Literature Review Part 1

Review of methods used in food security assessment to measure food consumption and estimate energy intake of households or individuals

Submitted by:

Joanne Arsenault, PhD

August 2, 2015

Submitted to:

Christine McDonald
FEWS NET Nutrition Advisor
Chemonics International Inc.
1. Introduction

Energy intake is typically measured in individuals or households by surveying respondents about food consumption over a specified period. The gold standard of dietary assessment methods is direct observation of individuals or household members and weighing of all foods consumed, but this method is time-consuming and labor intensive.

Two types of food consumption surveys that can be used to calculate energy intake for food security assessment are: 1) household consumption and expenditure surveys and 2) 24-hour dietary recall surveys. Household consumption and expenditure surveys are typically conducted routinely by governments for economic assessments. 24-hour dietary recall surveys are usually used in nutrition surveillance to collect individual-level data on a subset or all members of the household, but are not typically conducted routinely in low-income countries.

Both methods rely on respondents to self-report their food consumption and so share similar types of respondent bias such as memory and estimations of quantity. This review discusses issues that have been identified with both survey types and the degree of accuracy to which the surveys estimate actual energy intake of the respondents.

2. Household Consumption and Expenditure Surveys

Household consumption and expenditure survey (HCES) is a collective term that includes surveys that collect information on food acquisition or consumption as one of many components. HCES are conducted routinely in many countries by government statistical services or through the World Bank’s Living Standards Measurement Study surveys (LSMS). HCES contain a list of foods and respondents are asked how much of each food (by weight, volume, and/or monetary value) the household has acquired and/or consumed during some reference period. Household Budget Surveys (HBS) or Income and Expenditure Surveys (HIES) usually collect information on foods acquired, not consumed, by the household and the monetary value of the food, for the main purpose of providing expenditure weights for a Consumer Price Index. LSMS have a primary focus of measuring distribution of living standards and include other modules on health and welfare, and collect information on household consumption in quantity measures as well as monetary value.

Energy intake of the entire household is calculated from reported food consumption in a number of steps. First, the foods amounts must be converted to a gram weight basis. If the food contains any non-edible portion such as peels or bones, this weight is subtracted from the total amount using known factors for edible portions. Nutrient values for foods are assigned using country-specific nutrient composition tables. Energy consumed from all foods is summed for the entire household. To estimate intakes on the individual level assumptions must be made about the intra-household distribution of food. The most commonly used method, the adult-male equivalent (AME) method, assumes that food is allocated based on energy needs and each household member’s energy needs is expressed as a proportion of the adult male’s energy needs, summed, and food amounts are allocated based on each members fraction of the total energy needs of the household (1). If the survey asked about whether household members
were absent or if guests consumed any food during the reference period, this information is taken into account when calculating an average daily intake.

2.a. Issues with HCES

This section addresses many of the issues that have been raised regarding the use of HCES to obtain estimates of actual food and energy consumed and provide information for food security assessment (2-7). Given the recognition that HCES are widely available and being used for food security and nutrition assessment, but have several limitations, research efforts are underway to better understand how HCES can be improved for these purposes (8, 9).

Acquisition vs. consumption. If the survey asks respondents to recall foods acquired but not consumed, the survey will capture foods being stockpiled for use after the period of reference and information will not be captured on foods being consumed from stocks obtained prior to the reference period. Although households in developing countries do not typically have the capacity to store perishable items for long periods, this could impact estimates of staple grains that can be stored and may provide a substantial proportion of a household’s daily energy intake. Smith et al (3) point out that at any given time some households will be stockpiling foods and some will be drawing from stocks. Therefore, the mean estimate of a population should theoretically equate food acquisition to food consumption. However, at the individual household level, the estimate may not be precise if they are stockpiling or drawing from stocks. In addition, an estimation of the proportion of households with consumption below a threshold may be imprecise. Evidence from three surveys from Kenya, Philippines, and Bangladesh which collected both food acquisition and consumption data by 24-h dietary recall from the same households demonstrate that there were small differences (1-5%) in estimated energy intake per capita between estimated derived from acquisition and consumption; however, energy acquisition was systematically lower than energy consumption among the lowest income groups and the opposite for higher income groups (3).

Recall period. Most HCES use a recall method, asking respondents to recall foods acquired or consumed over some reference period such as the past week or month. This relies on the cognitive ability of respondents to remember accurately. Some research suggests a period of two weeks or less is preferable (4, 10, 11). Beegle, et al. reported lower consumption estimates with a 14-day recall than a 7-day recall in Tanzania (10). A review of 100 national surveys found a range of recall periods from 1 to 365 days, but 70% had an optimal reference period of two weeks or less (4).

Food list. HCES use a list of foods that are commonly consumed in the population and must be comprehensive enough to reflect usual consumption but not so long as to make the survey overly burdensome on respondents. The food list should reflect the dietary and cultural habits of the population that is being assessed. Food lists are not standardized across surveys, but the comprehensiveness and representativeness of standard food groupings could be improved. A survey of 100 HCES found a mean of 204 food items, ranging from 19 to 5407 items; however,
diary-type surveys had more items than interview-type recall surveys which had a mean of 102 food items (4). The detail of description of food items is important for the calculation of energy because the foods must be matched to appropriate food items in a nutrient composition database. Foods on the HCES may be broader than the selections in the nutrient database – for example, the HCES may list some major grain items but also have a category of ‘other grains’.

Conversions. A major issue with accurate assessment of energy from HCES is the conversion of the unit of measure obtained from the survey to an edible portion weight. The HCES may use standard weight or volume measurements or local units such as bags or baskets. HBS may only use monetary values which must be converted to weights and accuracy depends on collections of price data from local markets at the time of data collection. Local units must also have factors that must be obtained at the time of data collection to convert to standard quantities. If this information is not obtained during the survey, an indirect method will be needed which will lower the precision of the resulting estimation of energy. A second conversion step is necessary if the food contains inedible portions such as peels or bones. For some foods this can be a substantial proportion of the purchased weight.

Other issues. HCES may not collect data on food wasted or on food given to guests or pets. Furthermore, these surveys might not measure foods consumed outside of the home such as from school or restaurants, both of which can be rather substantial amounts of food. In India, failure to account for foods consumed outside the home in HCES is hypothesized to at least partially account for a decline in energy consumption per capita over two decades (10.3% in rural and 7.2% in urban areas) despite a decline in poverty (12). This is based on evidence from an analysis of a sample of 44 HCES from various countries that showed increasing energy consumption per capita as the quality of HCES improved in regard to asking about food away from home. An additional issue is seasonality, which may not be captured if the survey is not performed throughout the year and can introduce a large error. It is unclear how well a 2-4 week period of consumption translates to an annual estimate.

2.b. Studies comparing HCES to individual-level 24-hour recall dietary intakes

This section reviews some studies that compare HCES to dietary surveys using the 24-h recall method, which is the most commonly used method by nutritionists but not considered the gold standard and has its share of bias as will be discussed in section 3. Caution is warranted with the use of the term ‘consumption’ – economists may use the term to describe acquisition and it may not be clear what the word is actually capturing without seeing the wording of the question on the survey.

Expenditure surveys. Two studies reported higher estimates of energy consumption from household budget surveys than dietary intake surveys (13, 14). One study compared household budget survey (HBS) data in Germany and Greece with published dietary intake data (13). In Germany, HBS were compared to the Bavarian Food Consumption Survey 2002-3 which
assessed intake using three 24-h recalls among participants aged 14 years or older. In Greece, HBS were compared to a food frequency questionnaire, which is not designed to estimate energy intake and those results will not be presented here. The HBS was described as collecting data on ‘food availability’ and respondents were asked to record purchases, contributions from own production and foods received as gifts, but did not ask about foods consumed outside the home. Food amounts in the HBS were adjusted for wastage in the home by subtracting 10% of weight. Households from the German HBS were selected if household members ages were ≥ 14. In Germany, the mean energy intake from HBS was 3317 kcal and from the Bavarian Survey was 2000 kcal, which is an overestimation of recall by 66%.

The second study compared household food budget surveys (HBS) in 4 countries (Canada, Finland, Poland, and Spain) with individual level dietary studies and food balance sheets (14). Three of the countries included 24-h recalls as the method of individual level dietary survey and those results are reported here. The mean energy intakes from HBS were only 3% higher in Canada, but were 26% and 30% higher in Poland and Spain, respectively. It was unclear why the results varied by country.

Expenditure and consumption surveys. One study found similar mean energy intake from a household expenditure survey and a 24-h dietary recall survey in the same households in Kenya and the Philippines (15). The food expenditure survey asked about food purchased or consumed in the past week in Kenya and the past month in the Philippines, and the Philippines survey included questions on meals to guests to permit a downward calculation of family food availability. The authors referred to estimates from the HBS as ‘food availability’ although they had apparently asked about consumption as well as purchases. The mean daily ‘consumption’ per capita in the Philippines was almost equal (1805 kcal by expenditures and 1811 kcal from the dietary recall). However, among households in the lowest quartile of food expenditures, mean calorie consumption was lower by expenditure (1350 kcal) than dietary recall (1626 kcal), while in the highest quintile of food expenditures, consumption by expenditure survey was higher (2312 kcal) than by dietary recall (1983 kcal). Household availability estimates from Kenya were about 100 kcal less than estimates of intake by dietary recall.

A study compared the 2006 Uganda nationwide HCES study with 24-h recall data collected in 2008 among women of reproductive age and children 24-59 months of age from three regions of Uganda (16). The HCES collected information on food quantities and expenditures of foods consumed using a list of 52 foods. In order to assign nutrient values to foods that were broadly defined in the HCES, such as beans, the authors matched items that were commonly consumed in the 24-h recall using two methods: 1) using 4-5 of the most commonly consumed foods, and 2) using the single most commonly consumed food in the 24-h recall. Household intakes were adjusted using the adult male equivalent method (1) to predict intake of women and children from the HCES. The resulting mean energy intakes from both surveys were similar and no statistically differences were found for any region or population group. Among women, mean energy consumption for the 3 regions were between 1733 and 2185 kcal by dietary recalls, and by HCES were 1678 -1845 kcal (method 1) and 1745-2152 kcal (method 2). Among children, the regional mean energy intakes were 1010-1633 kcal by dietary recalls, and by HCES were 970-1125 kcal (method 1), and 997-1053 kcal (method 2). It is
unclear how results would compare if the 24-h recall data were not used to inform the selection of foods to assign to food categories in HCES, although using one food item in a food grouping in HCES (method 2) resulted in similar results to using 4-5 foods (method 1) to categorize a food group.

A study using data collected from the 2011/2012 Bangladesh Integrated Household Survey, which was a nationally representative sample of 5503 households, compared energy intakes from a household-level consumption 7-day recall to a 24-h recall of all individuals in the same households (17). The mean daily intake per person from the 7-day recall was 2357 kcal/d, which was 14% higher than the mean of 2064 kcal/d from the 24-h recalls. Using AME to estimate energy intake by age groups, the greatest deviation of energy intakes derived from the 7-day food recall from the 24-h recall was for the youngest children (< 4 years). The author speculated that the 7-day recall may have not accounted for food that was wasted or foods purchased rather than consumed.

In summary, these studies indicate a tendency of HCES to overestimate energy intake when compared to 24-h recalls, particularly in surveys that ask about food acquired rather than consumed. Overestimates of energy consumption ranged from 3-66% in expenditure or acquisition studies. In studies that incorporated questions on consumption, overestimation ranged from 0-14%. Most studies compared mean energy consumption and not the distributions of intakes. A “how-to’ guide on analyzing HCES suggested dropping outlier data for households with daily energy acquisition per adult equivalent of 12,000 kcal (5), which indicates acceptance of energy estimates well above what could actually be consumed by an individual. In food security assessment, interest is at the lower end of the spectrum of intakes. Despite the disparity of estimates of energy consumption at the lower income level in the Phillipines data (15), there were relatively small differences in the estimated proportion of the energy-deficient households (3).

A major problem with these comparison studies is the reference method used - the 24-h recall surveys. 24-h recall surveys are typically used in the nutrition community, but rely on similar sources of error as household surveys – memory and quantity estimation. The next section will discuss the 24-h recall method and validation studies for estimating energy intake.

3. 24-hour Dietary Recall Surveys

The most commonly used method to collect dietary information in the nutrition field is the 24-h recall. The survey is administered by interview and the respondent is asked to recall all foods consumed in the past 24 hours from the time of the interview or the day preceding the interview. The interview typically consists of multiple steps to aid the respondent in remembering foods – first by a quick listing of foods and then going back through each eating occasion for additional details of the foods and preparation methods. The interview then goes through each food item to estimate portion sizes, usually using visual aids to assist the respondent, and a final review of foods is conducted. An interactive 24-h recall method has been developed and used extensively to collect information in rural populations in developing countries (18).
One recall day does not characterize an individual’s usual food intake and the number of
days to characterize an individual’s nutrient intake varies by the variability in day-to-day intake
of foods. A one-day survey can approximate the mean intake of a population, but not the
distribution of intakes in the population due to intra-individual variation. A second day of recall
is usually sufficient to estimate the intra-individual intakes and statistically remove it in order to
assess a usual intake distribution of the population and estimate the proportion of the
population meeting a threshold intake level (19). Special software programs have been
developed to obtain usual intake distributions. Caution should be given to any estimate of
prevalence of inadequate nutrient intake from a survey that only contains one day of dietary
intake per individual, as the estimate can be inaccurate if the true prevalence is at the tails of
the distribution which will be artificially wide when intra-individual variation has not been
accounted for statistically.

The 24-h recall method is widely used because it is quicker and less expensive than the
more accurate method of weighed food records, where in a developing country setting an
observer remains with an individual throughout the day weighing all foods prepared and
consumed by the individual. The coding of information after collection is a tedious, time-
consuming process involving several steps: conversion of amounts consumed that have usually
been reported in volumes or other less quantitative measures such as ‘mounds of rice’,
calculation of average recipes because of lack of detailed recipe level at the household or foods
consumed outside of the household, preparation of a food composition table of nutrient values
of foods which is often not available in developing countries, and checking for implausible
values.

3.a. Issues with 24-hour recall surveys

Sources of measurement error in the 24-h method are similar to those that would be
encountered in HCES relating to the ability of respondents to recall correctly the foods
consumed, but these issues have been addressed more in the research literature regarding 24-
h recalls than HCES. The main sources of error are memory and portion size. Respondents may
unintentionally omit food items consumed or report food items they did not consume in the 24-
h period – referred to as errors of omission or addition. Estimation of portion size is difficult for
respondents, particularly without the use of life-like food models or replicas of typical portion
sizes. Respondent bias also occurs when respondent may feel the need to give socially
desirable answers, such as the mother of an undernourished or an overweight child reporting
feeding the child nutritious foods. The food composition tables of nutrient values can be a
source of error in any dietary assessment method, particularly in developing countries which
may not have laboratory analyzed nutrient values for local foods. In many cases, the US
Department of Agriculture food composition table is applied when local information is not
available.

The validity of 24-h recalls can be assessed against another dietary method that has less
measurement error and greater validity (a reference method) or against a biomarker which is
an independent measure. Dietary methods share some of the same biases because they rely
on subjects reporting and so it is difficult to truly assess one dietary method against another.
Direct observation of consumption is typically done in institutionalized settings where individuals may not be consuming their usual diet. In developing country settings, several validation studies have tested the 24-h recall against a whole-day observation and weighing in the subjects homes and may represent a more typical consumption. Biomarker studies of energy expenditure, such as the doubly labeled water method, can be used to estimate energy intakes under conditions of energy balance whereby energy intake equals energy expenditure (20). This method is very expensive and has been used to validate 24-h recalls in primarily developed country settings. The next section will review validation studies, with an emphasis of studies conducted in developing countries.

3.b. Studies comparing energy intakes by 24-h recall to ‘gold standard’ dietary assessment method

**Studies reporting overestimation of energy intakes with 24-h recall.** Several studies that compared energy estimates from 24-h recall to weighed records from direct observation of individuals reported an underestimation of energy intakes (21-24). Most studies used a 12-hour period of observation for the weighed records (WR) whereby a field worker remained in the home of the respondent and weighed all food prepared, portions served, and portions uneaten, and then returned to the home the following day to capture by recall any foods eaten after the worker left. This method is not completely independent of the recall, but it is impractical for workers to remain in the home for 24 hours and most foods were consumed during the observation period. Methods used to assess validity include comparisons of mean or median energy intakes, correlations, comparing classifications of intake (e.g., quartiles of energy intake), and the Cohens weighted kappa statistic or intra-class correlation coefficients to measure the agreement of classifications of intake.

In a study among school children and women in Kenya, the 24-h recall underestimated energy intake by 7% in the children and 5% in their mothers based on comparison of means (21). 15-21% of food items on the WR were omitted from recalls and 8-10% were additions (appeared on the recalls but not on the WR). During the hours of the observation period only (0800-1700 hours), the weighted kappas of 0.26-0.30 suggested poor agreement.

A study in Ethiopia among 58 women reported that median energy intakes from one recall day were significantly lower than from WR by about 10% (22). Thirty-one percent of the women’s energy intakes were within 10% of the WR, 64% were within 20%, and 81% were within 30% of WR. The Cohen’s weighed kappa for classification agreement of energy intake quartiles was 0.44 (fair). 36% of recalls had errors of omission and 25% had addition food items that were not on the WR. Underreporting was a particular problem with cereal items.

Two studies from Malawi reported underreporting in 24-h recalls for children (23) and women (24). Three days of recall and WR were obtained among 4-6 year old. Although the difference in median energy intakes between methods was reported to be non-significant, the sample size was small (29 children) and the median energy intakes from recalls were about 200 kcal/d less than WR (~14%). Only 17% of energy intakes were within 10% of WR on the same day. The intra-class correlation of energy intakes on the same day was 0.42. The authors noted a particular problem with underestimation of maize porridge. A study among 60 pregnancy...
women comparing two days of recall and WR reported median energy intakes by recall of about 300 kcal (16%) less than WR (24). The Cohen kappa of agreement was poor (0.22). Only 22% of recalls energy intakes were within 10% of WR, 32% were within 20%, and 81% were within 30%. The women reported significantly lower percentage of energy from cereals than by WR, which the authors noted is particularly problematic given the high proportion of energy from cereals in this population.

**Studies reporting mixed results or no underestimation of energy intakes by 24-h recall.** One study that compared two days of recall and WR in children reported different findings in two villages in Ghana (25). In one village, median energy was about 15% lower with recalls than WR by about 200 kcal (15%) and the intraclass correlation was 0.06. 35% of energy intakes by recall were within 10% of energy intakes by WR and only 20% of children with WR energy intakes in the lowest tertile had recalled intakes in this tertile. In the other village, median energy intakes were similar and the intraclass correlation of 0.25 was significant. However, only 27% of energy intakes were within 10% of WR and 50% of children were classified in the lowest tertile by both methods. Underreporting was due primarily to snack food consumption and a staple prepared from corn and cassava. In contrast to the previous studies, two studies with children under 2 years of age conducted in Senegal (26) and Malawi (27) reported good agreement between energy intakes by the two methods.

3.c. **Studies comparing energy intakes by 24-h recall to ‘gold standard’ energy expenditure method**

Only one study was identified from a developing country comparing doubly-labeled water (DLW) to self-reported food intake (28). A five-country study with 324 individuals of African ancestry reported under-reporting of energy intake in all 5 countries (28). The largest underreporting was in South Africa (52%), followed by Ghana (22.5%), Jamaica (17.9%), Seychelles (25%), and the US (18.5%).

In a review published in 2001, six studies that included a comparison of DLW to 24-h recalls reported underreporting energy intake by about 17% (29). Most of the studies included relatively small numbers of women. A later study with 20 US women that compared the automated multiple pass 24-h recall method used by the US National Health and Nutrition Survey (NHANES) to energy expenditure by DLW reported no bias with 24-h recall (30).

In summary, almost all of the reviewed validation studies that compared the 24-h recall method to the gold standard dietary assessment method or biomarker method reported underestimation of energy intakes. Estimates of underestimation in the studies ranged from 0-52%, but most were between 10-25%.
4. Conclusions

HCES surveys, both those that ask about expenditure/acquisition and those that include consumption, tend to overestimate energy intake compared to 24-h recall surveys. Estimates of energy intake from expenditure/acquisition surveys overestimated energy intake from 3-66%, but this only included comparisons in four countries and overestimates were 26-30% in two of the country comparisons. The degree of overestimation was less with HCES that included consumption questions, ranging from 0-14%. On the other hand, 24-h recall surveys underestimated energy intakes by 10-25% compared to gold-standard methods of direct observation/weighing and doubly-labeled water.

The small number of comparison studies that compared HCES to 24-h recall and the large variability of the results make it difficult to make a broad conclusion other than the HCES that incorporate consumption questions seem to be closer to energy intake estimates by 24-h recall. The variations in HCES with regard to food lists, length of recall period, questions pertaining to presence or absence of individuals partaking of food, and coding and analytic methods likely contribute to the variation seen in comparison studies. Only one of the reviewed studies included information about how well the HCES estimated intake the lower socioeconomic level or estimated the proportion of energy-deficient individuals or households. This information is critical to evaluating the validity of using HCES for food security assessment. It is also problematic that the comparison studies used the 24-h recall as the reference method to compare energy intakes. It is unlikely that national level HCES could be validated against a gold-standard dietary assessment method because of the expense of conducting these methods with a large sample, but improvements of HCES design should yield more consistent results for use in food security estimates.
References


