The American Journal of Clinical Nutrition 120 (2024) 1105–1113

American Society for Nutrition
Excellence in Nutrition Research and Practice

The American Journal of
CLINICAL NUTRITION

journal homepage: https://ajcn.nutrition.org/

Original Research Article

Computer vision–assisted dietary assessment through mobile phones in female youth in urban Ghana: validity against weighed records and comparison with 24-h recalls

Aulo Gelli [1], Odiche Nwabuikwu [1], Boateng Bannerman [2], Gabriel Ador [2], Vicentia Atadze [2], Millicent Asante [2], Silas Bempong [2], Peter McCloskey [3], Phuong Hong Nguyen [1,4], David Hughes [3], Gloria Folson [2,*]

[1] *International Food Policy Research Institute, Washington, DC, United States;* [2] *Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, Ghana;* [3] *PlantVillage, Penn State University, University Park, PA, United States;* [4] *Thai Nguyen University of Pharmacy and Medicine, Thai Nguyen, Vietnam*

A B S T R A C T

Background: Gaps persist in the data on diets and on the validity of dietary assessment methods in youth in low-income and middle-income countries (LMICs) due to costs constraints. Although computer vision–assisted dietary assessment tools have been proposed, limited evidence exists on their validity in LMICs.

Objectives: This study aimed to validate FRANI (Food Recognition Assistance and Nudging Insights), a mobile phone application with computer vision–assisted dietary assessment, against weighed records (WRs) and compare with 24-h recalls (24HR), in female youth in Ghana.

Methods: Dietary intake was assessed on 2 nonconsecutive days using FRANI, WR, and 24HR in females aged 18–24 y recruited at the University of Ghana, Accra ($n = 64$). Equivalence was examined by comparing intake mean ratios (FRANI/WR and 24HR/WR) with error margins of 10%, 15%, and 20%, using mixed-effect regression models adjusting for repeated measures. Agreement between methods was assessed using the concordance correlation coefficient (CCC).

Results: Equivalence for FRANI and WR was found at 10% bound for riboflavin and vitamin B-6 intakes and at 15% bound for protein, fat, calcium, folate, iron, thiamine, vitamin C, and zinc intakes. Energy, fiber, vitamin A, and niacin intakes were equivalent at 20% bound. Comparisons between 24HR and WR found no estimates within a 10% bound. Protein, iron, niacin, riboflavin, and zinc intakes were equivalent at a 15% bound; folate, thiamine, and vitamin B-12 intakes were equivalent at a 20% bound. CCCs between FRANI and WR ranged from 0.45 to 0.74 (mean: 0.60) and between 24HR and WR ranged from 0.48 to 0.76 (mean: 0.63). Omission errors were 15% for FRANI and 22% for 24HR. Intrusion errors were 22% for FRANI and 18% for 24HR.

Conclusions: FRANI-assisted dietary assessment accurately estimates nutrient intake and performed as accurately as 24HR in female youth in Ghana. Although improvements in computer vision–assisted diet assessment are possible, emerging evidence on FRANI suggests its readiness for scale-up.

Keywords: youth, adolescence, urbanization, diet, dietary assessment, low-income countries, validation

Background

Estimates suggest that diet-related risk factors cause 11 million deaths every year, higher than any other factor included in the Global Burden of Disease analyses [1]. These risks are further exacerbated by trends associated with the "nutrition transition," involving increased consumption of unhealthy, processed foods and decreased physical activity, leading to increases in rates of overweight and obesity [2]. Up to date dietary intake data are essential for effective evidence-based nutrition actions [3]. However, dietary data collection and analysis is complex and expensive [4]. Dietary assessment surveys commonly use the multipass 24-h recall (24HR) method that has been validated for use in adults reporting their intake and/or that of their young children [5], as well as in adolescents [6,7]. The costs of conducting a 24HR are of the order of $500 per respondent [8].

Abbreviations: CCC, concordance correlation coefficient; FRANI, Food Recognition Assistance and Nudging Insights; LMIC, low-income and middle-income country; LOA, limits of agreement; WR, weighed record; 24HR, 24-h recall.

* Corresponding author.

E-mail address: gfolson@noguchi.ug.edu.gh (G. Folson).

https://doi.org/10.1016/j.ajcnut.2024.08.011

Received 23 March 2024; Received in revised form 11 July 2024; Accepted 12 August 2024; Available online 8 October 2024

Although mobile phone–based dietary assessment has the potential to lower assessment costs, only a few tools have been validated and assessed for feasibility of use in low-income and middle-income countries (LMICs) [4]. In response to these challenges, the Nudging for Good project developed, validated, and assessed the feasibility of adopting innovative artificial intelligence (AI) computer vision, mobile technology to provide real-time diagnostics on dietary intake in adolescents living in urban settings in Ghana and Vietnam. The PlantVillage Food Recognition Assistance and Nudging Insights (or FRANI) application has demonstrated the capability to estimate food and nutrient intakes at least as accurately as a trained dietician undertaking a 24HR in adolescents 12–19 y in Ghana and Vietnam [7,9].

This study builds upon previous research by specifically focusing on young females aged 18–24 y, a demographic where understanding dietary intake is crucial for health outcomes and nutritional interventions. The primary aim is validating the use of FRANI to measure food intake in female youth aged 18–24y in urban Ghana against weighed records (WRs), considered the gold standard for dietary assessment, and comparing FRANI performance with that of 24HR. The study objectives were the following: *1*) to estimate food and nutrient intake using 3 methods (WR, FRANI, 24HR); *2*) to assess the extent of agreement between FRANI and 24HR with WR; and *3*) to examine sources of error for FRANI and 24HR.

Methods

Study population

Sixty-four female participants aged 18–24 y were recruited from the University of Ghana Legon campus student population. Eligibility criteria included living on or close to campus, being capable of using smartphones with FRANI application (provided by the project), and willingness to use it to record their food consumption including WR. Participants not willing to be followed up during the WR were excluded from the study. Recruitment efforts involved mounting flyers seeking interested students on notice boards and porters' lodges in all halls of residence on the university campus. Additionally, recruitment information was circulated on social media during holidays. A list of eligible participants was compiled from those expressing interest in the study. Participants were then randomly selected for participation, with selection stratified across the 4 university colleges. The sample size allowed detection of a 10% difference in energy intake for the dietary assessment methods and detect equivalence within 10% bounds ($\alpha = 0.05$; $\beta = 20\%$), as shown in a validation study in a similar study population [6].

Data collection

Each participant visited 5 times, including 1 d for recruitment, consent, and training, followed by 2 nonconsecutive days for WR (1 weekday and 1 weekend day) and 2 nonconsecutive days for the 24HR. During the recruitment visit, enumerators obtained consent, provided the smartphone preconfigured with FRANI and trained participants on FRANI application use. A FRANI video tutorial was also preloaded on the smartphones for reference. Dietary assessment was conducted ensuring that the reference day was consistent for each method. For example, following the recruitment visit, enumerators returned to the participant's home early in the morning to conduct the WR, recording all instances of food and beverage consumption, concurrently as participants recorded foods and beverages consumed using FRANI. On the following day, a different enumerator completed a multipass 24HR with the same participant. Data collection for WR and FRANI application thus took place simultaneously on each reference day, whereas the 24HR was completed the subsequent day, using the previous day as a reference. Background information on participants socioeconomic status was also collected using a structured questionnaire before the first 24HR.

Twenty-six enumerators and 2 supervisors, all with backgrounds in nutrition or health and experience in dietary assessment, were trained for 6 d. The training included presentations, role playing in mock interviews and practice in the field. Supervisors conducted daily checks to ensure accuracy and completeness of the survey data throughout the data collection period. Informed consent was requested from participants using a standardized form. Ethical clearance was obtained from IFPRI institutional review board (IRB) and Noguchi IRB (Federal-wide Assurance, number 00001824; NMIMR IRB CPN IRB00001276). Data collection took place between 1 November and 30 November, 2022. Two participants dropped out of the study after the first day of WR due to fatigue and time burden involved in the data collection (see participant flow diagram, Supplemental Figure 1).

Weighed records

Enumerators tracked participants from early morning (06:00) until after their last evening meal (20:00), recording weights of foods and beverages as consumed in chronological order [7]. Upon arrival, enumerators verified whether any meals had been consumed already. When a meal had been consumed before the enumerator's arrival, participants were prompted to recall the details of the meal and an alternative portion estimation method was used as per the 24HR. For mixed dishes, enumerators recorded recipes, including weights of individual ingredients, the final cooked weight of the mixed dish, and the amount consumed by the participant. Recipe details were provided by the household member responsible for meal preparation. When a food involved residual waste (e.g., bones), enumerators also recorded the weight of the waste. To conclude the WR, enumerators asked participants to confirm that they had eaten their last meal of the day before departing.

24-h recalls

The 24HR was collected using the multipass, quantitative 24HR recall method [5]. Briefly, this method entails interviewing participants, prompting them to elicit all foods and liquids consumed, as well as separately listing individual ingredients consumed in mixed dishes. Foods and beverages were listed regardless of the specific quantities that had been consumed, including condiments, spices, and sweeteners. Visual aids were used during interviews to assist in estimating portion sizes. For each food, a preferred quantification method was determined a priori, including direct weighing of similar foods. Dry powdery foods were estimated using "gari," a pan-roasted grated cassava flour. Household measures, or calibrated models, were also used to quantify portions of grains, roots, and tubers. Soups, sauces, and porridges were estimated using equivalent volumes of water. Throughout the data collection, all data were checked daily by the supervisor, research assistants, and principal investigator, for consistency and plausibility. Any queries were clarified with respondents within a 48-h period.

Food densities and standardized household measure weights were used to convert quantity data collected from volume or household measures into grams. The weights of items estimated through household measures, food models, and water volume were estimated using appropriate conversion factors collected before the dietary assessment. Finally, all data entries were crosschecked for completeness and accuracy by 3 research team members.

Food Recognition Assistance and Nudging Insights

Standard android mobile phones (Samsung Galaxy A21s) were provided to each participant, equipped with a preconfigured FRANI application, 1 d before the WR. Participants were trained on capturing images of the foods and beverages consumed at every meal using FRANI. The process involved users taking a picture of the meal they were about to consume, confirming the food classification returned by FRANI, and inputting the actual amount of food consumed as a proportion of the total portion served. In cases when FRANI's food classification was inaccurate or incomplete, users had the option to record the specific food item consumed by selecting from a comprehensive list of foods compiled from previous studies [9]. To allow for portion size estimation, a standardized visual prop, known as a "pop-socket," was used. This small disc, with a standard size (diameter: 3.96 cm), served as a fiduciary marker necessary for automatic portion size estimation. The FRANI food recognition incorporates an algorithm designed to scale each pixel in the respective images using the pop-socket as a reference. This adjustment enables the estimation of the 2D area covered by each food in the image and subsequently estimate the weight in grams based using a weight-per-pixel parameter, equivalent to a 2D density, derived from previous images used in training the food recognition algorithm [7,9].

Estimating food and nutrient intakes

For each assessment method, the gram quantities of different food items consumed were converted into nutrients using the West African food-composition table and the RING nutrient composition table, which is a compilation of food-composition databases relevant to Ghana [10]. The conversion into nutrients was undertaken using a program written in Stata. The dietary data were also used to examine food group consumption and dietary diversity. Each single food was classified into a specific food group according to the guidelines for the construction of the Minimum Dietary Diversity for Women indicator, which include the following groups: *1*) grains, white roots and tubers, and plantains; *2*) pulses; *3*) nuts and seeds; *4*) dairy; *5*) flesh foods; *6*) eggs; *7*) dark-green leafy vegetables; *8*) vitamin A–rich fruits and vegetables; *9*) other vegetables; and *10*) other fruits [11]. Oils (except for red palm oil and vitamin A–fortified oil, which were classified in the vitamin A–rich fruits and vegetables group), drinks, spices, condiments, and sweeteners were not classified into food groups. Foods with a total daily consumption of <10 g were considered condiments.

Statistical analysis

For each method, descriptive statistics were used to characterize energy and nutrient intakes by person-day. To ensure comparability across the 3 methods, consumption periods were matched across the 3 methods, excluding foods that were reportedly consumed outside the WR period. As the distributions of nutrient intakes were in most cases skewed, mean (SD) and median (interquartile range) intakes were reported. Bland–Altman plots were used to visualize differences in intakes for the different methods compared with the average intake by method. The limits of agreement (LOAs), calculated as the mean difference ± 1.96 SDs, represented the range where 95% of individual differences would be expected.

Equivalence testing was undertaken using the two 1-sided paired *t* test method, examining mean differences of log-transformed nutrient intakes with 10% (i.e. with 90% CI falling within a ratio of 0.9 to 1.1), 15% (i.e., with 90% CI falling within a ratio of 0.85 to 1.15), and 20% bounds (i.e., with 90% CI falling within a ratio of 0.8 to 1.2) using mixed-effect models including a random effect at the person level to account for the repeated measures [6,12,13]. The differences in log-transformed intakes are equivalent to ratios where the numerator is the intake as estimated by FRANI or 24HR and the denominator is the intake from the WR. In addition, we examined the extent of agreement between FRANI and WR and between 24HR and WR using the concordance correlation coefficient (CCC) with bootstrapped standard

TABLE 1
Characteristics of participants[1].

Characteristics	Mean (SD) or %
Adolescents	
Age (y)	21.6 (1.1)
Weight (kg)	64.1 (19)
Height (cm)	158.7 (15.8)
Proportion of participants living on campus	90.3
Ethnicity	
Ga/Ga-Adangme	12.9
Ewe	22.6
Akan	54.8
Northern	1.6
Other	8.1
University level	
Level 200 (year 2, bachelor level)	6.5
Level 300 (year 3, bachelor level)	40.3
Level 400 (year4, bachelor level)	51.6
Level 500 (year 1, masters level)	1.6
Previous semester GPA (range, 1–4)	3.2 (0.4)
Proportion of participants that own a smartphone	100.0
Household	
No. of people in household	3.8 (3.0)
Proportion of household owning assets	
Radio	24.2
Television	17.7
Air conditioner	9.7
Microwave	25.8
Fridge	64.5
Blender	51.6
Computer	88.7
Rice cooker	53.2
Electric stove	59.7
Vehicle	14.5
Parents	
Mother's level of education	
None	6.5
Primary	9.7
Junior high school	21.0
Senior high school	37.1
Tertiary	25.8
Mother's occupation	
Trader	58.1
Vocational	4.8
Unemployed	22.6
Private salaried worker	3.2
Other	11.2
Father's level of education	
None	3.2
Primary	1.6
Junior high school	14.5
Senior high school	37.1
Tertiary	43.6
Father's occupation	
Trader	1.6
Vocational	6.5
Unemployed	32.3
Private salaried worker	12.9
Casual wage earner	1.6
Other	45.2

Abbreviation: GPA, grade point average.
[1] $n = 62$.

errors [14]. Sources of error were examined by individual foods including *1*) instances of omissions (foods consumed but not reported) and intrusions (foods reported that were not consumed) for FRANI and 24HRs and *2*) portion estimation errors, assessed by comparing the reported food amounts by 24HR and FRANI to the observed amounts by WR for the top 20 most commonly consumed foods. Data were analyzed using STATA version 16.0 (StataCorp) and R version 4.2.2 (R Core Team).

Calibration of FRANI for optimal portion estimation

To explore potential gains in precision in FRANI from calibration, we repeated the FRANI model estimations, equivalence testing, and CCC comparisons with WR, using parameter values obtained from different percentiles of the weight-per-pixel coefficient of all the different foods in the FRANI image training library data set, including average (the default starting value), median, and 60th, 75th, 80th, and 100th percentiles. For each different weight-per-pixel parameter value, summary performance measures were obtained for model inferences combining results from equivalence tests (using the mean deviation across all nutrients considered and the count of nutrients falling within the given bounds), and the mean CCC across all nutrients considered. The weight-per-pixel parameter that minimized mean deviation while maximizing counts falling within the given bounds and the mean CCC was selected for the validation study.

Results

Descriptive statistics

Complete records were obtained for 124 person-days. Female participants were 21.6 y old on average (Table 1). The majority of participants lived on campus (90%) and were in the third (level 300) or fourth (level 400) year of a bachelor degree program (92%). All participants owned a smartphone. On average, participants lived in households of 3.8 members, with most households having access to various amenities, including a computer (88.7%), fridge (64.5%), electric stove (60%), iron (60%), blender (51.6%), and rice cooker (53.2%).

Analysis of the WR showed that, on average, there were 8.1 instances of food or beverage consumption per participant day with an average portion size of 140 g per instance. The mean reference period for the WR was 13 h per participant day. Mean daily energy intake was 2343.2 kcal (Table 2). Compared with WR, average daily energy intakes were lower for both FRANI and 24HR. The highest contributors to the percentage of daily energy intake were grains, roots, and tubers (24%) and other vitamin A–rich fruits and vegetables (24%) (Table 3). The lowest contributors to daily energy intake were other fruits (3%), pulses (3%), dark-green leafy vegetables (2%), and nuts and seeds (2%). On average, participants consumed 6.2 food groups daily and over 90% consumed 5 or more food groups per day. The proportion of participants consuming the different food groups was similar across the 3 methods (Figure 1). All 3 methods recorded that all participants consumed food grains daily. A high likelihood of food group consumption (>80%) was also found for vitamin A–rich fruits and vegetables, other vegetables, and meat, poultry, and fish.

Equivalence testing

The detailed results of the calibration of FRANI are summarized in Supplemental Tables 1–3 and Supplemental Figure 2. Calibration results suggested the choice of the maximum value available in the distribution from the FRANI image training library data set.

The distribution of energy intake estimated by FRANI and compared with WR as displayed in Bland–Altman plot (Supplemental Figure 3) highlighted relatively narrow (<1) LOAs, with the proportion falling outside the LOA was <10%. When comparing ratios of log-transformed intakes from FRANI with those of WR (Figure 2A, Supplemental Table 4), riboflavin and vitamin B-6 intakes were equivalent at the 10% error bound. Protein, fat, calcium, folate, iron, thiamine,

TABLE 2
Quantity of food group consumed from major food groups on 2 d by observed weighed records, FRANI application, and 24-h recall.

	Weighed records (n = 124)[1]		FRANI (n = 124)		24-h recall (n = 124)	
	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median
Quantity consumed (g)						
Grains, roots, and tubers	627.7 (317.4)	615.5	406.9 (343.4)	314.3	540.5 (314.3)	504.5
Pulses (beans, peas, and lentils)	134.6 (186.7)	0.0	91.5 (126.0)	20.2	125.1 (183.2)	0.0
Nuts and seeds	20.1 (81.5)	0.0	17.1 (61.7)	0.0	24.5 (92.4)	0.0
Dairy	85.0 (136.0)	14.0	66.1 (110.5)	0.0	45.2 (89.3)	21.5
Meat, poultry, and fish	363.7 (242.3)	321.5	326.5 (305.6)	267.5	311.2 (223.2)	264.9
Eggs	155.1 (204.1)	50.0	113.6 (178.5)	38.4	103.4 (153.9)	43.0
Dark-green leafy vegetables	28.2 (87.5)	0.0	13.0 (40.1)	0.0	27.2 (62.3)	0.0
Other vitamin A–rich fruits and vegetables	445.7 (259.7)	409.0	438.8 (342.8)	371.3	342.5 (237.8)	294.3
Other vegetables	446.0 (300.2)	429.5	372.7 (331.6)	296.4	367.3 (277.5)	275.8
Other fruits	63.6 (145.2)	0.0	44.4 (114.4)	0.0	97.3 (198.4)	0.0
Percentage of energy intake						
Grains, roots, and tubers	23.5 (14.3)	20.7	22.3 (13.3)	20.5	22.6 (13.8)	20.2
Pulses (beans, peas, and lentils)	3.3 (5.9)	0.0	3.0 (5.5)	0.0	3.1 (5.7)	0.0
Nuts and seeds	1.9 (4.7)	0.0	2.1 (5.3)	0.0	1.8 (5.1)	0.0
Dairy	3.7 (7.9)	0.0	3.8 (7.8)	0.0	9.0 (11.0)	0.0
Meat, poultry, and fish	14.2 (13.0)	12.9	14.1 (12.6)	14.2	15.2 (14.6)	14.5
Eggs	6.5 (9.2)	0.9	5.5 (7.9)	0.7	5.4 (7.3)	1.0
Dark-green leafy vegetables	2.4 (6.2)	0.0	1.5 (5.2)	0.0	2.3 (6.2)	0.0
Other vitamin A–rich fruits and vegetables	24.4 (13.7)	22.0	25.1 (13.6)	24.5	21.6 (13.5)	20.3
Other vegetables	12.2 (8.6)	10.6	15.8 (11.1)	13.5	11.8 (9.6)	8.7
Other fruits	2.6 (7.0)	0.0	2.5 (7.0)	0.0	2.5 (6.5)	0.0

Abbreviation: FRANI, Food Recognition Assistance and Nudging Insights.
[2]No. of person-days = 124, equal to number of subjects (62) multiplied by number of recalls (2).

TABLE 3
Nutrient intakes from observed weighed records, FRANI application, and 24-h recall.

	Weighed records (n = 124)[1]		FRANI (n = 124)		24-h recall (n = 124)	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Energy (kcal)	2343 (1193)	2198 (1550)	2197 (1307)	1938 (1525)	2152 (1447)	1807 (1702)
Protein (g)	56.2 (26.9)	54.5 (32.0)	54.4 (30.8)	47.7 (38.2)	56.4 (29.6)	53.8 (41.8)
Fat (g)	114.5 (92.6)	88.6 (126.2)	118.4 (105.6)	79.7 (131.3)	104.4 (100.2)	70.2 (97.2)
Carbohydrates (g)	197.7 (88.0)	185.3 (100.9)	168.2 (96.8)	150.8 (126.0)	170.9 (98.9)	142.0 (114.5)
Fiber (g)	24.7 (16.6)	21.3 (18.6)	22.3 (13.8)	19.0 (21.4)	22.6 (18.2)	17.3 (20.1)
Calcium (mg)	492.1 (270.4)	455.1 (320.3)	545.6 (408.3)	463.4 (430.8)	446.2 (292.5)	381.6 (412.9)
Folate (μg DFE)	297.8 (288.1)	179.8 (282.0)	277.6 (226.5)	194.1 (241.7)	288.5 (280.6)	161.3 (349.9)
Iron (mg)	17.1 (13.7)	13.2 (11.1)	17.4 (13.4)	15.1 (14.4)	19.1 (21.9)	13.4 (13.9)
Niacin (mg)	18.4 (9.2)	17.1 (11.2)	17.6 (12.7)	14.7 (12.5)	18.7 (11.1)	17.3 (15.6)
Riboflavin (mg)	0.7 (0.3)	0.7 (0.5)	0.8 (0.5)	0.6 (0.6)	0.7 (0.4)	0.6 (0.5)
Thiamin (mg)	1.0 (0.7)	0.8 (0.7)	0.9 (0.6)	0.8 (0.7)	0.9 (0.7)	0.7 (0.8)
Vitamin A (μg RAE)	1274 (1212)	904 (1087)	1636 (1853)	868 (1777)	1382 (1746)	724 (1331)
Vitamin B-6 (mg)	1.8 (1.0)	1.8 (1.1)	2.0 (1.3)	1.6 (1.7)	1.7 (1.2)	1.5 (1.5)
Vitamin B-12 (μg)	3.2 (3.5)	1.8 (4.4)	2.8 (4.2)	1.3 (2.9)	3.4 (5.4)	1.5 (3.4)
Vitamin C (mg)	97.4 (65.0)	91.0 (69.7)	111.5 (87.0)	85.0 (96.7)	93.5 (80.3)	72.9 (89.8)
Zinc (mg)	6.9 (3.6)	6.5 (3.7)	6.8 (3.9)	6.1 (4.9)	6.8 (3.8)	6.6 (6.0)

Abbreviations: FRANI, Food Recognition Assistance and Nudging Insights; IQR, interquartile range; RAE, retinol activity equivalents; SD, standard deviation.
[1] No. of person-days = 124, equal to number of subjects (62) multiplied by number of recalls (2).

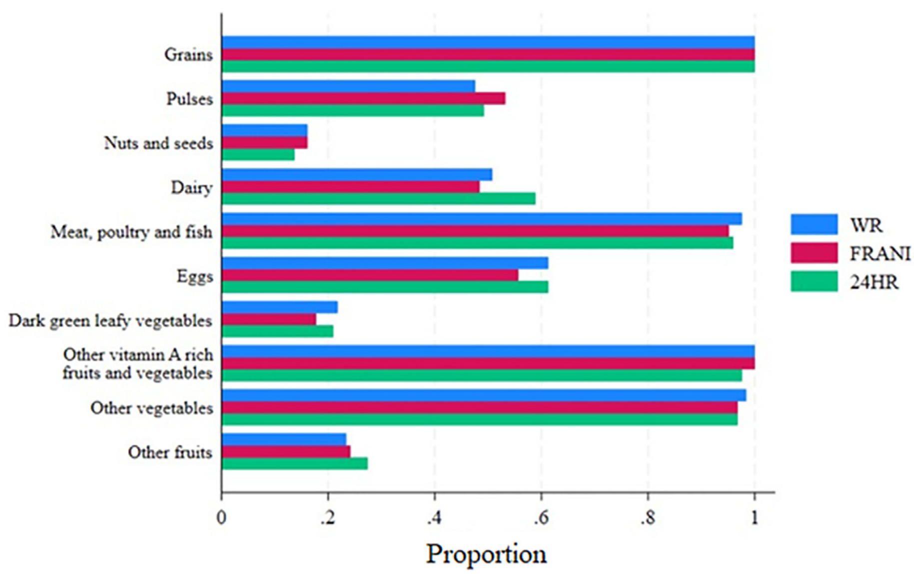


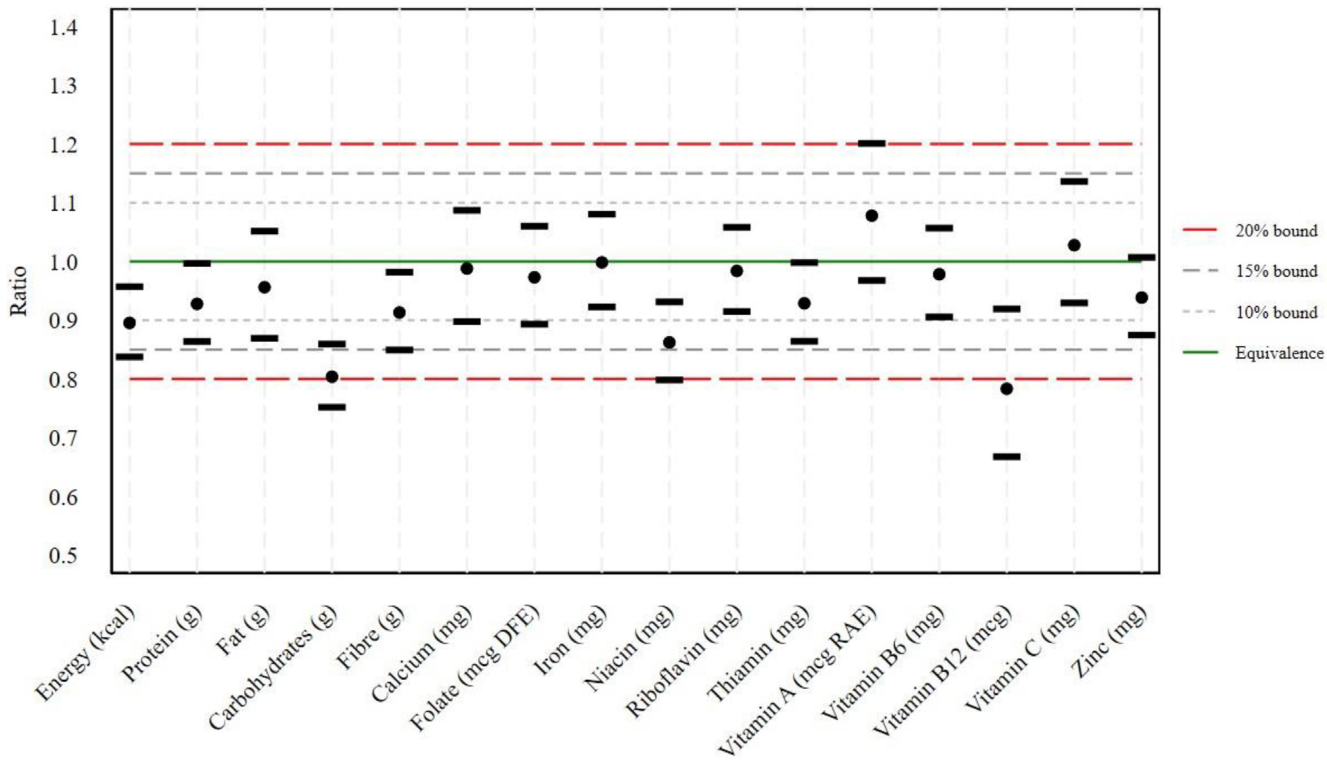
FIGURE 1. Proportion of food groups consumed by weighed records, FRANI app, and 24-hr recall. FRANI, Food Recognition Assistance and Nudging Insights; WR, weighed records; 24HR, 24-h recall. No. of person-days = 124, equal to number of subjects (62) multiplied by number of recalls (2).

vitamin C, and zinc intakes were equivalent at the 15% bound. Energy (mean ratio: 0.90; 90% CI: 0.84, 0.97), fiber, niacin, and vitamin A intakes were equivalent at the 20% bound. For FRANI method, vitamin B-12 was the only micronutrient considered where the equivalence ratio did not fall within the 20% bound. Comparisons between 24HR and WR intakes found that no nutrients fell within a 10% error bound. Protein, iron, niacin, riboflavin, and zinc intakes were equivalent at the 15% bound; folate, thiamine, and vitamin B-12 were equivalent at the 20% bound (Figure 2B, Supplemental Table 4). For 24HR method, equivalence ratios for intakes of calcium, vitamin A, vitamin B-6, and vitamin C did not fall within the 20% bound. The CCCs by nutrient between FRANI and WR ranged between 0.45 and 0.74 (mean: 0.60) and between 0.48 and 0.76 (mean: 0.63) for 24HR and WR (Figure 3, Supplemental Table 5).

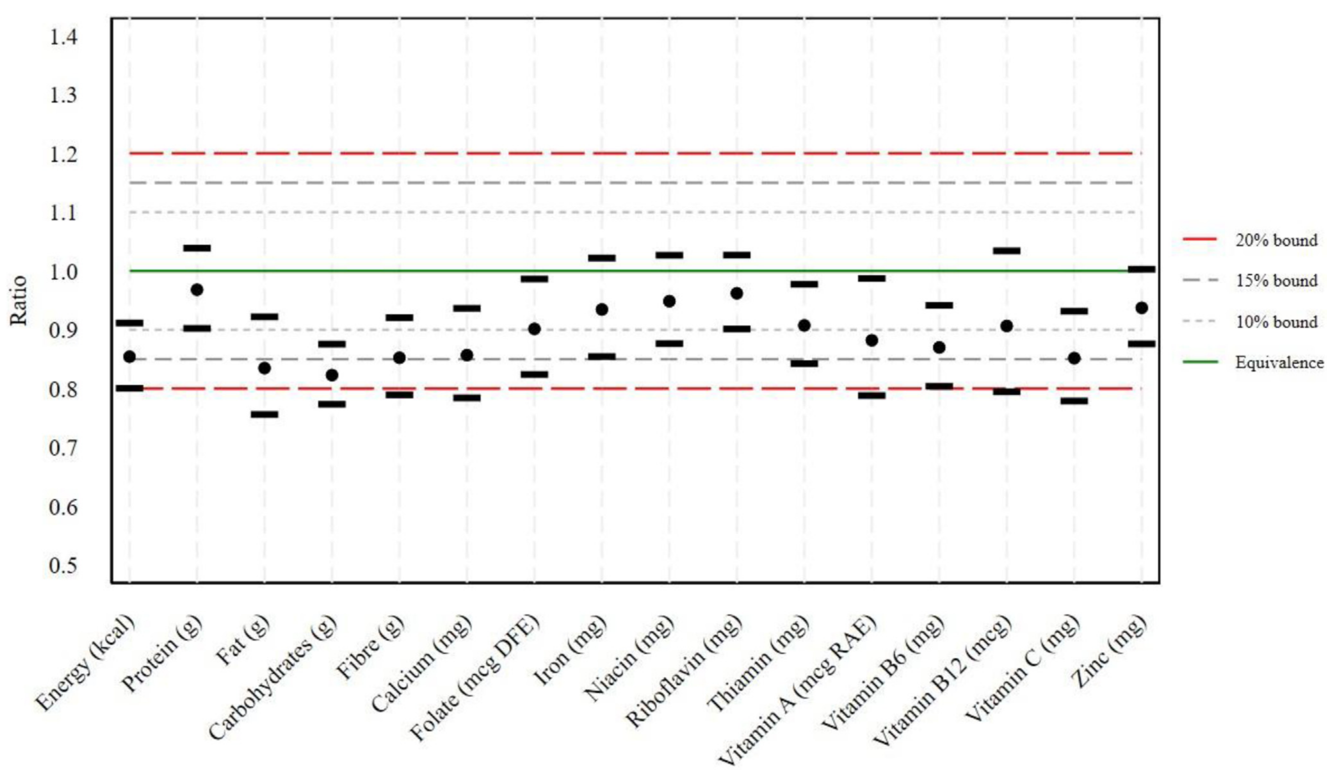
Sources of error

Examination of the proportion of energy intake consumed by food group highlighted mostly minor variations across FRANI and 24HR compared with those of WR (Table 3). Two notable exceptions among these comparisons with WR were the overestimation of the share of energy intake from dairy in the 24HR (4% compared with 9%) and the overestimation of the "other vegetables" share of energy intake in FRANI (12% compared with 16%). More substantive variations were found in the quantities of food consumed by foods with highest frequency of consumption (Table 4). FRANI and 24HR methods appeared to underestimate most quantities consumed of all food groups. Comparisons of food consumption episodes from FRANI with WR found levels of omission and intrusion errors of 15% and 22%, respectively. Conversely, when comparing 24HR with WR, omission errors were

A



B



(caption on next page)

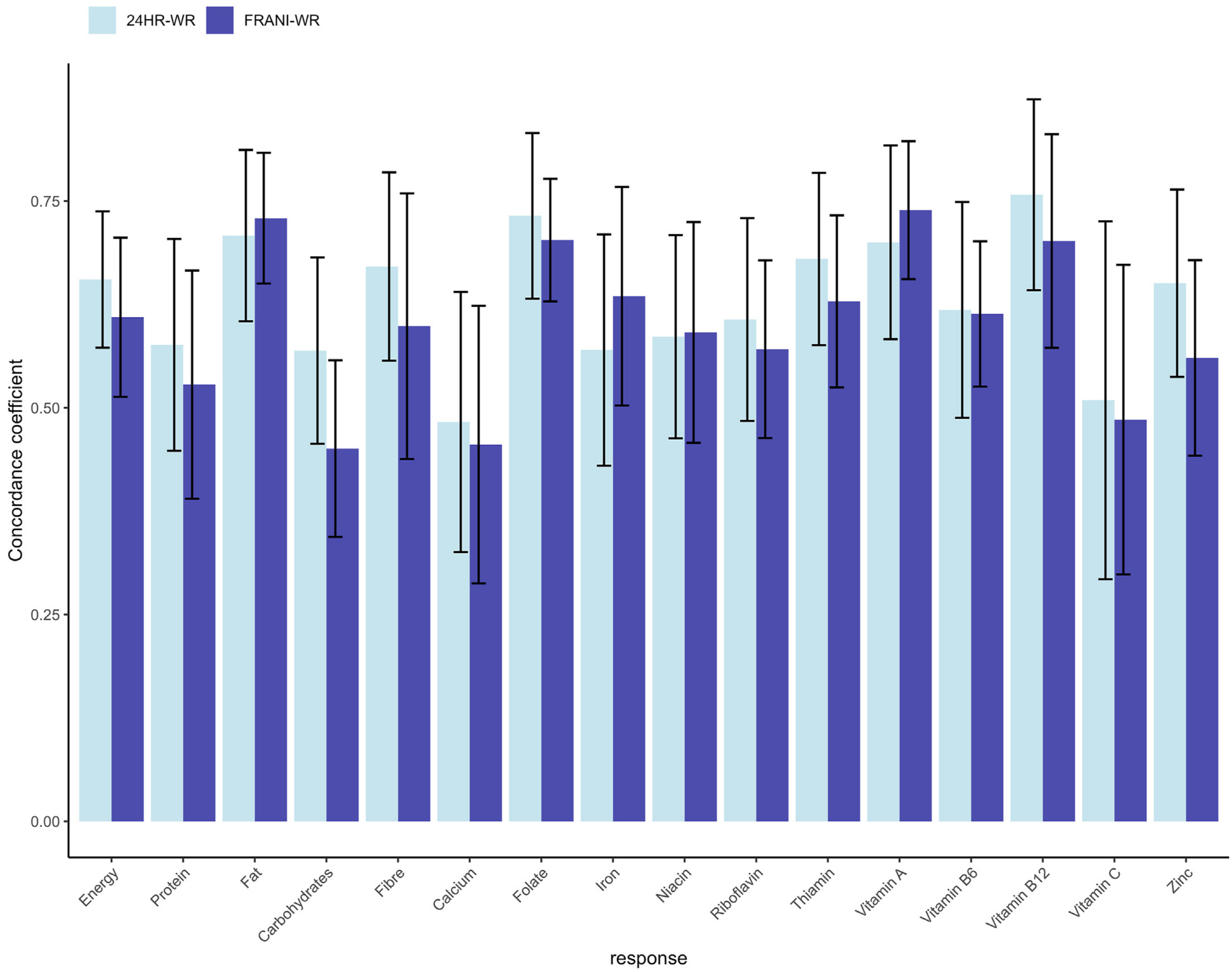


FIGURE 3. Concordance correlation coefficients of nutrient intakes on 2 d by weighed records, FRANI application, and 24HR. The concordance correlation coefficient (CCC) was estimated for each nutrient with adjustment for repeated measures. Error bars represent the 95% CIs. FRANI, Food Recognition Assistance and Nudging Insights; WR, weighed records; 24HR, 24-h recall. No. of person-days = 124, equal to number of subjects (62) multiplied by number of recalls (2).

higher (22%), while intrusion errors were lower (18%) than those reported for FRANI.

Discussion

This study contributes novel insights into the relative validity of using FRANI, an artificial intelligence–assisted mobile phone application, to assess the diets of female youth in LMICs. This validation study found that FRANI can accurately estimate energy intake in females aged 18–24 y in urban Ghana. Furthermore, for FRANI, equivalence was determined for 2 of the 11 micronutrients at the 10% bound, for 8 of the 11 micronutrients at 15% bound, and for 10 of the 11 micronutrients at the 20% bound. For the 24HR method, equivalence was determined for no micronutrients at the 10% bound, for 4 of the 11 micronutrients at 15% bound, and for 7 of the 11 micronutrients at the 20% bound. Analysis of the CCC for FRANI and 24HR compared with that for WR found broadly equivalent results. Errors of intrusion were lower for FRANI than that for 24HR, although omission errors were higher in FRANI than in those in 24HR. Although portion estimation for FRANI and 24HR involved both underestimation and overestimation, mean intakes trended toward underestimation.

In this study, FRANI demonstrated comparable accuracy found in previous validation studies targeting adolescent females aged 12–19 y in Ghana [9] and Vietnam [7]. In the previous study in Ghana in

FIGURE 2. Equivalence testing of ratios of nutrient intake on 2 d by weighed records, FRANI application, and 24-h recall: (A) FRANI application/weighed records; (B) 24-h recall/weighed records. Dotted lines showed 20% (red dash), 15% (gray long dash), and 10% (gray small dash) equivalence bounds. For example, (1 − ratio) × 100 is equal to the %error, and ratios between 0.9 and 1.1 are equivalent to a 10% bound around the mean %error. A 90% CI is used because two 1-sided tests are performed (each with α of 0.05). The ratio is back-transformed from the difference in the log-FRANI nutrient minus the log-weighed record nutrient intake or log 24-h–recalled nutrient minus the log-weighed record nutrient intake. Mean differences by method were estimated for each nutrient with regression models including random effects at the person level to account for repeated measures. FRANI, Food Recognition Assistance and Nudging Insights; WR, weighed records; 24HR, 24-h recall. No. of person-days = 124, equal to number of subjects (62) multiplied by number of recalls (2).

TABLE 4
Quantities of key food intakes from observed weighed records, FRANI application, and 24-h recall.

	No. of consumption episodes	Quantity (g), mean (SD)			Ratio	
		Weighed records	FRANI application	24-h recall	FRANI application/weighed records	24-h recall/weighed records
Tomato stew	43	77.4 (41.9)	87.0 (68.7)	105.7 (81.9)	1.1	1.4
Sugar	41	25.9 (12.3)	24.7 (16.8)	15.2 (9.9)	1.0	0.6
Chicken fried	40	53.1 (23.3)	59.0 (81.4)	86.9 (50.6)	1.1	1.6
Shito	32	26.1 (18.7)	76.0 (60.2)	19.6 (11.1)	2.9	0.8
Egg boiled	30	51.9 (9.0)	65.0 (109.3)	42.1 (7.4)	1.3	0.8
Rice jollof	26	244.9 (88.3)	456.1 (244.6)	218.6 (90.9)	1.9	0.9
Plantain ripe fried	25	114.9 (68.9)	145.5 (79.1)	120.5 (88.4)	1.3	1.0
Indomie	25	367.3 (155.5)	190.8 (71.6)	260.5 (158.3)	0.5	0.7
Bread butter	25	100.4 (42.3)	82.6 (51.9)	85.3 (39.1)	0.8	0.9
Rice plain	22	297.8 (80.8)	14.9 (8.1)	224.3 (88.2)	0.1	0.8
Fried egg	21	79.9 (34.2)	242.5 (206.5)	48.0 (18.9)	3.0	0.6
Milo beverage	19	210.0 (118.4)	171.9 (77.8)	31.2 (14.5)	0.8	0.1
Salad	19	69.9 (137.2)	22.0 (29.1)	86.9 (46.2)	0.3	1.2
Banku	18	365.2 (97.1)	250.4 (133.6)	344.3 (98.4)	0.7	0.9
Milk	18	49.3 (65.6)	42.6 (52.5)	31.6 (15.6)	0.9	0.6
Fried rice	17	262.5 (128.1)	198.6 (106.4)	236.5 (125.6)	0.8	0.9
Mixed vegetables stew	15	124.5 (70.3)	119.9 (48.5)	82.9 (49.0)	1.0	0.7
Spaghetti	15	192.2 (169.7)	116.8 (66.5)	187.4 (95.9)	0.6	1.0
Pepper sauce	15	61.3 (30.9)	75.8 (63.7)	29.3 (15.6)	1.2	0.5
Groundnut	14	27.3 (15.7)	54.9 (39.9)	66.9 (25.7)	2.0	2.4
Beef	14	68.4 (30.5)	68.3 (27.7)	75.7 (46.6)	1.0	1.1

Abbreviation: FRANI, Food Recognition Assistance and Nudging Insights.
[1]No. of person-days = 124, equal to number of subjects (62) multiplied by number of recalls (2).

females aged 12–19 y, equivalence for FRANI and WR was determined at the 10% bound only for energy intake, at 15% bound for 5 nutrients (iron, zinc, folate, niacin, and vitamin B-6), and at 20% bound for protein, calcium, riboflavin, and thiamine intakes [9]. In Vietnam, equivalence between FRANI and WR was determined at the 10% bound for energy, protein, and fat and 4 nutrients (iron, riboflavin, vitamin B-6, and zinc), at 15% bound for 3 nutrients (calcium, niacin, and thiamin), and at 20% bounds for all micronutrients except for vitamins A and B-12 [7]. However, omission errors for FRANI (15%) were considerably lower in this study than those in the previous study in Ghana (31%), whereas intrusion errors (22%) were higher than those previously reported in Ghana (16%) and at the same level to those reported in Vietnam (21%). There are some important differences between this study and the previous FRANI validation studies in the literature, including the age group of participants and the controlled environment in which participants were using FRANI. Unlike the previous validation studies in the literature, the procedures involved in this study removed any FRANI-related interaction between the WR enumerators and FRANI participants. In other words, once participants were trained on how to use FRANI, they used FRANI independently, without assistance from the research team.

The accuracy gains in estimation by FRANI reported in this study were a result of a combination of improvements in various components, including advancement in the underlying computer vision model for food classification, refinement in the calibration procedure for the portion estimation parameters, and enhancement in the underlying food databases including recipes, ingredients, and packaged foods. These findings highlight the potential for further improvements of the FRANI technology, as the project continues to improve computer vision capabilities and expand the lists of foods available for classification and user entry. Research is underway on portion size estimation procedures, where substantive gains can still be made, including the possibility of using depth information for volume estimation and expanding on the food weight coefficients used during the calibration of FRANI. Another important area of ongoing research includes using FRANI technology to provide real-time assessment of the quality of school meals being provided nationwide by the Government of Ghana.

To our knowledge, no other studies in the literature have assessed the validity of mobile technology for dietary assessment specifically considering its performance in youth in LMIC. A systematic review including 14 validation studies of mobile phone applications in high-income countries—with 2 focusing on adolescents in Sweden [15] and Korea [16]—examined performance of mobile applications mostly in relation to 24HR, with only 2 comparing with WR [17]. The review found that mobile applications slightly underestimated food consumption when compared with traditional dietary assessment methods. Underestimation has also been reported in validation studies of traditional 24HR in adults in LMIC using similar procedures to those used in this study [18,19]. For energy intake, ratios of 24HR/WR were 0.90 in rural Ethiopian females ($n = 58$), 0.94 in females in rural Kenya ($n = 42$), with similar results found in adolescents in Burkina Faso ($n = 105$) [6]. The evidence presented in this validation study is consistent with that reported in the emerging literature on this topic, highlighting the potential for new technologies to improve diet assessment methods and potentially reduce data collection costs.

Strengths and limitations

The strengths of this study are notable. First, it leverages comparisons with WR, the gold standard for dietary assessment, and the 24HR method commonly used in large-scale surveys. This dual approach enhances the robustness and reliability of the findings. Second, the study allowed participants the freedom to interact with FRANI independently. In contrast to previous studies of FRANI conducted in a relatively controlled environment, where FRANI use could be assisted by WR enumerators, this study reflects real-world usage scenarios, potentially providing a more realistic assessment of FRANI's performance. Although the study procedures emphasized minimizing interactions with regards to FRANI use and the WR enumerators, the WR inevitably involved interaction with participants, which may in turn have affected the FRANI records, although possibly not the accuracy of FRANI but rather may have reminded participants to use FRANI to capture their intake. Similarly, the 24HR is likely to have been affected by both the WR procedures and the image capture through FRANI, likely improving the 24HR performance compared with a typical 24HR survey without WR or FRANI use. However, evidence from a study in Vietnam where participants used FRANI independently (without WR) over a 30-d period found that consumption was on similar levels to that reported in a parallel validation study in the same population that included WR [7,20]. Finally, this study population is also double that of previous validation studies of FRANI. However, certain limitations also require consideration. Participants were randomly selected from the University of Ghana Legon campus in urban Accra, possibly limiting the external validity of these findings to this specific population. In addition, although the expectation was that the WR would cover a full 24-h period, for logistical reasons, enumerators were only able to cover the interval between early morning (06:00) to late evening after the last meal had supposedly been consumed. To ensure comparability across the 3 methods, we matched the consumption times across the 3 methods, excluding foods that were reportedly consumed outside this WR period.

In conclusion, both FRANI-assisted dietary assessment and 24HRs were found to accurately estimate nutrient intake in female youth 18–24 y in urban Ghana. Although improvements in the computer vision–assisted classification and portion estimation are possible, emerging evidence in this innovative research area suggests that this technology is ready for scale-up.

Author contributions

The authors' responsibilities were as follows – AG, GF: designed the research; PMC, SB, DH: developed FRANI AI; GF, BB, GA, VA, MA: conducted the research; AG, ON: analyzed and interpreted the data; AG, ON: wrote the article; GF, PHN, PMC, DH: provided critical revision of the manuscript for important intellectual content; and all authors: read and approved the final manuscript.

Funding

This study was conducted as part of the CGIAR Research Initiative on Resilient Cities. We would like to thank all funders who supported this research through their contributions to the CGIAR Trust Fund: https://www.cgiar.org/funders/

Data availability

Data described in the manuscript, code book, and analytic code will be made available upon request pending application.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajcnut.2024.08.011.

References

[1] GBD 2017 Diet Collaborators, Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017, Lancet 393 (10184) (2019) 1958–1972, https://doi.org/10.1016/S0140-6736(19)30041-8.

[2] B.M. Popkin, C. Corvalan, L.M. Grummer-Strawn, Dynamics of the double burden of malnutrition and the changing nutrition reality, Lancet 395 (10217) (2020) 65–74, https://doi.org/10.1016/S0140-6736(19)32497-3.

[3] D. Hargreaves, E. Mates, P. Menon, H. Alderman, D. Devakumar, W. Fawzi, et al., Strategies and interventions for healthy adolescent growth, nutrition, and development, Lancet 399 (10320) (2022) 198–210, https://doi.org/10.1016/S0140-6736(21)01593-2.

[4] W. Bell, B.A. Colaiezzi, C.S. Prata, J.C. Coates, Scaling up dietary data for decision-making in low-income countries: new technological frontiers, Adv. Nutr. 8 (6) (2017) 916–932, https://doi.org/10.3945/an.116.014308.

[5] R.S. Gibson, E. Ferguson, An interactive 24-hour recall for assessing the adequacy of iron and zinc intakes in developing countries, Intereconomics, International Life Sciences Institute (2008). Washington, DC.

[6] J.E. Arsenault, M. Moursi, D.K. Olney, E. Becquey, R. Ganaba, Validation of 24-h dietary recall for estimating nutrient intakes and adequacy in adolescents in Burkina Faso, Matern, Child Nutr. 16 (4) (2020) e13014, https://doi.org/10.1111/mcn.13014.

[7] P.H. Nguyen, L.M. Tran, N.T. Hoang, D.T.T. Trương, T.H.T. Tran, P.N. Huynh, et al., Relative validity of a mobile AI-technology-assisted dietary assessment in adolescent females in Vietnam, Am. J. Clin. Nutr. 116 (4) (2022) 992–1001, https://doi.org/10.1093/ajcn/nqac216.

[8] K.P. Adams, W. Bell, J.W. Somé, B. Colaiezzi, S. Wafa, B. Rogers, et al., The cost and cost efficiency of conducting a 24-h dietary recall using INDDEX24, a mobile dietary assessment platform, compared with pen-and-paper interview in Viet Nam and Burkina Faso, Br. J. Nutr. 129 (3) (2022) 535–549, https://doi.org/10.1017/S0007114522001362.

[9] G.K. Folson, B. Bannerman, V. Atadze, G. Ador, B. Kolt, P. McCloskey, et al., Validation of mobile artificial intelligence technology-assisted dietary assessment tool against weighed records and 24-hour recall in adolescent females in Ghana, J. Nutr. 153 (8) (2023) 2328–2338, https://doi.org/10.1016/j.tjnut.2023.06.001.

[10] B. Stadlmayr, U. Charrondiere, V.N. Enujiugha, R.G. Bayili, E.G. Fagbohoun, B. Samb, et al., West African food composition table/Table de composition des aliments d'Afrique de ĺOuest, FAO, Rome, Italy, 2012.

[11] FHI 360 FAO, Minimum dietary diversity for women: a guide to measurement, FAO, Rome, Italy, 2016, p. 82.

[12] M.J. Batterham, C. Van Loo, K.E. Charlton, D.P. Cliff, A.D. Okely, Improved interpretation of studies comparing methods of dietary assessment: combining equivalence testing with the limits of agreement, Br. J. Nutr. 115 (7) (2016) 1273–1280, https://doi.org/10.1017/S0007114516000040.

[13] F.E. Thompson, S. Dixit-Joshi, N. Potischman, K.W. Dodd, S.I. Kirkpatrick, L.H. Kushi, et al., Comparison of interviewer-administered and automated self-administered 24-hour dietary recalls in 3 diverse integrated health systems, Am. J. Epidemiol. 181 (12) (2015) 970–978, https://doi.org/10.1093/aje/kwu467.

[14] L.I.K. Lin, A concordance correlation coefficient to evaluate reproducibility, Biometrics 45 (1) (1989) 255–268. http://www.jstor.org/stable/2532051.

[15] Å. Svensson, C. Larsson, A mobile phone app for dietary intake assessment in adolescents: an evaluation study, JMIR Mhealth Uhealth 3 (4) (2015) e93, https://doi.org/10.2196/mhealth.4804.

[16] J.E. Lee, S. Song, J.S. Ahn, Y. Kim, J.E. Lee, Use of a mobile application for self-monitoring dietary intake: feasibility test and an intervention study, Nutrients 9 (7) (2017) 748, https://doi.org/10.3390/nu9070748.

[17] L. Zhang, A. Misir, H. Boshuizen, M. Ocké, A systematic review and meta-analysis of validation studies performed on dietary record apps, Adv. Nutr. 12 (6) (2021) 2321–2332, https://doi.org/10.1093/advances/nmab058.

[18] C.A. Gewa, S.P. Murphy, C.G. Neumann, A comparison of weighed and recalled intakes for schoolchildren and mothers in rural Kenya, Public Health Nutr 12 (8) (2009) 1197–1204, https://doi.org/10.1017/S1368980008003698.

[19] A.A. Alemayehu, Y. Abebe, R.S. Gibson, A 24-h recall does not provide a valid estimate of absolute nutrient intakes for rural women in southern Ethiopia, Nutrition 27 (9) (2011) 919–924, https://doi.org/10.1016/j.nut.2010.10.015.

[20] B.C. Braga, P.H. Nguyen, L.M. Tran, N.T. Hoang, B. Bannerman, F. Doyle, et al., Feasibility of using an artificial intelligence-based telephone application for dietary assessment and nudging to improve the quality of food choices of female adolescents in Vietnam: evidence from a randomized pilot study, Curr. Dev. Nutr. 8 (6) (2023) 102063, https://doi.org/10.1016/j.cdnut.2023.102063.