like to acknowledge the laboratory team for their support in biochemical assessments as well as the information technology team for their assistance in the dietary data extraction.

AUTHOR DISCLOSURES

All authors declare that there is no conflict of interest.

REFERENCES

1. Anjana RM, Deepa M, Pradeepa R, Mahanta J, Narain K, Das HK et al. Prevalence of diabetes and pre diabetes in 15 states of India: results from the ICMR-INDIAB population-based cross-sectional study. Lancet Diabetes Endocrinol. 2017; 5:585-96. doi: 10.1016/S2213-8587(17)30174-2.
2. Chauhan S, Aeri BT. Prevalence of cardiovascular disease in India and its economic impact-A review. International Journal of Scientific and Research Publications. 2013;3: 2250-3153.
3. Willett W. Nutritional epidemiology. 3rd ed. New York, USA: Oxford University Press; 2013.

4. Dubey KG. The Indian Cuisine. New Delhi,: PHI Learning Pvt. Ltd; 2010.
5. Report of Third Repeat Rural Survey. Hyderabad,India: National Nutrition Monitoring Bureau, National Institute of Nutrition; 2012.
6. Ramachandran. P. The double burden of malnutrition in India. Rome: Food and Agriculture organization; 2006.
7. Sudha V, Radhika G, Sathya RM, Ganesan A, Mohan V. Reproducibility and validity of an interviewer-administered semi-quantitative food frequency questionnaire to assess dietary intake of urban adults in southern India. Int J Food sci nutr. 2006;57:481-93. doi: 10.1080/09637480600969220.
8. Nutritional Intake in India. NSS 68th Round. India: National Sample Survey Organization, Ministry of Statistics and Program Implementation, Government of India. [25th April 2018]; Available from: http://mospi.nic.in/sites/default/files/publication_reports/nss_report_560_19dec14.pdf. 2011-2012.
9. International Institute for Population Sciences (IIPS) and ICF. National Family Health Survey (NFHS-4), 2015-16: India. Mumbai: IIPS; 2017.
10. Bowen L, Bharathi AV, Kinra S, DeStavola B, Ness A, Shah E. Development and evaluation of a semi quantitative food frequency questionnaire for use in urban and rural India. Asia Pac J Clin Nutr. 2012;21:355-60.
11. Sowmya N, Lakshmipriya N, Arumugam K, Venkatachalam S, Vijayalakshmi P, Ruchi V et al. Comparison of dietary profile of a rural south Indian population with the current dietary recommendations for prevention of non-communicable diseases. Indian J Med Res. 2016;144:112-9. doi: 10.4103/0971-5916.193297.
12. Bharathi AV, Kurpad AV, Thomas T, Yusuf S, Saraswathi G, Vaz M. Development of food frequency questionnaires and a nutrient database for the Prospective Urban and Rural Epidemiological (PURE) pilot study in South India: methodological issues. Asia Pac J Clin Nutr. 2008;17:178-85.
13. Willett W, Stampfer M, Chu NF, Spiegelman D, Holmes M, Rimm E. Assessment of questionnaire validity for measuring total fat intake using plasma lipid levels as criteria. Am J Epidemiol. 2001;154:1107-12. doi: 10.1093/aje/154.12.1107.
14. Nettleton JA, Rock CL, Wang Y, Jenny NS, Jacobs DR. Associations between dietary macronutrient intake and plasma lipids demonstrate criterion performance of the Multi-Ethnic Study of Atherosclerosis (MESA) food-frequency questionnaire. Br J Nutr. 2009;102:1220-7. doi: 10.1017/S0007114509382161.
15. Radhika G, Sathya RM, Ganesan A, Saroja R, Vijayalakshmi P, Sudha V, Mohan V. Dietary profile of urban adult population in South India in the context of chronic disease epidemiology (CURES–68). Public Health Nutr. 2011;14:591-8. doi: 10.1017/S136898001000203X.
16. International Institute for Population Sciences (IIPS) and ICF. National Family Health Survey (NFHS-4), 2005-2006: India. Mumbai: IIPS; 2007.
17. Sudha V, Mohan V, Anjana RM, Kamala Krishnaswamy. Dr. Mohan's Atlas of Indian. Chennai,India: ; 2015.
18. Anjana RM, Pradeepa R, Deepa M, Datta M, Sudha V, Unnikrishnan R et al. The Indian Council of Medical Research—India Diabetes (ICMR-INDIAB) Study: Methodological Details. J Diabetes Sci Technol. 2011;5:906-14. doi: 10.1177/1932296 81100500413.
19. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem. 1972;18:499-502.
20. Anjana RM, Rani CS, Deepa M, Pradeepa R, Sudha V, Nair HD et al. Incidence of diabetes and pre diabetes and predictors of progression among Asian Indians: 10-year follow-up of the Chennai Urban Rural Epidemiology Study (CURES). Diabetes Care. 2015;38:1441-8. doi: 10.2337/dc14-2814.
21. Anjana RM, Sudha V, Lakshmipriya N, Subhashini S, Pradeepa R, Geetha L et al. Reliability and validity of a new physical activity questionnaire for India. Int J Behav Nutr Phys Acty. 2015; 12:40. doi: 10.1186/s12966-015-0196-2.
22. Human energy requirements. Scientific background papers from the Joint FAO/WHO/UNU Expert Consultation. Public Health Nutr. 8: 929–1228. 2005. [15th April 2018]; Available from: http://www.fao.org/3/a-y5686e.pdf
23. Indian Council of Medical Research. Nutrient Requirements and Recommended Dietary Allowances for Indians: A Report of the Expert Group of the Indian Council of Medical Research. NewDelhi, India: Indian Council of Medical Research; 2010.
24. Goldberg GR, Black AE, Jebb SA, Cole TJ, Murgatroyd PR, Coward WA, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. Eur J Clin Nutr. 1991;45:569-81.
25. EpiNu [food and nutrient database]. Version 1. India: Madras Diabetes Research Foundation; 2006
26. BREHM MA, Scholtes VA, Dallmeijer AJ, Twisk JW, Harlaar J. The importance of addressing heteroscedasticity in the reliability analysis of ratio-scaled variables: an example based on walking energy-cost measurements. Dev Med Child Neurol. 2012;54:267-73.
27. Ibiebele TI, Parekh S, Mallitt KA, Hughes MC, O'Rourke PK, Webb PM. Reproducibility of food and nutrient intake estimates using a semi-quantitative FFQ in Australian adults. Public Health Nutr. 2009;12:2359-65. doi: 10.1017/S1368980009005023.
28. Hebden L, Kostan E, O'Leary F, Hodge A, Allman-Farinelli M. Validity and reproducibility of a food frequency questionnaire as a measure of recent dietary intake in young adults. PLoS One. 2013;8:e75156. doi: 10.1371/journal.pone.0075156.
29. Shaikh NI, Frediani JK, Ramakrishnan U, Patil SS, Yount KM, Martorell R, Narayan KV, Cunningham SA. Development and evaluation of a Nutrition Transition-FFQ for adolescents in South India. Public Health Nutr. 2017; 20:1162-72. doi: 10.1017/S1368980016003335.
30. Esfahani FH, Asghari G, Mirmiran P, Azizi F. Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the Tehran Lipid and Glucose Study. J Epidemiol. 2010;20:150-8. doi: 10.2188/jea.je20090083.
31. Dehghan M, del Cerro S, Zhang X, Cuneo JM, Linetzky B, Diaz R, Merchant AT. Validation of a semi-quantitative Food Frequency Questionnaire for Argentinean adults. PLoS One. 2012;25:7:e37958. doi: 10.1371/journal.pone.0037958.
32. Mohseni-Takalloo S, Mirmiran P, Hosseini-Esfahani F, Azizi F. Dietary fat intake and its relationship with serum lipid profiles in tehranian adolescents. J Food Nutr Res. 2014;2:330-4.
33. Ma Y, Li Y, Chiriboga DE, Olendzki BC, Hebert JR, Li W, Leung K, Hafner AR, Ockene IS. Association between carbohydrate intake and serum lipids. J Am Coll Nutr. 2006;25:155-63.
34. Liu S, Manson JE, Stampfer MJ, Holmes MD, Hu FB, Hankinson SE, Willett WC. Dietary glycemic load assessed by food-frequency questionnaire in relation to plasma high-density-lipoprotein cholesterol and fasting plasma

triacylglycerols in postmenopausal women. Am J Clin Nutr. 2001;73:560-6. doi: 10.1093/ajcn/73.3.560.
35. Brand-Miller JC. Glycemic load and chronic disease. Nutr Rev. 2003;61:S49-55. doi: 10.1301/nr.2003.may.S49-S55.
36. Mohan V, Radhika G, Sathya RM, Tamil SR, Ganesan A, Sudha V. Dietary carbohydrates, glycaemic load, food groups and newly detected type 2 diabetes among urban Asian Indian population in Chennai, India (Chennai Urban Rural Epidemiology Study 59). Br J Nutr. 2009;102:1498-506. doi: 10.1017/S0007114509990468.
37. Sun Q, Spiegelman D, van Dam RM, Holmes MD, Malik VS, Willett WC, Hu FB. White rice, brown rice, and risk of type 2 diabetes in US men and women. Arch Intern Med. 2010;170:961-9. doi: 10.1001/archinternmed.2010.109.
38. Mente A, Dehghan M, Rangarajan S, McQueen M, Dagenais G, Wielgosz A et al. Association of dietary nutrients with blood lipids and blood pressure in 18 countries: a cross-sectional analysis from the PURE study. Lancet Diabetes Endocrinol. 2017;5:774-87. doi: 10.1016/S2213-8587(17)30283-8.
39. Sharlin J, Posner BM, Gershoff SN, Zeitlin MF, Berger PD. Nutrition and behavioral characteristics and determinants of plasma cholesterol levels in men and women. J Am Diet Assoc. 1992;92:434-40.
40. Sackner-Bernstein J, Kanter D, Kaul S. Dietary intervention for overweight and obese adults: comparison of low-carbohydrate and low-fat diets. A meta-analysis. PLoS One. 2015;10:e0139817. doi: 10.1371/journal.pone.0139817.
41. Schwingshackl L, Hoffmann G. Comparison of effects of long-term low-fat vs high-fat diets on blood lipid levels in overweight or obese patients: a systematic review and meta-analysis. J Acad Nutr Diet. 2013;113:1640-61. doi: 10.1016/j.jand.2013.07.010.
42. Mendez MA. Invited commentary: dietary misreporting as a potential source of bias in diet-disease associations: future directions in nutritional epidemiology research. Am J Epidemiol. 2015;181:234-6. doi: 10.1093/aje/kwu306
43. Livingstone MB, Black AE. Markers of the validity of reported energy intake. J Nutr. 2003;133:895S-920S. doi: 10.1093/jn/133.3.895S.
44. Black AE, Goldberg GR, Jebb SA, Livingstone MB, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2.Evaluating the results of published surveys. Eur J Clin Nutr. 1991;45:583-99.
45. Bedard D, Shatenstein B, Nadon S. Underreporting of energy intake from a self-administered food-frequency questionnaire completed by adults in Montreal. Public Health Nutr. 2004;7:675-81. doi: 10.1079/PHN2003578.
46. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test–retest reliability. PLoS One. 2013;8:e73990. doi: 10.1371/journal.pone.0073990.

Supplementary table 1. Region-wise intake of nutrients and food groups

Nutrient and food group intake	East		West		North		South		Northeast	
	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Energy (kcal)	2412	1014	2158	676	2681	592	2263	750	2062	903
Carbohydrate (g/d)	373	29	351	29.6	355	27.5	375	37.8	380	53.1
Protein (g/d)	58.4	8.2	58.0	10.2	61.2	9.6	63.9	11.6	66.6	12.7
Fat (g/d)	65.1	12.2	77.2	13.4	72.7	11.5	63.8	14.7	56.3	19.3
SFA (g/d)	19.1	7.0	28.4	5.5	24.7	7.9	20.8	8.7	15.7	8.4
MUFA (g/d)	28.9	6.6	21.5	8.1	30.1	6.2	21.0	8.0	16.3	8.4
PUFA (g/d)	13.1	2.5	22.5	9.8	13.5	3.5	16.2	10.5	18.2	7.4
Dietary fibre (g/d)	37.9	8.2	40.3	8.4	41.0	10.7	29.7	9.3	25.1	10.4
Glycemic Index	62.0	3.5	57.8	5.1	58.9	3.4	61.4	2.8	63.1	4.2
Glycemic load (g/d)	203.0	48.7	179	37.6	1736	41.5	212	37.3	227	56.0
Refined cereals (g/d)	197.1	81.2	118	86.8	118	48.4	229	79.6	311	162
Whole grains (g/d)	120.4	80.0	89.1	108	157	61.2	31.5	39.8	22.3	54.1
Pulses and legumes (g/d)	43.7	25.8	44.3	32.3	42.6	30.4	46.1	20.6	30.5	21.5
Milk and milk products (g/d)	157	189	302	250	345	229	246	199	138.3	166.6
Fats and edible oils (g/d)	39.8	9.4	46.2	9.1	41.3	10.7	34.5	13.7	32.5	14.4
Fruits (g/d)	139	70.7	152	71.4	151	100	117	65.2	108	59.7
Leafy vegetables (g/d)	6.9	10.8	17.3	16.7	10.4	31.4	23.1	19.5	11.7	15.3
Other vegetables (g/d)	29.9	48.9	23.7	27.6	38.5	53.7	47.1	32.1	104	164
Roots and tuber s(g/d)	98.5	53.2	110	51.5	96.0	48.8	85.6	33.6	110	59.2
Animal foods[†] (g/d)	17.0	31.9	8.1	22.0	5.1	24.0	40.2	54.0	60.5	67.8

IQR: Inter Quartile Range; SFA: saturated fatty acids; MUFA: monounsaturated fatty acids; PUFA: Polyunsaturated fatty acids

[†]Animal foods include meat, poultry, egg, fish and other sea foods.