



Validation of the INDDX24 mobile app v. a pen-and-paper 24-hour dietary recall using the weighed food record as a benchmark in Burkina Faso

Beatrice Rogers^{1*}, Jérôme W. Somé^{1,2}, Peter Bakun¹, Katherine P. Adams³, Winnie Bell¹, David Alexander Carroll II¹, Sarah Wafa¹ and Jennie Coates¹

¹Tufts University, Gerald J. and Dorothy R. Friedman School of Nutrition Science and Policy, Boston, MA, USA

²Institut de Recherche en Sciences de la Santé, Ouagadougou, Burkina Faso

³University of California, Davis, Institute for Global Nutrition, Department of Nutrition, Davis, CA, USA

(Submitted 21 June 2021 – Final revision received 14 November 2021 – Accepted 23 November 2021 – First published online 26 November 2021)

Abstract

Effective nutrition policies require timely, accurate individual dietary consumption data; collection of such information has been hampered by cost and complexity of dietary surveys and lag in producing results. The objective of this work was to assess accuracy and cost-effectiveness of a streamlined, tablet-based dietary data collection platform for 24-hour individual dietary recalls (24HR) administered using INDDX24 platform v. a pen-and-paper interview (PAPI) questionnaire, with weighed food record (WFR) as a benchmark. This cross-sectional comparative study included women 18–49 years old from rural Burkina Faso (*n* 116 INDDX24; *n* 115 PAPI). A WFR was conducted; the following day, a 24HR was administered by different interviewers. Food consumption data were converted into nutrient intakes. Validity of 24HR estimates of nutrient and food group consumption was based on comparison with WFR using equivalence tests (group level) and percentages of participants within ranges of percentage error (individual level). Both modalities performed comparably estimating consumption of macro- and micronutrients, food groups and quantities (modalities' divergence from WFR not significantly different). Accuracy of both modalities was acceptable (equivalence to WFR significant at $P < 0.05$) at group level for macronutrients, less so for micronutrients and individual-level consumption (percentage within $\pm 20\%$ for WFR, 17–45% for macronutrients, 5–17% for micronutrients). INDDX24 was more cost-effective than PAPI based on superior accuracy of a composite nutrient intake measure (but not gram amount or item count) due to lower time and personnel costs. INDDX24 for 24HR dietary surveys linked to dietary reference data shows comparable accuracy to PAPI at lower cost.

Key words: Dietary recall: Survey methods: Tablet-based data collection: Accuracy: Cost-effectiveness

To be effective in addressing nutrition and health challenges, food and nutrition policies must be based on an empirical understanding of food consumption patterns: dietary adequacy, the nutritional value of foods, food safety and food sources. In low- and middle-income countries (LMIC), the framing of food security and nutrition problems increasingly encompasses not only under-nutrition but also diet quality⁽¹⁾ and overweight and obesity, with their attendant chronic diseases^(2,3). Issues of the safety of the food supply, including microbiological and mycotoxin contamination^(4,5) and potentially risky additives⁽⁶⁾, are emerging concerns of food policy in LMIC. Detailed information on individual food consumption, collected in a survey that is appropriately representative of the relevant population, is critical for making informed policies and developing effective programmatic responses to assure nutritional health^(7,8); lack of individual-level dietary data is a barrier to developing effective solutions⁽⁹⁾.

A wide variety of qualitative and semi-quantitative methods are currently used to assess the adequacy of individual dietary intakes in LMIC including dietary diversity scores^(10–15) and FFQ^(16,17), but their use as a measure of diet quality has been questioned⁽¹⁸⁾, and quantitative estimates yielded by these methods are imprecise^(19–22). Household food consumption data derived from household consumption and expenditure surveys have been used to estimate individual consumption but do not account for the intrahousehold allocation of food^(23–25) and may miss consumption of food away from home⁽²⁶⁾. Given these limitations, the detailed information derived from individual quantitative dietary recalls is more appropriate for many types of analyses and policy decisions.

Food policies at the national level manifest their effects at the level of the individual; to make informed food policies, governments need quantitative individual data. Such data allow for

Abbreviations: CAPI, computer-assisted personal interview; FMDB, Food Matters Database; LMIC, low- and middle-income country; NSR, non-standard recipe; PAPI, pen-and-paper interview; PSEM, portion size estimation method; WFR, weighed food record; 24HR, 24-hour dietary recall.

* **Corresponding author:** Beatrice Rogers, email beatrice.rogers@tufts.edu

measuring important policy-relevant variables, including how foods are prepared (reflecting the preservation of their nutritional quality) and their source (important for understanding potential vehicles for fortification or other interventions and potential contamination risk). In most LMIC, individual quantitative dietary data have been collected infrequently if at all, due to issues related to cost, time, and complexity^(7,27).

Quantitative methods for assessing diets include single or multiple 24-hour dietary recalls (24HR), food diaries and weighed food record (WFR)^(28,29). The interviewer-administered quantitative 24HR is often viewed as most appropriate for use in dietary surveys, providing more detailed information than the qualitative methods described above; recalls of time periods beyond the previous 24 h are not commonly used as they are subject to serious recall error^(30,31). Barriers to use of the 24HR include the time required for the interview, need for more preparation in advance of data collection, more complex data analysis, and the need for a comprehensive set of dietary reference data to convert consumption data into detailed nutrient information.

A 24HR may be administered as a paper-based questionnaire (pen-and-paper interview (PAPI)) or on a tablet or other device (computer-assisted personal interview (CAPI)). Paper-based questionnaires require manual data entry, an error-prone process entailing a long lag between data collection and the availability of results. In contrast, CAPI-based methods have the advantage of electronic data capture and a direct link to reference data such as food composition information, eliminating the need for manual data entry and greatly reducing the time for data processing and analysis.

The INDDEx Project created a mobile data collection application (INDDEx24) for administering a 24HR using the multiple-pass method^(30–33), linked to the Global Food Matters Database (FMDB) where the dietary reference data are stored. The foods entered into the recall were drawn from a food list developed in advance and coded so as to link directly to the dietary reference data for analysis; the mobile application provides a standardised electronic data collection platform, customisable to meet the specific needs of a given survey, an architecture that links the data directly to the dietary reference data, and an automated analysis tool that streamlines the production of key results.

The present study reports the results of a comparison of a 24HR administered using a paper-based questionnaire (PAPI) with the CAPI approach using INDDEx24 in terms of accuracy (using the WFR as the benchmark) and cost-effectiveness. Comparisons of time required from initial preparation through production of key results and of cost are reported elsewhere⁽³⁴⁾. Few studies have compared the 24HR CAPI *v.* PAPI approach in a LMIC context. Several studies have assessed relative cost of 24HR using CAPI *v.* PAPI and identified key cost contributors^(32,35–39). None of these studies related cost to the accuracy of the data collected.

Subjects and methods

Study setting, design and participants

The study took place in 2019 in Ouhitenga province in the Plateau Central region of Burkina Faso. The validation study

component was conducted in Moackin village in the rural commune of Absouya, and the interview time assessment study component in Ziniare in the urban commune of Ziniare and in Goue in the rural commune of Loumbila. While the 24HR can be administered as part of a representative survey, the goal of the study was not to represent the population but rather to compare the two administration methods; thus, the study site was purposively selected to be roughly comparable in key socio-demographic characteristics (e.g. economic status and education level) with the average levels in Burkina Faso.

The validation study was a cross-sectional comparative study calibrating two dietary recall methods against a benchmark WFR. The study had two arms. For the first arm (WFR-PAPI), respondents participated in the WFR observation during the first day of data collection and were interviewed about the foods consumed on the observation day using the 24HR PAPI the second day. For the second arm (WFR-INDDEx24), the same procedures were carried out except that participants were interviewed using the INDDEx24 the second day of data collection. To implement the interview time assessment study, a similar design was used with two sub-arms (INDDEx24 and PAPI) but with no WFR for this component.

Participants were women aged 18–49 years living in the study areas. One female member within the eligible age range was enrolled from each household. Exclusion criteria for participation included any mental or physical health issues that could affect the participant's ability to complete the 1-d observation and the interview. Participants needed to agree to participate in both phases of data collection (observation for WFR and 24HR interview).

Recruitment and allocation of study participants

A household listing was conducted in Moackin village to identify households with eligible participants and was used to build the sampling frame from which households were randomly selected for the validation study. For the time assessment study, thirty households were systematically selected in each of the two other sites (Ziniare and Goue). In each household, one female member within the eligible age range who was available and consented was enrolled.

For the validation study, participants were contacted by a 'consenting agent' who was assisted in locating the selected households by a local guide. The consenting agent first obtained verbal agreement from the household head and then the agent talked with one eligible, available and interested woman of the household to obtain a written consent. Study procedures were clearly explained during the consenting process. Participants were informed that they would be participating in 2 d of data collection, the first for the WFR and the second for the 24HR. After that, the agent scheduled the WFR visit for a day within the next 2 d and arranged the time for the interviewer to arrive before the respondent started preparing the first meal of the day. The consenting agent provided the participant with a bowl and plate, instructed her to use them for taking her meals and explained that she should not change her diet. Equal numbers of participants were recruited in each study arm.

Data collection

Data collection for WFR and 24HR (INDDEX24 and PAPI) took place in September–October 2019 and was carried out on all days of the week. A total of eight 24HR interviewers and sixteen WFR interviewers were trained for 14 and 10 d, respectively, to collect the data. 24HR interviewers were trained on both INDDEX24 and PAPI and were assigned to conduct one or the other modality alternating by day so that all interviewers conducted interviews using both modalities on all days of the week. Data collection was overseen by two supervisors for the 24HR and three for the WFR. Two individuals trained on 24HR but not retained were hired as consenting agents in charge of participants' recruitment.

Weighted food record. Interviewers arrived at the participant's house in time to observe the preparation and/or the consumption of the first meal of the day and left after consumption of the last main meal of the evening. Interviewers were trained to remain as discreet and unobtrusive as possible when in the participant's house between meal preparation and/or consumption periods. They accompanied the participant if she left the house and were trained to record any food consumed away from home. Throughout the WFR observation day, interviewers recorded all details about foods and mixed dishes prepared in the home (quantity of each ingredient, cooking method, total quantity of dish prepared, weight of the pot/pan), quantity of the food consumed by weighing the food and beverages before and after the participant had consumed them using a digital scale (MyWeigh KD 7000 with 1-g accuracy and 7-kg capacity), and time and place of preparation and consumption for all food items. For purchased foods, foods prepared outside of the participant's household or leftovers from the previous day, the interviewers were instructed to record as many details as possible regarding the food or mixed dish, including the main ingredients, cooking method, and any other distinguishable characteristics, along with quantities. At the end of the WFR observation day, the interviewer arranged with the respondent the time for the 24HR visit the next day and later provided this information to the 24HR team for planning the visit of the 24HR interviewers.

24-Hour dietary recall

For the 24HR, a different group of interviewers visited the participant in her house the day after the WFR observation and carried out the 24HR. Before administering the 24HR, interviewers completed the registration of the participant, which consisted of recording a same set of information as for the WFR (e.g. date, household ID, location, birthdate, pregnant/lactating) to link 24HR and WFR data.

For the INDDEX4 interview, interviewers recorded first the full list of all foods, beverages and mixed dishes recalled as consumed by the participant during the previous 24 h and then probed for more details using the food and recipe list developed by the study and integrated in the INDDEX24 app (online Supplemental Information); if the food or recipe was not in the list, the interviewer directly input the required details. Following selection of the food or recipe from the food and recipe list in INDDEX24 (or inputting of a missing food item or non-

standard recipe (NSR) name), the amount (portion size) consumed by the participant was estimated using the portion size estimation methods (PSEM) developed by the study team and integrated in the INDDEX24 app (online Supplemental Information). PSEM options used for INDDEX24 included direct weight, life-sized photos, proxy weight using play dough, proxy weight using dried sorghum, proxy weight using water, and standard units. In the case of mixed dishes, interviewers were instructed to select the standard recipe that best fit the description of the mixed dish, based on the participant's report of ingredients used, with flexibility to accept minor divergences from the standard recipe. The acceptable differences included the addition/subtraction of minor ingredients (e.g. herbs, seasoning and spices), the use of a different type of cooking oil, the omission/addition of oil, and the use of condiments such as tomato, onion, bell pepper generally used in small amounts in the study context. If no close match was found for a reported mixed dish, it was considered as a NSR, and its details (ingredient description and quantities, and total quantity of recipe prepared) were collected by the interviewers with the participants or the dish preparer when possible. If the NSR was like an existing standard recipe, the interviewer could copy some/all the ingredients and use the copied version as a base from which to build the NSR. For recipes purchased or prepared by another person not available to provide details and considered as non-standard, participants were asked to report as many details as possible (ingredients, cooking method, place of purchase/acquisition and whatever other information could be provided).

For 24HR PAPI interview, interviewers used a paper data collection form and printed food and recipe booklets that contained all the same items as the INDDEX24 food and recipe list. During the 24HR PAPI interview, interviewers recorded the detailed description of the food or recipe reported by the participant and identified the best match in the food or recipe books, and then recorded the relevant code on the paper form. If the interviewer could not find a food or recipe in the books, they would describe it in detail and code it as '9999' for food and '99 999' for recipe. The same approach was taken for the standard and NSR selection in INDDEX24, and the same rules were applied to the PAPI data collection. To confirm whether mixed dishes described by the participants were among those listed as 'standard recipes', interviewers located the closest recipe in their printed recipe booklet and asked the respondent to list all the ingredients used in preparation of her recipe. Interviewers were trained to consider the recipe as non-standard if there were significant differences between the lists of ingredients (as described earlier); in that case, the recipe details were recorded including the amount prepared, the quantities and description of ingredients used. To estimate amounts consumed of foods and mixed dishes reported, PAPI interviewers had the same list of PSEM options (online Supplemental Information) as for INDDEX24 and were trained to identify which would be the best PSEM option for each food. A list of standard units and their codes was developed for use during PAPI interviews. When a standard unit was selected as the PSEM option, the interviewer asked the respondent about the standard unit size/volume (e.g. small or large bottle of Coca-Cola, French baguette – with weight or volume conversions automatically applied) of the food and the



portion of the standard unit they consumed (e.g. 1 unit, half a unit, etc.). Then, interviewers selected the corresponding standard unit code within the list and recorded the code and the portion consumed in the PAPI form.

After the 24HR PAPI and INDDDEX24, the interviewers administered the socio-demographic module and concluded the visit. For each day of 24HR data collection, interviewers were assigned two to three respondents with whom to conduct interviews using one type of 24HR modality (either INDDDEX24 or PAPI). On average, fourteen WFR observations were conducted per day along with six to eight INDDDEX24 and six to eight PAPI interviews following the WFR the next day.

Outcomes

Food group and nutrient intakes were estimated based on foods, beverages, and mixed dishes consumed and recorded during the WFR and both 24HR modalities (INDDDEX24 and PAPI) using nutrient composition from the food composition table compiled by the study team (online Supplemental Information). The FAO/WHO GIFT food groups classification was used to group the foods, beverages and mixed dishes recorded in order to estimate intake by food group. For the time assessment study, the mean 24HR interview time and time spent on individual passes were estimated for both INDDDEX24 and PAPI.

Sample size

The sample size was estimated as follows. First, we assumed that any correlation lower than 0.6 between the WFR-INDDDEX24 and WFR-PAPI would be considered unsatisfactory. Second, we considered an increase in correlation of 0.15 or greater (correlation 0.75 or greater) a practically important improvement. Lastly, we set the power at 80% and α level at 5%. Given these conditions, we estimated the sample size to be 104 for each arm. It was inflated to 117 to account for approximately 12% non-response. The final number of respondents was 116 for INDDDEX24 and 115 for PAPI (Fig. 1).

Ethical approvals and consent to participate

The study protocol and instruments were approved by the Ethics Committee for Health Research of the Ministry of Health in Burkina Faso and the Institutional Review Board of Tufts University in USA. All participants provided a written informed consent before the enrolment in the study.

Data management, cleaning and processing

For the INDDDEX24, supervisors did some spot checking on tablets immediately after the interview, but most of the review was done by the investigators at the end of daily data collection after exporting the data in CSV. format. INDDDEX24 data were reviewed for any outliers in gram weights reported, errors in food categorisation/coding, and completeness of NSR. Supervisors discussed and resolved any issues by early the following day.

Supervisors checked data quality of paper forms for WFR and 24HR PAPI daily, and any issues identified were discussed with the interviewers and resolved. WFR and 24HR PAPI data were

entered into a programme built on CSPro version 7.3 using double entry. Hard copies of WFR and PAPI forms were used by the data entry supervisor to resolve any discrepancies. Study coordinators carried out subsequent cleaning to ensure appropriate use of food, recipe and ingredient codes, PSEM codes (PAPI only), value plausibility for amounts/quantities and number of portions, and household ID.

Any new foods or NSR (including food away from home and food prepared on the previous day) or PSEM conversion factors that did not already exist in the INDDDEX24 dietary reference data were added to ensure that they had a match with the WFR, PAPI and/or INDDDEX24. For the food data processing and conversion into nutrients, the INDDDEX24 Global FMDB of dietary reference data (portion conversion factors, standard recipe information and food composition data) was used for WFR and PAPI using SAS v9.4. INDDDEX24 analytical reporting feature was used to match all foods and convert them into nutrients automatically. The nutrient intake data sets for WFR, INDDDEX24 and PAPI were then merged in order to complete the study analysis.

Statistical analyses

We conducted a comparison of the socio-demographic characteristics between the two study arms (WFR-INDDDEX24 *v.* WFR-PAPI) to check the success of randomisation. Nutrient intake data were checked for outliers by reviewing histograms and scatter plots by nutrient. By convention, energy intakes <2092 kJ or >20920 kJ (< 500 or > 5000 kcal)/d were flagged as outliers^(40–42). Any INDDDEX4 or PAPI that differed from the respective WFR by > 4184 kJ (1000 kcal) in either direction was also flagged for review. Both types of outliers were investigated by looking at the associated WFR or PAPI forms and the INDDDEX24 analytical report exports for INDDDEX24. Outliers that were due to data entry errors were corrected after reference to the paper forms. Remaining outliers were examined but maintained in the data set. The energy intakes from WFR for three INDDDEX24 and five PAPI respondents were flagged as implausible values, and these respondents were excluded from the validation comparison analysis. Shapiro–Wilk tests were used to test for normality, and all nutrients were transformed (log or cube root) to correct for skewness. Nutrients included in the analysis were energy (kJ [kcal]), fat (g), protein (g), carbohydrates (g), total fibre (g), vitamin A (mcg RAE), vitamin C (mg), Ca (mg), Fe (mg) and Zn (mg). Bland–Altman plots were generated to assess visually the agreement between the WFR and the 24HR method for grams consumed, energy, and all nutrients of interest^(43,44). Bland–Altman plots show individual differences between the two methods and were used to examine the mean bias, 95% limits of agreement, and the distribution of bias.

To assess the relative accuracy of the two recall methods, we compared each method to the benchmark (WFR-INDDDEX24 and WFR-PAPI) for each nutrient of interest and for the energy contribution from the different FAO/WHO/GIFT food groups. Following the method of Arsenault *et al.*⁽⁴⁵⁾, equivalence tests were done using the two one-sided paired *t* test between recall method and WFR at the group level. A 10% bound was used to assess equivalence, as suggested by other researchers^(45,46). A significant $P < 0.05$ for the equivalence test indicates that the

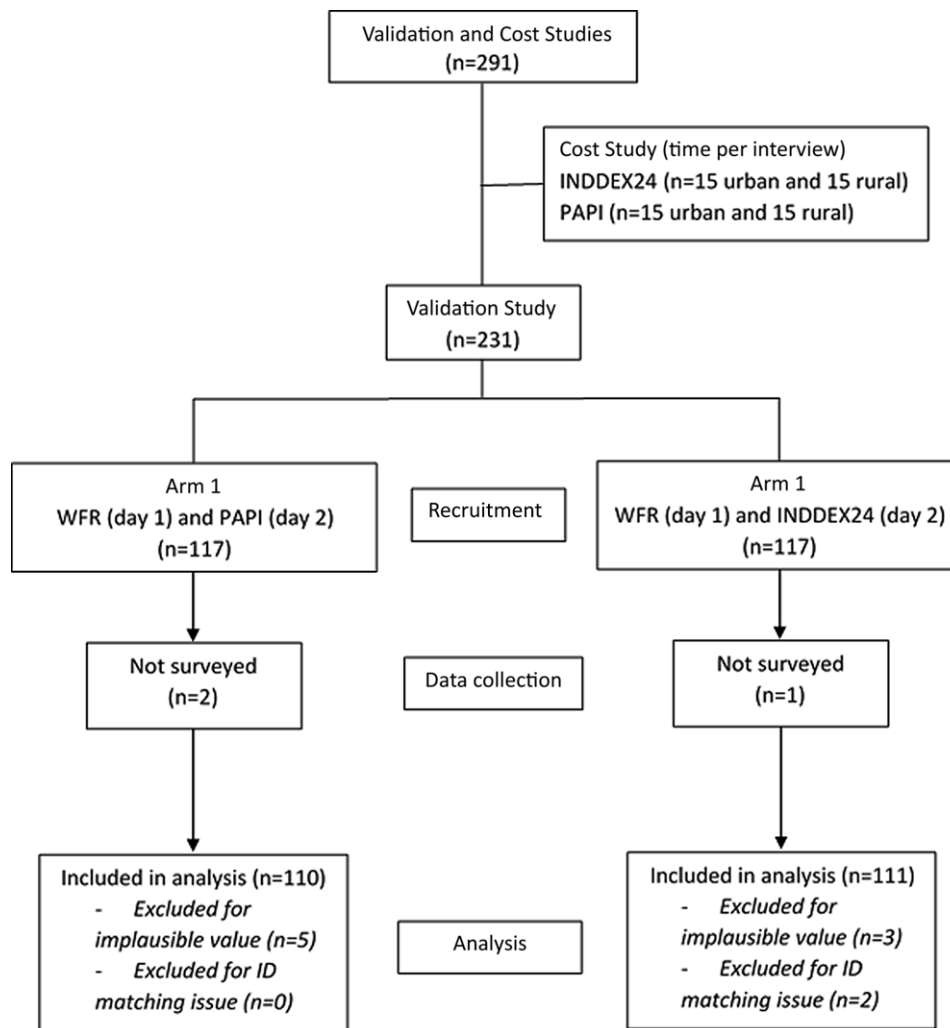


Fig. 1. Flow diagram for INDDEX24 validation and cost studies in Burkina Faso. PAPI, pen-and-paper interview; WFR, weighed food record.

WFR is equivalent to the 24HR at 10% bound; non-significance means the two are statistically significantly different. After equivalence testing, we compared the two recall methods to each other and assessed the relative magnitude of their differences from the WFR using a difference-in-differences approach; comparing the WFR-PAPI difference to the WFR-INDDEX24 difference using a random intercept mixed effects regression model to compare slopes between 24HR methods; this comparison was also done at the group level. In addition, the percentage of INDDEX24 and PAPI respondents with energy and nutrient intakes that fell within specific percentage error categories compared with the WFR was used to assess accuracy of both 24HR modalities at the individual level. This approach allows for assessing the extent to which each modality (INDDEX24 and PAPI) is accurate at the individual level relative to the WFR estimates. All analyses were performed using SAS 9.4 or Stata 15 se.

Cost-effectiveness analyses

In order to estimate and compare the cost of using the CAPI modality via INDDEX24 and the PAPI modality to conduct a 24HR and produce a clean, analysable 24 h data set, we

conducted a costing study alongside the validation study in Burkina Faso⁽³⁴⁾. We took an activity- and ingredients-based approach to collecting and processing the cost data for each modality, and costs were collected from a societal perspective, meaning that all costs, including the opportunity cost of respondents' time, were accounted for⁽⁴⁷⁾.

Prior to data collection, we defined the set of activities required to conduct the 24HR and prepare an analysable data set, including the preparation of dietary reference data, survey preparation (including the procurement of supplies and equipment), training, survey execution, data entry, and data cleaning and processing. Then, during the validation study, we recorded the types and quantities of inputs, or ingredients (e.g. personnel, facilities, transportation, equipment and supplies), that were used to carry out each activity. Throughout the validation study, this modality-specific time use and expenditure information was recorded by study personnel using a set of data collection instruments (paper-based quick logs for use in the field, time use logs housed in Google Forms, and Excel-based time and expenditure reporting logs). Time use data for study personnel were valued using wages and salaries, while respondent time was valued at



the minimum wage in Burkina Faso⁽⁴⁸⁾. All expenditures were converted from West African CFA francs to USA dollars where necessary, adjusted to 2019 USA dollars, and summed in order to estimate the total cost of conducting the 24HR using CAPI and using PAPI. In addition to calculating the total cost of conducting the 24HR using each modality, we also disaggregated costs based on total time (person days), total time costs (person days valued at personnel wage rates/salaries) and total non-time monetary costs.

In addition to the 'base model' in which the cost of conducting the 24HR reflected the conditions under which the validation study was conducted, we modelled several cost scenarios to reflect anticipated development of the FMDB. The FMDB, which houses dietary reference data for INDDX24, is just beginning to be populated with data. However, as future INDDX24 users share their dietary reference data (e.g. food composition tables, standard recipes, portion conversion factors), other users of the platform may be able to borrow some of these data directly from the FMDB, thereby reducing the costs associated with the preparation of these data for users of INDDX24. We modelled three alternative scenarios in which we assumed that 25, 50 or 75% of dietary reference data were borrowed directly from the FMDB for the INDDX24 modality.

To estimate the cost-effectiveness of using the CAPI modality relative to PAPI, we set these estimates of cost alongside measures of the accuracy of each modality compared with the WFR benchmark. We based our primary measures of cost-effectiveness on three accuracy measures: (1) average percentage accuracy in estimating the number of food items consumed (item count), (2) average percentage accuracy in estimating total gram amount of food intake, and (3) a composite measure of the average percentage accuracy in estimating each of ten nutrients of interest (energy, fat, protein, carbohydrate, fibre, vitamin A, vitamin C, Ca, Fe, and Zn), in the spirit of mean nutrient adequacy described in Hatløy, Torheim, and Oshaug⁽¹³⁾ and applied for validation of dietary metrics⁽⁴⁹⁻⁵¹⁾.

The average percentage accuracy in estimating the number of food items and gram amount was calculated based on the average percentage error relative to the WFR, that is, we calculated the average percentage error as the absolute value of one minus the ratio of the average estimate based on each 24HR modality to the average estimate based on the WFR, and then average percentage accuracy was calculated as 100 minus the average percentage error. For example, if the average gram amount among the study sample in the PAPI arm was 3500 based in the 24HR and 4000 based on the WFR, then the average PAPI percentage error was calculated as $\left| \left(\frac{1-3500}{4000} \right) \times 100 \right| = 12.5\%$, and the average PAPI percentage accuracy, or effectiveness, was $100 - 12.5 = 87.5\%$.

For the composite measure of nutrient intake, average percentage accuracy was calculated as the overall average of the average percentage accuracy for each nutrient as described for item count and gram amount. In addition to these primary measures of cost-effectiveness, we used the same methodology to calculate nutrient-specific effectiveness for each of the ten nutrients of interest. The average percentage accuracy of each

outcome based on the validation study was assumed to be the same for each of the alternative scenarios.

We calculated cost-effectiveness, or cost per average percentage points of accuracy, based on total cost and each measure of disaggregated cost (time measured in person-days, time cost and non-time (monetary) cost). For the scenarios in which dietary reference data were assumed to be borrowed from the FMDB, effectiveness was based on the average percentage accuracy of each outcome from the validation study.

Results

The randomisation was well balanced, with no statistically significant differences in participants' socio-economic and demographic characteristics between the two arms of the validation study (Table 1). A total of 231 women (116 INDDX24 and 115 PAPI) were enrolled in the validation study in Burkina Faso. Due to implausible values and a matching issue between WFR and INDDX24 because of errors in respondent ID, five INDDX24 and five PAPI respondents were excluded from the analysis (Fig. 1). The mean age of participants in the full sample was 29.7 (SD 8.5) years. Most participants (83.5%) did not have any formal education, and the rest of them had primary or secondary level education. The average household size was 11.7 (SD 5.8). About one-third of participants (35.2%) were breast-feeding, and 13.4% were pregnant at the time of the study. Almost all the participants' households (95%) owned agricultural land. Borehole was the primary source of drinking water for all the participants. Slightly less than two-thirds of participants (64.7%) reported preparing food for others all or most of the time. The main household cooking fuel was wood, which was reported by 96.8% of participants. Participants reported using mainly three types of lighting sources: lamp with rechargeable battery (37.5%), flashlight with non-rechargeable battery (35.6%) and solar energy (16.8%).

Differences in estimates by INDDX24 v. pen-and-paper interview 24-h dietary recall modalities compared with the weighed food record benchmark: nutrient intakes

In general, INDDX24 and PAPI were equivalent to WFR based on two one-sided paired *t* equivalence test for the item count, gram amount intake and all macronutrient intakes except fat intake for PAPI, as indicated by the significant equivalence test *P*-value (Table 2). Regarding the micronutrient and fibre intakes, neither of the 24HR modalities was equivalent to WFR except for vitamin A and Zn intakes. INDDX24 and PAPI were equivalent to WFR for vitamin A intake, while only INDDX24 was equivalent for Zn intake. For the difference-in-differences comparisons, there were very few instances where INDDX24 and PAPI differed significantly from one another in approximating the WFR estimate except for energy and carbohydrates. For energy intake, INDDX24 underestimated by 157.8 kcal while PAPI overestimated by 137.8 kcal when compared with WFR. Both INDDX24 and PAPI underestimated carbohydrate intake, but INDDX24 underestimated by 73.9 g, while PAPI underestimated by 10 g (Table 2). Thus, for these nutrients (energy,

Table 1. Socioeconomic and demographic characteristics of participants in INDDX24 validation study in Burkina Faso

	PAPI* n = 115	INDDX24 n = 116	P-value†
Average age (mean (SD)) years	29.8 (8.1)	29.5 (8.8)	0.776
Number of people in household (mean (SD))	12.3 (5.8)	11.1 (5.8)	0.127
Pregnant respondent (%)	13.9	13.0	0.361
Breastfeeding respondent (%)	35.2	39.4	0.778
Education level (%)			0.584
Primary	11.1	15.5	
Secondary school	2.8	3.6	
High school	–	–	
University	–	–	
Professional	–	–	
No education	86.1	80.9	
Agricultural land possession by household (%)	95.4	94.5	0.919
Source of drinking water (%)			1
borehole	100	100	
Weekly frequency of food preparation by respondent (%)			0.999
All or most of the time	64.8	64.5	
Sometimes	34.2	34.5	
Rarely	1.0	1.0	
Household cooking fuel (%)			0.228
Wood	97.2	96.4	
Straw	1.0	1.8	
Gas	0.0	1.8	
Other	1.8	0.0	
Household lightening source (%)			0.771
Lamp with rechargeable battery	37.0	34.3	
Flashlight with non-rechargeable battery	37.0	38.0	
Solar energy	15.7	17.6	
Solar-rechargeable flashlight	9.3	8.3	
Other source‡	1.0	1.8	
Household floor material (%)			0.420
Banco	13.9	9.1	
Cement or concrete	59.3	58.2	
Dirt sand	26.8	32.7	
Household roof material (%)			0.553
Metal sheets	81.5	75.5	
Mat or straw	16.7	21.8	
Dirt sand or banco	1.8	2.7	
Household wall material (%)			0.522
Earth-made brick or banco	75.0	74.5	
Half banco and half cement	18.5	20.0	
Cement	3.7	4.5	
Other wall material§	2.8	1.0	
Household toilet facility (%)			0.553
Pit latrine with slab	22.6	30.0	
Pit latrine without slab	17.9	16.4	
Bush or field	46.2	44.5	
Other toilet facility**	3.9	1.8	
No toilet facility	9.4	7.3	

*PAPI = Pen and Paper Interview

†P-value from t-test (binary variables) or chi-2 test (categorical variables) comparison between and PAPI arms

‡Other lightening sources included generator, woodfire and flashlight from mobile phone

§Other wall materials included hides/tarpaulin and mat/straw

**Other toilet facilities included ECOSAN latrine, manually poured flush toilet, and simple SanPlat latrine

carbohydrates), INDDX24 and PAPI did not yield comparable results, though both were equivalent to the WFR at the 10 % bound level.

For the individual-level comparison, the proportion of INDDX24 and PAPI respondents with energy intake within 10 % of their WFR estimates was, respectively, 13 and 21 %; 29 and 43 %, respectively, had energy intakes within the 20 % error range (Table 3; see also Bland–Altman plots in online Supplementary Fig. S1–S4). The proportions of respondents with micronutrient intakes within 10 % of their WFR were 20 % or less

for all nutrients, with INDDX24 having slightly lower percentages within ± 10 and ± 20 percentage points. Both modalities showed a higher percentage of respondents within those error ranges for macronutrients (fat, protein, carbohydrate) than for micronutrients. For the micronutrient intakes, estimates falling within the 10 % error range were under 10 %, and those falling within the 20 % error range were under 20 %, with no consistent difference between modalities. Accuracy of the estimates was even lower for Ca, where less than 5 % of PAPI and INDDX24 respondents had intakes within 10 % of their WFR

Table 2. Comparison of mean differences of INDEX24 and PAPI to WFR in estimated nutrient consumption for the participants in INDEX24 validation study in Burkina Faso (Mean values and standard deviations)*

	Arm 1: PAPI (n 110)							Equivalence test <i>P</i>	Arm 2: INDEX24 (n 111)				WFR- INDEX24 difference		Equivalence test <i>P</i>	Diff-in-diff test (<i>P</i>)
	WFR†		PAPI		WFR-PAPI difference		WFR mean		INDEX24							
	Mean	SD	Mean	SD	Mean	SD	Mean		SD	Mean	SD	Mean	SD			
Item count	31.1	9.2	23.5	5.4	7.6	7.1	< 0.0001	32.7	10.9	23.9	5.9	8.8	9.2	0.0072	0.4021	
Gram amount	4673.5	1324.7	3395.5	1338.6	1278.0	1395.4	< 0.0001	4680.9	1367.3	3360.2	1305.0	1320.7	1495.4	< 0.0001	0.9100	
Energy (kcal)	2446.1	906.0	2583.9	1021.8	-137.8	1024.4	< 0.0001	2425.8	956.8	2267.9	1068.1	157.8	1042.5	< 0.0001	0.0305	
Fat (g)	47.6	36.0	56.2	39.0	-8.6	39.2	0.0548	45.2	34.7	51.1	38.5	-5.9	37.6	0.0008	0.2688	
Protein (g)	78.3	33.2	93.2	41.6	-14.9	44.8	< 0.0001	75.9	34.7	82.8	54.0	-6.9	51.0	< 0.0001	0.0616	
Carbohydrates (g)	396.3	144.6	386.3	162.6	10.0	167.0	< 0.0001	396.0	147.6	322.1	145.6	73.9	156.7	< 0.0001	0.0080	
Total fibre (g)	39.0	31.4	74.1	43.0	-35.1	43.2	1	40.8	41.2	66.7	61.4	-25.9	63.6	0.9976	0.1068	
Vitamin A (mcg RAE)	298.2	320.8	471.8	588.4	-173.6	587.9	0.0123	260.0	208.4	434.6	445.6	-174.7	419.7	0.0083	0.8016	
Vitamin C (mg)	62.4	316.2	48.6	45.2	13.8	320.7	0.997	23.3	21.3	46.5	50.8	-23.3	48.4	0.9637	0.3442	
Ca (mg)	641.8	587.3	1430.2	1061.1	-788.3	1034.1	0.9342	698.4	665.6	1198.1	1301.9	-499.8	1344.3	0.0545	0.0205	
Fe (mg)	21.7	15.4	62.6	59.6	-41.0	58.8	1	22.4	16.6	46.6	37.0	-24.2	36.8	1	0.0199	
Zn (mg)	8.6	5.6	15.8	6.9	-7.2	7.4	0.6382	8.4	6.3	13.6	8.8	-5.2	8.8	0.0398	0.3040	

* Note the WFR is the minuend, and therefore a negative difference between WFR and 24HR modalities indicates an overestimate, and a positive number indicates an underestimate. Equivalence test *P*-value is from the paired two one-sided *t* test (TOST) reported using the natural log geometric mean and 10 % bound. A significant *P* < 0.05 for the equivalence test indicates that the WFR is equivalent to the 24HR at 10 % bound; non-significance means the two are statistically significantly different. The difference-in-difference *P*-value is from calculating the difference between each 24HR modality and the WFR and then calculating the overall difference of the differences. A random intercept mixed regression model was used for the difference-in-difference comparison. Statistical significance *P* < 0.05 indicates that the ININDEX24 and PAPI differed from each other with respect to the accuracy of their estimates. FCT completeness for each nutrient based on the foods reported in the WFR: energy (100 %), fat (100 %), protein (100 %), carbohydrates (97.5 %), fibre (92.1 %), vitamin A RAE (98.3 %), vitamin C (84.2 %), Ca (100 %), Fe (99.2 %) and Zn (99.6 %). The distribution of missing FCT values for each nutrient occurred equally across the ININDEX24 and PAPI modalities.

† WFR, weighed food record



Table 3. Percentage of respondent falling within ranges of percentage error in estimating energy and nutrient intakes with INDDDEX24 and PAPI 24HR modalities compared with WFR*

Item	Arm 1: PAPI (n 110)					Arm 2: INDDDEX24 (n 111)						
	0–±10 %	± 10.1–20 %	± 20.1–30 %	± 30.1–40 %	± 40.1–50 %	> ±50 %	0–±10 %	± 10.1–20 %	± 20.1–30 %	± 30.1–40 %	± 40.1–50 %	> ±50 %
Item count	20.9	23.6	20.9	22.7	8.2	3.6	10.8	25.2	26.1	25.2	8.1	4.5
Gram amount	20.0	18.2	12.7	14.5	12.7	21.8	14.4	17.1	15.3	14.4	17.1	21.6
Energy (kcal)	20.9	21.8	10.9	11.8	7.3	27.3	12.6	16.2	14.4	24.3	11.7	20.7
Fat (g)	14.5	15.5	8.2	10.9	11.8	39.1	8.1	9.0	17.1	9.9	6.3	49.5
Protein (g)	17.3	13.6	12.7	8.2	10.9	37.3	10.8	14.4	9.0	12.6	10.8	42.3
Carbohydrates (g)	17.3	13.6	20.0	16.4	12.7	20.0	17.1	11.7	15.3	16.2	17.1	22.5
Total Fibre (g)	5.5	4.5	4.5	6.4	7.3	71.8	9.0	7.2	7.2	8.1	5.4	63.1
Vitamin A (mcg RAE)	4.5	8.2	3.6	9.1	2.7	71.8	7.2	5.4	7.2	7.2	6.9	66.0
Vitamin C (mg)	8.2	7.3	4.5	4.5	6.4	69.1	0.0	4.5	3.6	7.2	7.2	75.7
Ca (mg)	4.5	10.0	3.6	5.5	4.5	71.8	4.5	6.3	8.1	5.4	6.3	69.4
Fe (mg)	2.7	2.7	5.5	2.7	8.2	78.2	7.2	9.9	0.9	3.6	7.2	71.2
Zn (mg)	6.4	8.2	6.4	3.6	6.4	69.1	5.4	12.6	4.5	3.6	9.0	64.0

* FCT completeness for each nutrient based on the foods reported in the WFR: energy (100 %), fat (100 %), protein (100 %), carbohydrates (97.5 %), fibre (92.1 %), vitamin A RAE (98.3 %), vitamin C (84.2 %), Ca (100 %), Fe (99.2 %) and Zn (99.6 %). The distribution of missing FCT values for each nutrient occurred equally across the INDDDEX24 and PAPI modalities.

estimates; none of the INDDDEX24 respondents had vitamin C intake within 10 % of the WFR estimate. Both PAPI and INDDDEX24 overestimated intakes far more frequently than they underestimated them at the individual level based on an error range of ± 50 %. This was the case for both macronutrients and micronutrients (online Supplementary Tables S3 and S4).

Differences in estimates by INDDDEX24 v. pen-and-paper interview 24-h dietary recall modalities compared with the weighed food record benchmark: food group intakes

Table 4 shows the results from the INDDDEX24 and PAPI compared with the WFR for the contribution of each food group to total energetic consumption. The largest contribution of energy comes from cereals: about 77 % according to the WFR and 65–66 % for either 24HR modality. The 24HR underestimate was statistically significant for both modalities and comparable between them. The second largest contributor of energy, fats and oils, showed no significant difference between the WFR and either modality, with estimated contributions between 7.1 and 9.2 % of total energy content. The average differences in the contribution of energy from the rest of the food groups were around or less than 5 %. The differences between the WFR and either modality were statistically significant only for vegetables, beverages, spices and condiments, and fish and shellfish. The contribution of beverages and spices/condiments was negligible for WFR and both 24HR modalities; the contribution of vegetables and fish/shellfish was negligible for the WFR but meaningful (around 4 % for fish/shellfish and 6–8 % for vegetables) for 24HR modalities, not different by modality. Overestimates by both 24HR modalities were far more frequent than underestimates (except for cereals), and the two modalities were consistent with each other in overestimating v. underestimating the contributions except for the roots and tubers group, which was consumed by only around 5 % of respondents in either study arm. Because small numbers render significance tests unreliable, P-values for the differences between the WFR and each 24HR are reported only for those food groups consumed by at least 10 % of respondents.

Differences in cost and cost-effectiveness estimates by INDDDEX24 v. pen-and-paper interview 24-h dietary recall modalities compared with the weighed food record benchmark: food group intakes

The total economic costs of using the INDDDEX24 and the PAPI modalities to conduct the 24HR are presented in Table 5. Reflected in the base model, the cost of using INDDDEX24 was slightly lower (by \$1360) than the cost of using PAPI when no dietary reference data were borrowed from the FMDB. In all four scenarios, while the cost of non-personnel inputs (e.g. supplies, equipment, transport, facilities) was higher for INDDDEX24 than PAPI, conducting the 24HR using PAPI was more human-capital intensive, and the higher personnel costs associated with PAPI meant the total cost of INDDDEX24 was lower (Table 5).

The average percentage accuracy and cost-effectiveness of each modality based on the three primary measures of effectiveness (item count, gram amount, and a composite measure of nutrient intake) are presented in Table 6. PAPI was slightly more

Table 4. Median percentage of energy intake from major FAO/WHO GIFT food groups (Median values and 25th, 75th percentiles)

	Arm 1: PAPI (n 110)							Arm 2: INDDX24 (n 111)						
	Respondent consuming (%)	WFR median	WFR 25th, 75th	PAPI median	PAPI 25th, 75th	Median difference	WRS* test P	Respondent consuming (%)	WFR median	WFR 25th, 75th	INDDX24 median	INDDX24 25th, 75th	Median difference	WRS test P
Cereals	100	77.4	61.6, 87.1	66.4	55.3, 77.2	11	0.0001	100	78.8	61.1, 87.7	65	53.6, 78.1	13.8	0.0001
Pulses, seeds and nuts	100	2.8	0.7, 13.8	5.9	3.3, 11.1	-3.1	0.2898	99.1	1.8	0.6, 10.2	4.7	1.3, 8.3	-3.9	0.0549
Vegetables	100	2.6	1.5, 4.8	8.1	3.4, 13.2	-5.5	0.0001	100	2.7	1.5, 5.7	6.2	2.4, 13.3	-3.5	0.0006
Beverages	100	0.6	0.3, 1.4	0.0	0.0, 0.0	0.6	0.0001	100	0.9	0.3, 2.1	0	0, 0	0.9	0.0001
Spices and condiments	100	0.2	0.1, 0.4	0.6	0.3, 1.2	-0.4	0.0001	100	0.2	0.1, 0.5	0.7	0.3, 1.0	-0.5	0.0001
Fish and shellfish	91.8	0.6	0.1, 2.4	4.5	1.4, 7.0	-3.9	0.0001	83.8	0.8	0.2, 2.7	4.1	1.9, 8.4	-3.3	0.0001
Fats and oils	70.9	8.2	2.5, 14.6	9.2	0.0, 18.9	-1	0.2256	70.3	7.1	2.5, 10.1	8.9	0, 18.6	-1.8	0.9088
Fruits	26.4	2.4	1.3, 4.5	3.4	0.7, 5.2	-1	0.4270	25.2	3.2	1.4, 5.2	4.8	2.3, 7.3	-1.6	0.0872
Sweets and sugars	8.2	1.6	0.5, 2.1	2.1	0.0, 3.9	-0.6	0.8203	16.2	2.3	0, 10.5	3.5	2.4, 6.0	-1.2	0.8145
Roots and tubers	4.5	5.3	2.9, 13.0	2.1	0.0, 11.8	3.2	-	5.4	3.9	1.9, 6.8	7.1	4.6, 8.8	-3.2	-
Meat	3.6	6.0	1.9, 13.5	4.2	1.2, 6.1	1.8	-	6.3	3.6	1.5, 6.2	4.2	1.9, 5.9	-0.6	-
Milk and milk products	0.9	1.7	1.7, 1.7	1.9	1.9, 1.9	-0.2	-	7.2	0	0, 0	3	1.6, 4.5	-3	-
Food for particular nutritional use	0.9	36.5	36.5, 36.5	0.0	0.0, 0.0	0.9	-	0.9	4.4	4.4, 4.4	0	0, 0	4.4	-

* WRS = Wilcoxon rank sum test is a nonparametric test that compares two paired groups. Bolded WRS P-value indicates a statistically significant difference between the WFR and the 24HR at or below, $P = 0.05$. Note the WFR is the minuend and therefore a negative difference between WFR and 24HR modalities indicates an overestimate and a positive number indicates an underestimate. Significance of the difference between WFR and 24HR is not shown for food groups consumed by less than 10 % of respondents. These are roots and tubers (5.4 % INDDX24 and 4.5 % PAPI), meat (6.3% INDDX24 and 3.6% PAPI), milk and milk products (7.2 % INDDX24 and 1 % PAPI) and food for particular nutritional use (0.9 % INDDX24 and 0.9 % PAPI).

Table 5. Economic cost of conducting a 24HR* using INDDEx24 and PAPI† under alternative scenarios

Scenario	Modality	Total time	Total time cost	Total non-time cost	Total cost
		(person days)	(2019 USD)	(2019 USD)	(2019 USD)
Base model	INDDEx24	463	50,124	27,981	78,105
	PAPI	526	62,162	17,303	79,465
	Difference‡	-63	-12,038	10,678	-1,360
Borrow dietary inputs from FMDB§ (25%)**	INDDEx24	437	47,610	27,175	74,785
	PAPI	526	62,162	17,303	79,465
	Difference‡	-89	-14,552	9,872	-4,680
Borrow dietary inputs from FMDB (50%)**	INDDEx24	412	45,096	26,369	71,466
	PAPI	526	62,162	17,303	79,465
	Difference	-114	-17,065	9,066	-7,999
Borrow dietary inputs from FMDB (75%)**	INDDEx24	386	42,583	25,564	68,146
	PAPI	526	62,162	17,303	79,465
	Difference‡	-140	-19,579	8,260	-11,319

* 24HR, 24-hour dietary recall; FMDB, Food Matters Database.

† PAPI=Pen and Paper Interview

‡ The difference was calculated as INDDEx24-PAPI.

§ FMDB=Food Matters Data Base

** These scenarios assumed that, for the INDDEx24 modality, the specified percentage of dietary reference data (e.g., food composition data, standard recipes, conversion factors) were borrowed from the FMDB.

accurate than INDDEx24 (~1–2 percentage points) for gram amount and item count, while the average percentage accuracy of INDDEx24 based on the composite nutrient intake indicator was ~9 percentage points more accurate than PAPI. Given the small differences in cost of the two modalities for the base model, PAPI was more cost-effective than INDDEx24 in terms of cost per unit of accuracy in estimating item count, while INDDEx24 was more cost-effective for gram amount and the composite indicator of nutrient intake.

For the modelled scenarios in which 25–75 % of dietary reference data were assumed borrowed from the FMDB for the INDDEx24 modality, the cost of INDDEx24 further decreased relative to PAPI, based on the assumption that the dietary reference data are integrated into the INDDEx24 platform, while the preparation and use of dietary reference data for PAPI would still require implementers to seek out the information and incorporate it into their analysis. Thus, as the percentage of dietary reference data borrowed from the FMBD for INDDEx24 increases, so does the divergence between it and PAPI, since the cost of preparing dietary reference data for PAPI is unchanged. Therefore, INDDEx24 was incrementally more cost-effective than PAPI based on each of the three primary measures of accuracy. Each of these estimates of cost-effectiveness based on time (measured in person days), time cost and non-time monetary cost is available in Table 6.

Discussion

We hypothesised that the INDDEx24 and PAPI approaches would be comparable in accuracy, and that the time savings resulting from the INDDEx24 approach would mean that cost-effectiveness would favour the CAPI approach, and this hypothesis was largely supported the study.

Beyond the current study and a parallel study in Vietnam, we are unaware of studies that compared the accuracy of a CAPI and PAPI approach using a reference benchmark of accuracy such as

the WFR. Some studies have assessed accuracy of one or the other modality using WFR or biomarkers^(52,53). Numerous validation studies have reported results similar to the present study – namely, that intakes of many nutrients measured by 24HR were not statistically significantly different from WFR, but that some nutrients did show differences and that some modifications may be needed to reduce error^(54–57). But these did not compare one 24HR modality with another, and few were conducted in LMIC. Comparisons of CAPI with PAPI have been done to assess acceptability and feasibility of the methods⁽⁵⁸⁾, and differences in measured outcomes⁽⁵⁹⁾, but without a benchmark to compare accuracy of the two methods.

A possible limitation of the study is that the comparability of the two modalities may be due in part to the fact that all interviewers conducted both INDDEx24 and PAPI interviews; learning undoubtedly occurred that affected (most likely improved) the quality of the interviews in both study arms. For example, the INDDEx24 platform provided tags and probes as part of the app; interviewers may have learned to ask these questions in the PAPI even when they did not appear in the paper form. Developing the INDDEx24 platform required the systematic collection of dietary reference data prior to survey implementation. These data facilitated the development of an analysable database for both the INDDEx24 and PAPI approaches; such data might not be available in advance in the case of a PAPI as commonly implemented. In particular, the PAPI interviews made use of the repository of standard recipes and followed the same approach to recording NSR using the standard recipes as a basis when possible; this option is not typically available in PAPI dietary surveys.

Another possible limitation is that using the WFR as a benchmark meant that all respondents to the 24HR in the validation study had already been exposed to dietary measurement during the WFR the previous day. This exposure might have increased their awareness of their food consumption, affecting their recall the following day. Previous studies have found that the accuracy of 24HR may be improved by asking respondents to record or otherwise pay attention to their consumption in the days prior



Table 6. Cost-effectiveness of conducting a 24HR using INDDX24 and PAPI in Burkina Faso

Scenario	Outcome	Modality	Average percentage accuracy	Cost per average percentage accuracy on the basis of:			
				Total time (person days)	Total time cost (2019 USD)	Total non-time cost (2019 USD)	Total cost (2019 USD)
Base model	Gram amount	INDDX24	72	6	698	390	1088
		PAPI	73	7	856	238	1094
		Difference*	-1	-1	-157	152	-6
	Item count	INDDX24	73	6	686	383	1069
		PAPI	76	7	823	229	1052
		Difference	-2	-1	-137	154	17
	Composite nutrient intake†	INDDX24	48	10	1042	582	1624
		PAPI	39	14	1596	444	2041
		Difference	9	-4	-554	137	-417
Borrow dietary inputs from FMDB (25%)‡	Gram amount	INDDX24	72	6	663	379	1042
		PAPI	73	7	856	238	1094
		Difference	-1	-1	-192	140	-52
	Item count	INDDX24	73	6	651	372	1023
		PAPI	76	7	823	229	1052
		Difference	-2	-1	-171	143	-28
	Composite nutrient intake	INDDX24	48	9	990	565	1555
		PAPI	39	14	1596	444	2041
		Difference	9	-4	-607	121	-486
Borrow dietary inputs from FMDB (50%)‡	Gram amount	INDDX24	72	6	628	367	996
		PAPI	73	7	856	238	1094
		Difference	-1	-2	-227	129	-98
	Item count	INDDX24	73	6	617	361	978
		PAPI	76	7	823	229	1052
		Difference	-2	-1	-206	132	-74
	Composite nutrient intake	INDDX24	48	9	938	548	1486
		PAPI	39	14	1596	444	2041
		Difference	9	-5	-659	104	-555
Borrow dietary inputs from FMDB (75%)‡	Gram amount	INDDX24	72	5	593	356	949
		PAPI	73	7	856	238	1094
		Difference	-1	-2	-262	118	-144
	Item count	INDDX24	73	5	583	350	932
		PAPI	76	7	823	229	1052
		Difference	-2	-2	-240	121	-119
	Composite nutrient intake	INDDX24	48	8	885	531	1417
		PAPI	39	14	1596	444	2041
		Difference	9	-5	-711	87	-624

* The difference was calculated as INDDX24 minus PAPI.

† Composite nutrient intake is calculated as the overall average of the average percentage accuracy of PAPI (or of INDDX24) in estimating intake for energy, fat, protein, carbohydrate, fibre, vitamin A, vitamin C, Ca, Fe and Zn relative to intake based on the weighed food record.

‡ These scenarios assume the specified percentage of dietary data inputs (e.g. food composition data, standard recipes, conversion factors) can be borrowed from existing dietary reference data in the INDDX24 database.

to an anticipated dietary survey^(60,61), so this ‘priming’ effect may even be an advantage. In the present study, our purpose was not to obtain information about usual or representative diets of the population, but specifically compare the accuracy with which diets were recalled as captured by INDDX24 and PAPI. While administration of the WFR might have improved the accuracy of the recall, the comparison would not have been possible without administering the WFR as a point of reference, since comparisons were made on detailed consumption information by nutrient and food group.

A final limitation is that the study was implemented in a single geographic area of Burkina Faso, and only to rural women 18–49 years old; the conclusions drawn from the comparison thus may not be applicable to different age or sex groups or to contexts

with more varied and complex (e.g. urban) dietary sources. A similar INDDX24 validation study was conducted in Vietnam, a different geographic context and a country with a more diverse diet, using the same procedures and protocols⁽⁶²⁾. This second study reached similar conclusions in that both found that INDDX24 and PAPI performed comparably to each other in terms of accuracy with respect to the WFR. In that study, though, 24HR modalities were equivalent to WFR for more micronutrients, and a larger proportion of respondents fell within ±20% of the WFR intakes than in the present study.

In addition to accuracy and cost-effectiveness, there are other benefits to using the INDDX24 platform. It is standardised, so that results from any study using this platform may be comparable with results from other countries or (in the future) other

time periods. The INDDEX24 platform is designed to be linked to dietary reference data available in the FMDB. As INDDEX24 becomes more widely used, dietary reference data (e.g. food composition data, portion conversion, standard recipes) will increasingly populate the FMDB, further reducing the initial costs of preparing for a dietary survey and enabling the smooth linkage from survey responses to food composition and other data needed to process the survey results. The intention is to build up the contents of the FMDB and make it widely accessible, allowing faster preparation, survey implementation, analysis and production of usable information. As more dietary reference data can be borrowed from the FMDB, cost-effectiveness becomes even more favourable to INDDEX24. This is important not only to reduce the initial cost of preparation but also to reduce significantly the lag time between completing data collection and producing usable results while they are still relevant. By streamlining the process of conducting individual dietary intake surveys, the intention is that the use of the INDDEX24 platform will contribute to greater availability and use of detailed dietary data to inform effective food and nutrition policies.

Conclusion

The goal of the study was to compare the INDDEX24 platform, a CAPI approach to conducting individual dietary recall surveys, with the commonly implemented PAPI approach. Results support the conclusion that INDDEX24 is not less accurate than PAPI and that it has distinct advantages in terms of both cost and the timeliness with which survey results can be obtained once data collection is complete. In the future, as the FMDB is populated, the time and cost required for survey preparation using INDDEX24 may be further reduced. Inaccuracies in the estimation of micronutrient consumption and individual-level consumption were comparable between the two modalities; further exploration into ways of increasing the accuracy of recall methods is needed. While innovative technologies for measuring food consumption are under exploration, the 24HR is likely the most feasible method for collecting individual dietary information on populations in LMIC, and the INDDEX24 platform demonstrates advantages that could facilitate such collection and thus encourage the use of dietary surveys in the formation of programmes and policies related to food and nutrition.

Acknowledgements

We are very grateful to the INSD and IRSS staff and interviewers who worked tirelessly to support this study and to respondents who generously shared their time, household space and dietary information for full-day observations and recalls. We also appreciate and Julia Matteson's work in the preparation of the manuscript.

The INDDEX Project is implemented by the Tufts University, Gerald J. and Dorothy R. Friedman School of Nutrition Science and Policy with funding from the Bill & Melinda Gates Foundation.

B. R., J. C., J. W. S. and W. B. designed the research; B. R., J. C., J. W. S., W. B., D. A. C., and S. W. provided essential materials (developed the INDDEX24 platform); J. W. S. and S. W.

conducted the research; P. B., K. P. A. and J. W. S. analysed data; B. R., J. W. S. and K. A. wrote the paper; B. R. had primary responsibility for final content. All authors read and approved the final manuscript.

The authors declare no conflict of interest.

Supplementary material

For supplementary material referred to in this article, please visit <https://doi.org/10.1017/S0007114521004700>

References

1. von Grebmer K, Saltzman A, Birol E, *et al.* (2014) *2014 Global Hunger Index: the Challenge of Hidden Hunger*. Washington, DC: Welthungerhilfe, International Food Policy Research Institute, and Concern Worldwide.
2. UNICEF (2018) UNICEF Strategic Plan 2018–2021 – Executive Summary. –https://www.unicef.org/media/48126/file/UNICEF_Strategic_Plan_2018-2021-ENG.pdf (accessed August 2020).
3. USAID (2014) Multi-Sectoral Nutrition Strategy 2014–2025. –https://www.usaid.gov/sites/default/files/documents/1867/USAID_Nutrition_Strategy_5-09_508.pdf (accessed August 2020).
4. Andrews-Trevino JY, Webb P, Shively G, *et al.* (2019) Relatively low maternal aflatoxin exposure is associated with small-for-gestational-age but not with other birth outcomes in a prospective birth cohort study of Nepalese infants. *J Nutr* **149**, 1818–1825.
5. Andrews-Trevino JY, Webb P, Baral K, *et al.* (2020) Early life exposure to mycotoxins and child linear growth in nepal: methods and design of a prospective birth cohort study. *J Food Secur* **8**, 1–10.
6. Hoffmann V, Moser C & Saak A (2019) Food safety in low and middle-income countries: the evidence through an economic lens. *World Dev* **123**, 104611.
7. Coates JC, Colaiezzi BA, Bell W, *et al.* (2017) Overcoming dietary assessment challenges in low-income countries: technological solutions proposed by the International Dietary Data Expansion (INDDEx) Project. *Nutrients* **9**, 289.
8. HLPE (2017) *Nutrition and Food Systems. A Report by the High-Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security*. Rome, Italy: Committee on World Food Security.
9. IFPRI (2019) Accelerating the End of Hunger and Malnutrition: a Global Event: Synopsis. Proceedings of IFPRI-FAO Bangkok Conference. Bangkok, Thailand; Washington, DC: Food and Agriculture Organization (FAO) and International Food Policy Research Institute (IFPRI). doi: [10.2499/9780896293533](https://doi.org/10.2499/9780896293533).
10. Kennedy G, Ballard T & Dop M (2011) Guidelines for Measuring Household and Individual Dietary Diversity. Rome. <http://www.fao.org/3/i1983e/i1983e.pdf> (accessed July 2020).
11. Swindale A & Bilinsky P (2006) Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide (v.2). Washington, DC. https://www.fantaproject.org/sites/default/files/resources/HDDS_v2_Sep06_0.pdf (accessed July 2020).
12. Wiesmann D, Bassett L, Benson T, *et al.* (2009) Validation of the World Food Programme's Food Consumption Score and Alternative Indicators of Household Food Security. IDEAS Working Papers from RePEc. <http://cdm15738.contentdm.oclc.org/utis/getfile/collection/p15738coll2/id/32010/filename/32011.pdf> (accessed July 2020).

13. Hatloy A, Torheim LE & Oshaug A (1998) Food variety – a good indicator of nutritional adequacy of the diet? A case study from an urban area in Mali, West Africa. *Eur J Clin Nutr* **52**, 891–898.
14. Hoddinott J & Yohannes Y (2002) Dietary Diversity as a Food Security Indicator. Dietary Diversity as Household Food Security Indicator. 136. <http://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/47966/filename/43424.pdf> (accessed July 2020).
15. Arimond M, Wiesmann D, Becquey E, *et al.* (2010) Simple food group diversity indicators predict micronutrient adequacy of women's diets in 5 diverse, resource-poor settings. *J Nutr* **140**, 2059S–2069S.
16. Ang LC, Yap F, Najwa WN, *et al.* (2019) Validity, reliability of a Food Frequency Questionnaire (FFQ) to assess dietary intake of preschool children. *Int J Environ Res Public Health* **16**, 4722.
17. Forster H, Fallaize R, Gallagher C, *et al.* (2014) Online dietary intake estimation: the Food4Me food frequency questionnaire. *J Med Internet Res* **16**, e150.
18. Verger E, Dop M & Martin-Prével Y (2017) Not all dietary diversity scores can legitimately be interpreted as proxies of diet quality. *Public Health Nutr* **20**, 2067–2068.
19. Eng JY & Moy FM (2011) Validation of a food frequency questionnaire to assess dietary cholesterol, total fat and different types of fat intakes among malay adults. *Asia Pac J Clin Nutr* **20**, 639–645.
20. Sharma S, Iwasaki M, Kunieda C, *et al.* (2009) Development of a quantitative food frequency questionnaire for assessing food, nutrient, and heterocyclic aromatic amines intake in Japanese Brazilians for a colorectal adenoma case-control study. *Int J Food Sci Nutr* **60**, 128–139.
21. Wulandari H & Prasanto H (2018) Comparison between brief food frequency questionnaire and food record to assess the energy and protein intake of hemodialysis patients at Dr. Sardjito Hospital in Indonesia. *Saudi J Kidney Dis Transpl* **29**, 145–152.
22. Zang J, Luo B, Chang S, *et al.* (2019) Validity and reliability of a food frequency questionnaire for assessing dietary intake among Shanghai residents. *Nutr J* **18**, 30–12.
23. Coates J, Patenaude BN, Rogers BL, *et al.* (2018) Intra-household nutrient inequity in rural Ethiopia. *Food Policy* **81**, 82–94.
24. Harris-Fry H, Shrestha N, Costello A, *et al.* (2017) Determinants of intra-household food allocation between adults in South Asia – a systematic review. *Int J Equity Health* **16**, 107–121.
25. Harris-Fry HA, Paudel P, Shrestha N, *et al.* (2018) Status and determinants of intra-household food allocation in rural Nepal. *Eur J Clin Nutr* **72**, 1524–1536.
26. Mwangi D, Fiedler J & Sununtnasuk C (2017) Imputing Nutrient Intake from Foods Prepared and Consumed Away from Home and Other Composite Foods. Washington, DC. <http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/131036> (accessed August 2020).
27. Pisa PT, Landais E, Margetts B, *et al.* (2018) Inventory on the dietary assessment tools available and needed in africa: a prerequisite for setting up a common methodological research infrastructure for nutritional surveillance, research, and prevention of diet-related non-communicable diseases. *Crit Rev Food Sci Nutr* **58**, 37–61.
28. Thompson FE & Subar AF (2013) Dietary assessment methodology. In *Nutrition in the Prevention and Treatment of Disease*, pp. 5–45 [AM Coulston, CJ Boushey & MG Ferruzzi, editors]. Waltham, MA: Elsevier, Inc.
29. Wrieden W, Peace H, Armstrong J, *et al.* (2003) A Short Review of Dietary Assessment Methods Used in National and Scottish Research Studies. <https://www.food.gov.uk/sites/default/files/multimedia/pdfs/scotdietassessmethods.pdf> (accessed August 2020).
30. Gibson RS, Charrondiere R & Bell W (2017) Measurement errors in dietary assessment using self-reported 24-hour recalls in low-income countries and strategies for their prevention. *Adv Nutr* **8**, 11.
31. Gibson RS, & Ferguson EL (2008) An Interactive 24-hour Recall for Assessing the Adequacy of Iron and Zinc Intakes in Developing Countries. <http://www.ifpri.org/sites/default/files/publications/tech08.pdf> (accessed July 2020).
32. Harris-Fry H, Beard B, Harrisson T, *et al.* (2018) Smartphone tool to collect repeated 24 h dietary recall data in Nepal. *Public Health Nutr* **21**, 260–272.
33. Steinfeldt L, Anand J & Murayi T (2013) Food reporting patterns in the USDA automated multiple-pass method. *Proc Food Sci* **2**, 145–156.
34. Adams K, Bell W, Somé J, *et al.* (2021) The cost of conducting a 24-hour dietary recall using INDDX24, a mobile dietary assessment platform, compared to pen-and-paper in Vietnam, Burkina Faso. *Curr Dev Nutr* **5**, 620.
35. Dickinson FM, McCauley M, Madaj B, *et al.* (2019) Using electronic tablets for data collection for healthcare service, maternal health assessments in low resource settings: lessons learnt. *BMC Health Serv Res* **19**, 1–7.
36. Leisher C (2014) A comparison of tablet-based and paper-based survey data collection in conservation projects. *Soc Sci* **3**, 264–271.
37. Zhang S, Wu Q, van Velthoven MH, *et al.* (2012) Smartphone *v.* pen-and-paper data collection of infant feeding practices in rural China. *J Med Internet Res* **14**, e119.
38. Caeyers B, Chalmers N & De Weerd J (2012) Improving consumption measurement and other survey data through CAPI: evidence from a randomized experiment. *J Dev Econ* **98**, 19–33.
39. Mukasa O, Mushi HP, Maire N, *et al.* (2017) Do surveys with paper and electronic devices differ in quality and cost? Experience from the Rufiji Health and demographic surveillance system in Tanzania. *Glob Health Action* **10**, 1387984.
40. Bell W, Coates JC, Rogers BL, *et al.* (2019) Getting the food list 'right': an approach for the development of nutrition-relevant food lists for household consumption and expenditure surveys. *Public Health Nutr* **22**, 246–256.
41. Bermudez O, Lividini K, Smits M, *et al.* (2012) Estimating micro-nutrient intakes from Household Consumption and Expenditures Surveys (HCES): an example from Bangladesh. *Food Nutr Bull* **33**, S208–S213.
42. Tucker KL, Bianchi LA, Maras J, *et al.* (1998) Adaptation of a food frequency questionnaire to assess diets of Puerto Rican and Non-Hispanic adults. *Am J Epidemiol* **148**, 507–518.
43. Bland JM & Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* **8**, 135–160.
44. Bland JM & Altman DG (2007) Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat* **17**, 571–582.
45. Batterham MJ, Van Loo C, Charlton KE, *et al.* (2016) Improved interpretation of studies comparing methods of dietary assessment: combining equivalence testing with the limits of agreement. *Br J Nutr* **115**, 1273–1280.
46. Tugault-Lafleur CN, Black JL & Barr SI (2017) A systematic review of methods to assess children's diets in the school context. *Adv Nutr* **8**, 63–79.
47. World Health Organization (2003) Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis. Geneva, Switzerland. https://www.who.int/choice/publications/p_2003_generalised_cea.pdf (accessed August 2020).
48. ILOSTAT (2018) *Country Profiles*. Burkina Faso: International Labour Organization.



49. Steyn NP, Nel J, Labadarios D, *et al.* (2014) Which dietary diversity indicator is best to assess micronutrient adequacy in children 1 to 9 years? *Nutrition* **30**, 55–60.
50. Torheim LE, Ouattara F, Diarra MM, *et al.* (2004) Nutrient adequacy and dietary diversity in rural Mali: association and determinants. *Eur J Clin Nutr* **58**, 594–604.
51. Mirmiran P, Azadbakht L, Esmailzadeh A, *et al.* (2004) Dietary diversity score in adolescents - a good indicator of the nutritional adequacy of diets: Tehran lipid and glucose study. *Asia Pac J Clin Nutr* **13**, 56–60.
52. Timon CM, van den Barg R, Blain RJ, *et al.* (2016) A review of the design and validation of web- and computer-based 24-h dietary recall tools. *Nutr Res Rev* **29**, 268–280.
53. Xue H, Yang M, Liu Y, *et al.* (2017) Relative validity of a 2-day 24-hour dietary recall compared with a 2-day weighed dietary record among adults in South China. *Nutr Diet* **74**, 298–307.
54. Arsenault JE, Moursi M, Olney DK, *et al.* (2020) Validation of 24-h dietary recall for estimating nutrient intakes, adequacy in adolescents in Burkina Faso. *Matern Child Nutr* **16**, e13014.
55. Kirkpatrick SI, Baranowski T, Subar AF, *et al.* (2019) Best practices for conducting and interpreting studies to validate self-report dietary assessment methods. *J Acad Nutr Diet* **119**, 1801–1816.
56. Lafrenière J, Lamarche B, Laramée C, *et al.* (2017) Validation of a newly automated web-based 24-hour dietary recall using fully controlled feeding studies. *BMC Nutr* **3**, 1–10.
57. Savard C, Lemieux S, Lafrenière J, *et al.* (2018) Validation of a self-administered web-based 24-hour dietary recall among pregnant women. *BMC Pregnancy Childbirth* **18**, 1–10.
58. Zang J, Song J, Wang Z, *et al.* (2015) Acceptability and feasibility of smartphone-assisted 24 h recalls in the Chinese population. *Public Health Nutr* **18**, 3272–3277.
59. Htet MK, Fahmida U, Do T, *et al.* (2019) The use of tablet-based multiple-pass 24-hour dietary recall application (MP24diet) to collect dietary intake of children under two years old in the prospective cohort study in Indonesia. *Nutrients* **11**, 2889.
60. Pries AM, Filteau S & Ferguson EL (2019) Snack food and beverage consumption and young child nutrition in low- and middle-income countries: a systematic review. *Matern Child Nutr* **15**, e12729.
61. Rollo ME, Williams RL, Burrows T, *et al.* (2016) What are they really eating? A review on new approaches to dietary intake assessment and validation. *Curr Nutr Rep* **5**, 307–314.
62. Bell W, Coates J, Bakun P, *et al.* (2021) The INDDX24 Dietary assessment platform is as accurate as pen-and-paper 24-hour dietary recall when compared to a weighed food record benchmark in Vietnam. *Curr Dev Nutr* **5**, 870.