

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
IFPRI

SHOW ME WHAT YOU EAT:

Assessing diets remotely through pictures

PROJECT NOTE







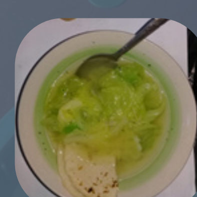
Proof of Concept – Guatemala 2019

Francisco Ceballos, Manuel A. Hernandez*

International Food Policy Research Institute (IFPRI)

May, 2021

*f.ceballos@cgiar.org; m.a.hernandez@cgiar.org

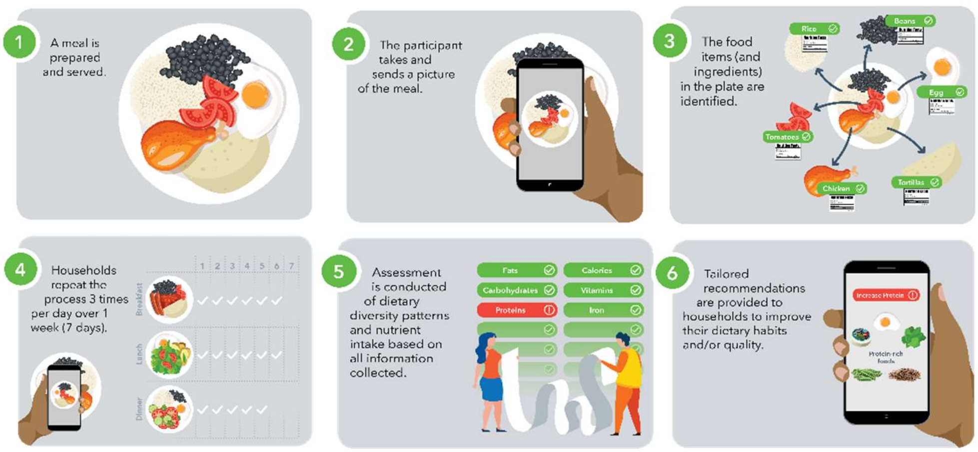




THE IDEA

Goal: Using real-time smartphone meal pictures sent by rural or urban households to better monitor and assess the quality of their diets, and provide tailored recommendations to improve them

Detailed information on household and individual dietary intake is crucial for adequate nutritional monitoring and designing interventions to improve diets. Common recall-based methods are generally time consuming, costly, and subject to non-negligible measurement errors and potential biases. In addition, the scope of information that can be obtained in a regular survey is typically limited. Detailed diaries, in turn, are effort- and time-intensive and prone to errors.

With increasing mobile penetration in both urban and rural areas, meal pictures can overcome some of these difficulties, providing real-time, detailed food intake information of individuals remotely and at a minimal cost. Moreover, pictures can be obtained over extended periods of time, beyond the standard short spans (i.e. 24-hours) in recall survey questions, with little to no data quality loss.



Such rich consumption data can help identify and better understand vulnerabilities and nutritional imbalances —including specific macronutrient or micronutrient gaps or excesses—, and open the door for low-cost, individually-tailored digital interventions to promote healthier diets. Moreover, crowdsourced data allow to identify locally available, affordable foods rich in specific nutrients consumed by similar households in the area. Interventions, in turn, can be delivered through text messages, interactive voice response (IVR), or phone calls, or videos or interactive games integrated into an app, benefitting from a two-way communication channel with individuals.

PILOT

In late 2019, we conducted household surveys among smallholder farmers in the Western Highlands of Guatemala, as part of an ongoing study in the region. In this context, we invited 286 households owning smartphones to participate in a pilot project. One individual from each household was asked to send pictures of their three main daily meals (served plates) for 5 consecutive days. A detailed printed protocol on how to take pictures was provided, requesting pictures to be taken under good lighting conditions and directly from above, to include the whole plate (in addition to any tortillas and drinks), and to include a coin for scaling purposes. Pictures were shared through Whatsapp using a specialized platform on the receiving end. The individual owning the phone was provided with a data plan to send pictures free of cost, with a second one provided once all pictures were sent, as appreciation for participating.

Household characteristics	
Dept.: Huehuetenango	29.8%
Dept.: San Marcos	11.6%
Dept.: Quiché	58.6%
Indigenous	76.2%
HH size (members)	5.9
HH head is male	79.6%
HH head age	46.2
HH head complete primary	35.9%
Agric. land size (hectares)	1.2
Finished walls	49.2%
Finished ceiling	13.3%
Finished floors	63.5%
Connected to electricity	84.5%
Connected to water	86.7%
Connected to sewer	30.4%

Participating individuals	
TOTAL	**178**
Female	**137**
9–18 y	25
19–30 y	78
31–50 y	27
>51 y	7
Pregnant	1
Lactating	29
Neither	107
Male	**41**
9–18 y	7
19–30 y	22
31–50 y	10
>51 y	2

In a second stage, we engaged local nutritionists to identify valid pictures and label the main ingredients and estimated weights in each of them.

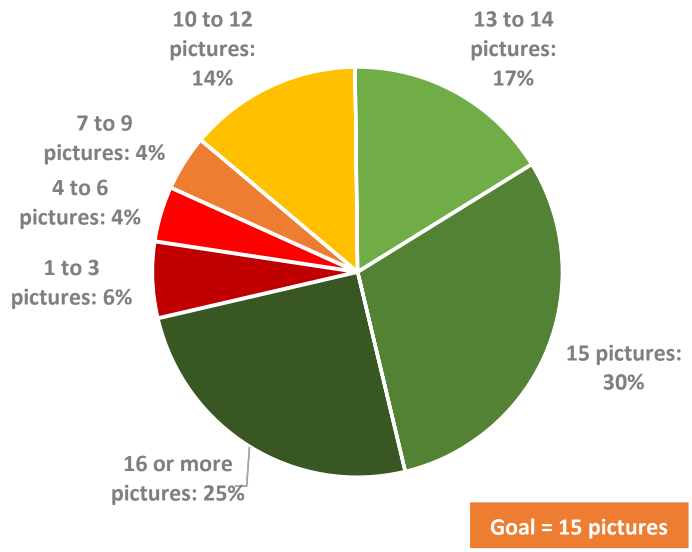
Finally, we use these labeled pictures to calculate an array of nutritional indicators, some of which would not be available using standard data collection methods. This note highlights the main lessons learnt from this proof of concept study.

PARTICIPATION, COMPLIANCE, AND PICTURE QUALITY

PARTICIPANTS ARE WILLING AND ABLE TO SEND VALID PICTURES

Out of 286 individuals invited to participate, 196 (or 69%) accepted, and 178 sent pictures, equivalent to a compliance rate above 90%. Overall, **2,394 valid pictures**, i.e., a picture of a plate with a meal, were received, or 13.4 pictures on average per participating individual.

15 pictures were requested (3 meals a day over 5 days), and more than half of the participants sent this amount or more, while 86% sent at least 10 pictures (equivalent to 3 complete days of meals). Less than 14% contributed 9 or less pictures.



HOUSEHOLD PROFILES AROUND PARTICIPATION AND COMPLIANCE

When analyzing household characteristics correlated with participation and compliance, we observe that while individuals from indigenous households are on average more willing to participate, they are less prone to send pictures, perhaps indicating less familiarity with cellphones, lower digital education, or other digital use or access gaps.

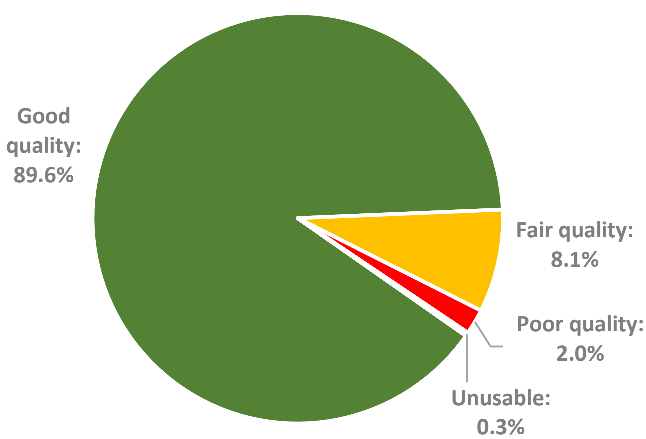
Male-headed and wealthier households (as captured by different proxies) also seem less likely to participate.

Other characteristics like age and education of the household head do not seem to matter for either participation or compliance.

	Accepted to participate in pilot	Participated and sent > 1 pictures	Participated and sent 15 or more pictures
Indigenous household	+ 9 pp.	- 14 pp.	- 24 pp.
Finished ceiling	-	- 14 pp.	-
Connected to sewer	- 10 pp.	-	-
Household size	+ 2 pp. (per member)	-	-
Household head is male	- 8 pp.	- 11 pp.	-
Household head age	-	-	-
Household head education	-	-	-
Size of agricultural land	-	+ 9 pp.	-
Mean probability	**70%**	**92%**	**47%**

Note: Marginal effects from a Probit specification. "pp" means percentage points. "-" indicates a non-significant effect, indistinguishable from zero.

PICTURES SENT ARE OF SUFFICIENT QUALITY



Labelers assessed 90% of pictures to be of good quality and only less than 3% as problematic for identifying food elements. Overall, 3 out of every 4 participating individuals sent more than 80% of pictures of good quality.

POTENTIAL ISSUES:

1. Depending on the plate layout, some elements (ingredients) are not always identifiable
2. When the meal is a soup or a stew, some ingredients may be missed
3. Some pictures correspond to a *family* meal instead of to an individual plate (very few cases)

LABELING PICTURES

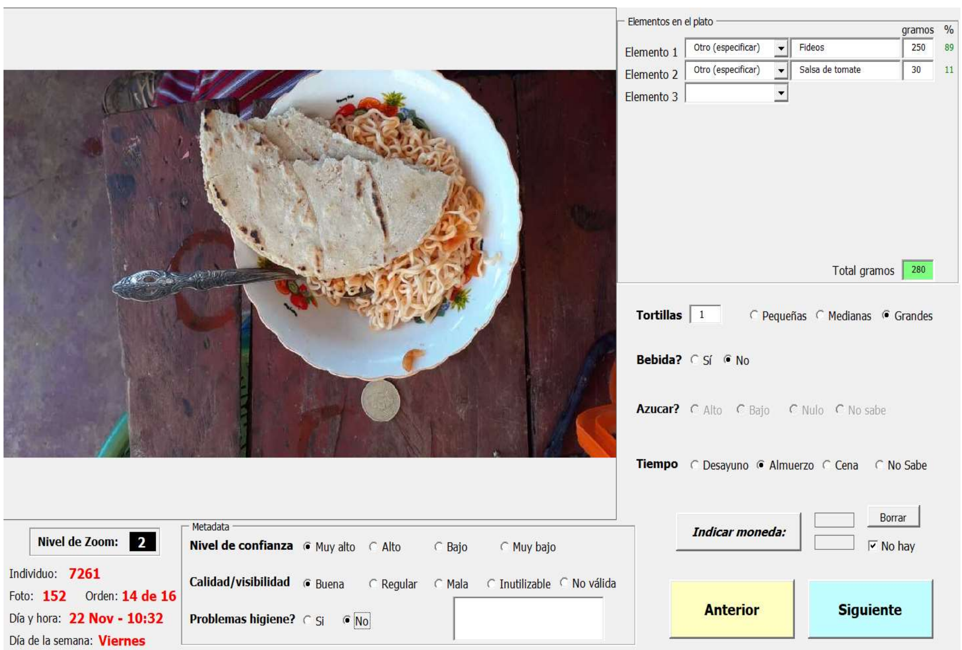
MANUAL LABELING IS FEASIBLE AND COST EFFECTIVE

While automatic labeling using machine learning is an ideal medium- to long-term goal—once sufficient labeled pictures become available—, manual labeling is a fast, reliable, and cost-effective alternative.

Four local nutritionists labeled around 700 pictures each, using a dedicated software. The median time needed *per picture* was 1 minute, with an average time of 2 minutes. Minimal variations were found across nutritionists, indicating that these times are reasonable. Even though we relied on nutritionists, more cost-effective alternatives—such as university students from health-related disciplines subject to rigorous training around labeling—may suffice for scaling.

All in all, labeling a *full day of meals* for a single individual uses 3-6 minutes of staff time. This comes in contrast to the much longer time required to conduct full survey-based consumption modules, in addition to fieldwork costs.

METHODOLOGY



We developed a dedicated Visual Basic solution to label pictures. For each picture, we asked for a full list of food elements, including an estimated weight in grams; whether a drink or tortillas were included; and other relevant metadata such as picture quality and labeler's confidence around food identification.

Each food *element* (e.g. tomato sauce) was decomposed into *ingredients* (e.g. tomato, oil, onion, salt) following standard local recipes and the estimated weights. Subsequently, the nutritional content of each ingredient was obtained from 2012 food composition tables of the Instituto de Nutrición de Centro America y Panama (INCAP), and each meal's full nutritional content derived by aggregating that of all its ingredients.

LABELED ELEMENTS CLOUD

ARE LABELS RELIABLE?

In order to assess the reliability of labels, we asked all four labelers to assess a common subset of 95 pictures. Overall, the degree of agreement around the general contents of meals was very high, with labelers agreeing about the main food ingredients in almost 98% of cases.

However, we observe important differences in the *estimated weight* of meals. Such differences can be minimized through consistent training and by providing visual cues from pre-weighted food items to be used as reference. Specific bias correction techniques can also be used if a given labeler is consistently over- or under-estimating weights.

Correlation between labelers

	Labeler pairs					
	1 vs 2	1 vs 3	1 vs 4	2 vs 3	2 vs 4	3 vs 4
Total grams in meal	0.43	0.84	0.78	0.39	0.50	0.83
% energy from Meat products	0.70	0.72	0.81	0.78	0.77	0.76
% energy from Milk, dairy, and eggs	0.85	0.92	0.88	0.83	0.82	0.84
% energy from Grains and cereals	0.76	0.82	0.75	0.77	0.74	0.71
% energy from Fruits and vegetables	0.93	0.90	0.73	0.87	0.70	0.61
% energy from Pulses, legumes, and nuts	0.78	0.79	0.71	0.81	0.72	0.68
% energy from Fats, sweets, and spices	0.08	-0.03	0.20	0.08	0.51	0.29

Despite differences in estimated food *weight*, labelers tend to coincide more over the meal *content*. This is shown by much higher pairwise correlations in the percentage of energy from different food groups (derived from the labels). An exception is the group of Fats, sweets, and spices, which is difficult to observe directly from meal pictures (discussed below).

LABELING UNCERTAINTIES

Not every ingredient in a dish can always be ascertained. In particular, we found the following sources of uncertainty:

- ▷ Distinguishing among meats (e.g. cow vs. pig)
- ▷ Distinguishing among dark green vegetables (chard, spinach, hierbamora)
- ▷ Identifying exact content in soups or stews
- ▷ Identifying filling of certain foods (e.g. tamal)

Most dishes, however, were identified with reasonable confidence given labelers' knowledge of the local context and foods. This highlights the need to work with local labelers and with protocols to handle uncertainty around item identification (e.g. allowing labelers to indicate different plausible ingredients when unsure).

POTENTIAL LIMITATIONS

Certainly, pictures of meals cannot completely capture an individual's diet.

On one hand, such a method misses snacks, drinks, or desserts consumed outside normal meals. On the other hand, it is not always possible to visually determine all ingredients gone into the preparation of meals, such as the amount of oils, fats, salt, or sugar used, the type or quantity of spices added, and the cooking methods.

Such information could be obtained, though, by relying on a simple set of questions asked after submitting each meal picture, either through phone calls or SMS. Such questions would be further facilitated by relying on a dedicated app.

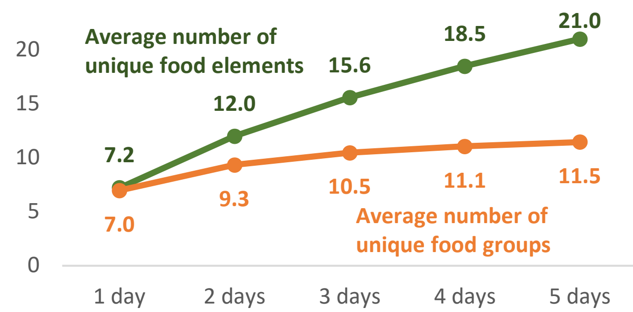
More broadly, while the submission of pictures and other relevant information could be free of cost through the provision of data plans or a free data-usage app, still not all targeted households would necessarily participate and/or send all required information for a comprehensive dietary assessment. In this regard, additional incentive mechanisms to encourage participation and compliance may be required.

DIETARY DIVERSITY

To test the effectiveness of this approach, we first compare dietary diversity (DD) measures for the same individual, as derived from the conventional survey-based recall and the picture-based methods. During the initial visit to the household, a standard 24-hour recall DD module (19 food groups) was administered by trained enumerators. After the survey, an individual from the household sent meal pictures for the following 5 days. The ingredients in these pictures, as identified during the labeling, were then matched to the same 19 food groups, allowing us to derive picture-based DD measures.

RECALL LENGTH AND HOUSEHOLDS' *FOOD UNIVERSE*

DD measures of pictures from 5 consecutive days are however not directly comparable to 24-hour measures. To illustrate this, the figure shows the average number (across households) of unique food elements (as identified by labelers) and unique food groups when considering all meals consumed over one or more days. It is interesting to note that, in line with previous studies, including more days increases the observed diversity of diets but only up to a certain point (note the concavity of the curve as a household reaches the extent of their "food universe").



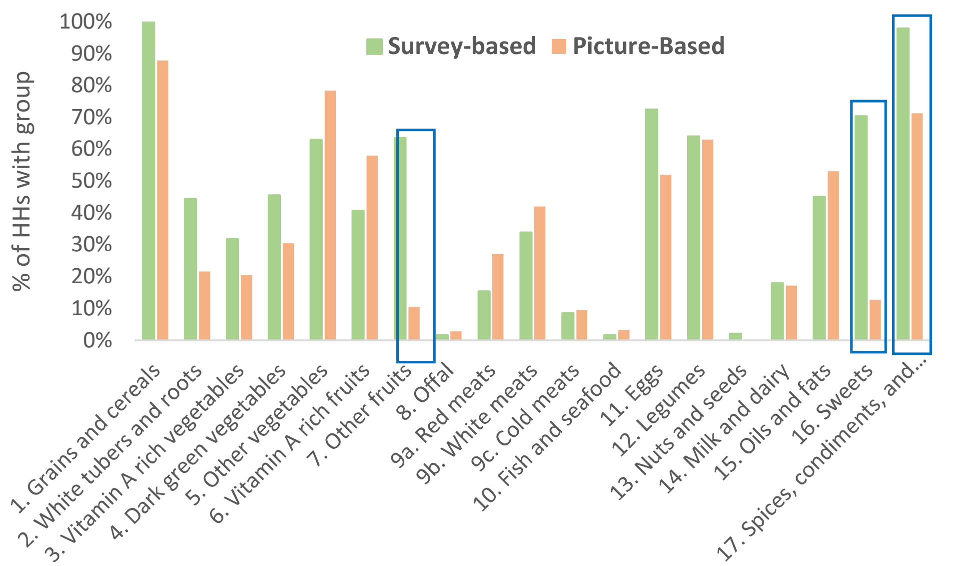
To make the two measures comparable, we thus report 1-day (with 3 reported meals) picture-based DD measures below.

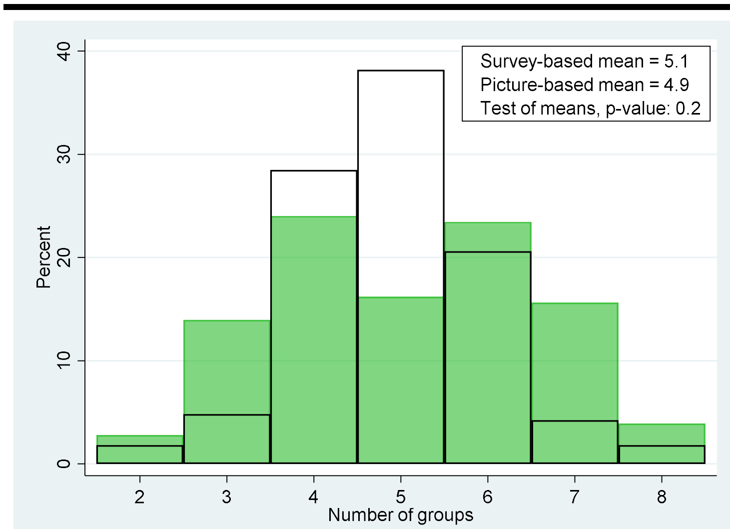
FOOD GROUP IDENTIFICATION

The figure shows that the food groups identified under both methods are very similar. Nevertheless, some food groups are harder to identify through pictures:

- ▷ **Other fruits**, generally consumed as snacks or dessert
- ▷ **Sugar and sweets**
- ▷ **Spices, condiments, and infusions** (though not clear in the graph since labelers tend to indicate salt as an ingredient)

We thus exclude these three groups from the DD comparisons below.





SURVEY-BASED RECALL VS PICTURE-BASED

The figure shows overlapped histograms of the number of food groups consumed by all households, as identified through a survey-based 24-hour recall (green) or through labeled meal pictures (transparent).

While the number of food groups consumed tends to be slightly higher under the survey-based identification, the results do not differ substantially between methods.

Certainly, while survey-based recall is the most commonly used method, it is not exempt from measurement error. Therefore, an in-the-lab comparison, where food items are objectively counted and weighted, would be needed to determine the superiority of either approach.

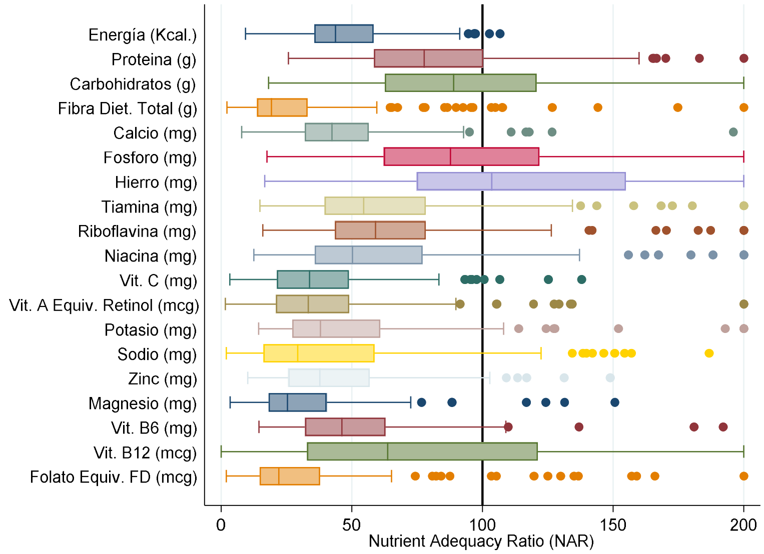
BEYOND DIETARY DIVERSITY

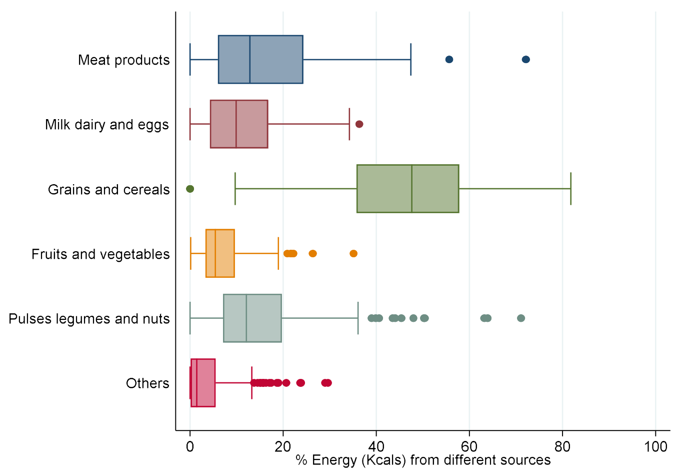
As described above, we rely on the 2012 INCAP food composition tables and the pictures' labels, including estimated weights for each food ingredient, to approximate intake per meal of an array of macronutrients and micronutrients.

NUTRIENT ADEQUACY

To derive Nutrient Adequacy Ratios (NAR), we compare intake across 3 meals (for days in which individuals sent a complete set of meal pictures) to the Recommended Daily Amounts (RDA) from the Food and Nutrition Board, Institute of Medicine, National Academy of Sciences Dietary Reference Intakes.

As observed, there is considerable variability in NARs, both across households and between nutrients. This evidence supports the potential for individual monitoring and tailored interventions to foster healthy and balanced diets, as an alternative to general nutritional information campaigns.



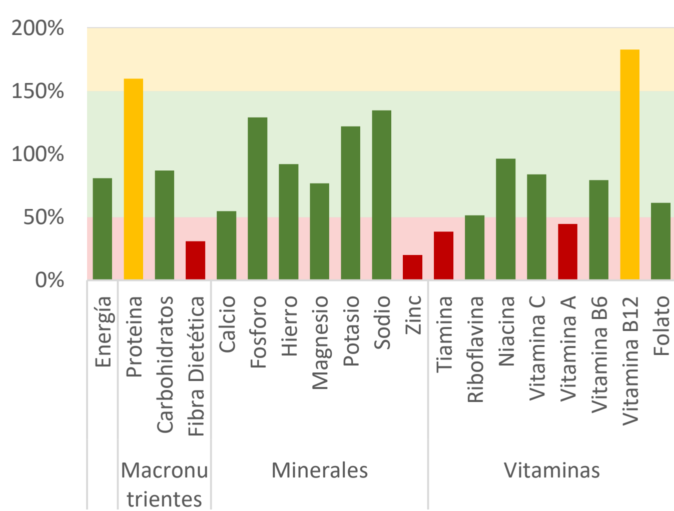


OTHER NUTRITIONAL MEASURES

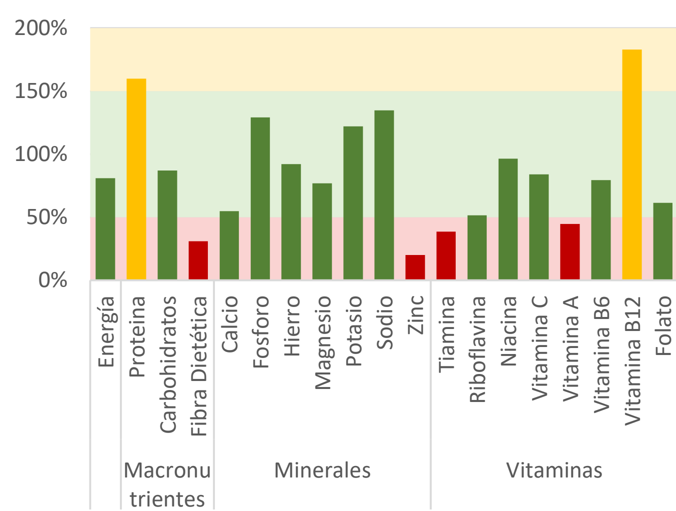
The rich data derived from pictures allows us to construct additional dietary measures, such as the percentage of energy (or macro or micronutrients) derived from different food sources (see figure on left).

In addition, we can construct individual-specific nutritional assessment reports (see example in figures below), showing observed nutrient gaps and excesses, that could eventually be accompanied with tailored recommendations based, for instance, on meals consumed by similar households in their geographic area (thereby incorporating dimensions of local access to and availability of foods as well as cultural factors).

Nutritional Assessment - Individual A



Nutritional Assessment - Individual B



Note: In this example, we use arbitrary NARs of 50/150% as thresholds. In practice, these would be adjusted following nutrient-specific literature.

SUMMARY

In this proof of concept study, we show the feasibility of deriving rich, detailed dietary information from smartphone pictures of daily meals sent by individuals from rural households over several days. We find reasonable participation rates and high compliance, with individuals willing and able to send pictures of sufficient quality in regular intervals.

While automated picture processing is the medium- to long-term goal, we show that manual labeling by local nutritionists is attainable, cost-effective, and even desirable when typical foods may vary across regions, such as in Guatemala. In addition, compared to other alternatives for collecting detailed dietary data such as survey-based recall or diaries, pictures can be obtained at a lower cost and over longer periods of time, and demand lower time and effort from individuals.

We identify some potential issues, such as: (i) accurately estimating the weight of food elements; (ii) distinguishing between visually-similar food elements; (iii) identifying unseen elements such as sugar, spices, or condiments; and (iv) not observing snacks and other foods or drinks consumed between meals. We argue, however, that most of these issues can be solved by improving the training of labelers and developing calibrated materials to be used as visual aids, developing protocols to handle labeling uncertainties (to derive plausible ranges of nutrient consumption), and complementary data collection tools such as phone calls, SMS, or in-app questions inquiring about consumption of snacks and elements not visible in the pictures, in addition to cooking methods or physical activity proxies to further adjust nutritional measures.

While this pilot study focused on rural households in the Western Highlands of Guatemala, such a tool can be easily rolled out into other rural areas as well as urban areas, informing issues around urban/rural dietary trends. In addition, while undernutrition is a more pressing problem among our study population, this tool would be equally effective for identifying dietary imbalances leading to the increasingly important problems of overweight and obesity in developing countries.

Overall, the proposed tool should not be viewed as a means to replace existing, proven methods for the collection and assessment of nutritional data, but rather as a practical complementary resource in the toolkit for remote collection of rich nutritional data, identification of individual-level nutritional imbalances, and the design of more direct, personalized recommendations for healthy and balanced local diets.

WAY FORWARD

- Extend study to other rural and urban contexts for replication and scaling-up purposes.
- Develop a free (ideally data-free) smartphone app to collect geotagged pictures, metadata, and gather additional information of relevance. The app could potentially allow for:
  - Submitting pictures for different household members, informing intra-household allocation of foods
  - Capturing basic socioeconomic and individual-specific information via short surveys
  - Questions inquiring about snacks consumed between meals and additional meal information such as sugar, salt, and condiments added, together with cooking methods and physical activity proxies
  - Providing incentives, in the form of airtime top-ups or mobile money (in settings where such infrastructure exists on a large scale), for submitting pictures, answering surveys, or showing adoption of recommendations
- Design a system to identify nutritional gaps or excesses and provide personalized feedback as a behavioral-change tool to promote healthy diets, with a double-burden perspective. In particular, we propose:
  - Delivering personalized reports indicating specific dietary gaps or excesses
  - Providing tailored recommendations on how to adjust diets based on local food availability, focusing on common nutrient-rich foods consumed by similar households in the area or meal recipes of local foods rich in a particular nutrient, aided by additional materials such as videos or simple nutritional education games
- Aim for the impact evaluation of existing nutritional interventions or some of the above elements, for instance by automatically randomizing new users of the app into different treatment arms and requesting new sets of pictures over time in regular intervals, resulting in rich nutritional panel datasets.
- A dedicated app would also allow for testing of different diffusion models, including reaching out to specific target groups, using mass media campaigns, or relying on alternative channels such as social media.