Original Article

The validity of a structured interactive 24-hour recall in estimating energy and nutrient intakes in 15-month-old rural Malawian children

Chrissie M. Thakwalakwa[*†]**, Heli M. Kuusipalo**[†‡]**, Kenneth M. Maleta**[*]**, John C. Phuka**[*†]**, Per Ashorn**[†‡] **and Yin Bun Cheung**[§¶]

*Department of Community Health, College of Medicine, University of Malawi, Blantyre, Malawi, †Department of International Health, University of Tampere Medical School, Tampere, Finland, ‡Department of Paediatrics, Tampere University Hospital, Tampere, Finland, §Department of Biostatistics, Singapore Clinical Research Institute, Singapore, and ¶Office of Clinical Sciences, Duke-NUS Graduate Medical School, Singapore

Abstract

This study aimed to compare the nutritional intake values among 15-month-old rural Malawian children obtained by weighed food record (WFR) with those obtained by modified 24-hour recall (mod 24-HR), and to develop algorithm for adjusting mod 24-HR values so as to predict mean intake based on WFRs. The study participants were 169 15-month-old children who participated in a clinical trial. Food consumption on one day was observed and weighed (established criterion) by a research assistant to provide the estimates of energy and nutrient intakes. On the following day, another research assistant, blinded to the direct observation, conducted the structured interactive 24-hour recall (24-HR) interview (test method). Paired *t*-tests and scatter-plots were used to compare intake values of the two methods. The structured interactive 24-HR method tended to overestimate energy and nutrient intakes (each $P < 0.001$). The regression-through-the-origin method was used to develop adjustment algorithms. Results showed that multiplying the mean energy, protein, fat, iron, zinc and vitamin A intake estimates based on the test method by 0.86, 0.80, 0.68, 0.69, 0.72 and 0.76, respectively, provides an approximation of the mean values based on WFRs.

Keywords: weighed food record, structured interactive 24-hour recall, validity, adjustment algorithms, predictive power.

Correspondence: Ms Heli M. Kuusipalo, Department of International Health, Arvo Building, Lääkärinkatu 1, University of Tampere, FIN-33014, Tampere, Finland. E-mail: heli.kuusipalo@uta.fi

Introduction

Accurate, effective, comprehensive and simple to use dietary intake assessment methods are needed in the communities to assess dietary status for both research and health care purposes. Weighed food record (WFR) is considered by some researchers to be the most accurate method for providing quantitative dietary data (Ferguson *et al.* 1995). It provides accurate group mean estimates of energy and other nutrients. However, the method is very time consuming and expensive. It also has a high-respondent burden and could be disruptive to the participants.

Therefore, it is difficult to apply this method to studies that involve a large number of participants and/or days of assessment.

The 24-hour recall (24-HR) is a retrospective dietary assessment method that provides information on the respondent's food intake during the previous 24-hour period. The participant recalls all foods and drinks consumed and identify portion sizes with the aid of a trained interviewer. 24-HR requires less equipment, personnel and time than WFR. These advantages render it an attractive alternative for large-scale surveys (Dop *et al.* 1994).

However, the 24-HR has some constraints because it relies on the respondent's ability to remember and accurately quantify the food items consumed on the previous day. A modified 24-hour recall (mod 24-HR) can reduce these memory lapses (Ferguson *et al.* 1995). In this method, training of interviewers on portion size estimation is provided and picture charts are supplied to respondents for use as a checklist. Standard spoons, bowls and plates are used to help respondents estimate the amount of food consumed (Ferguson *et al.* 1989, 1995).

The mod 24-HR has previously been used to quantify dietary intakes among pregnant women and 3- to 4-year-old children in rural Malawi (Ndekha *et al.* 2000; Maleta *et al.* 2003a). It has also been used in other settings, but its predictive power in estimating energy and nutrient intake in children below the age of 2 years has not been assessed. In many developing countries, including Malawi, growth stunting and undernutrition are common in children aged less than 2 years (Maleta *et al.* 2003b; Phuka *et al.* 2008). User-friendly tools to assess the nutritional intakes of these children are important for understanding the epidemiology of growth and nutrition. This age range is a prime target for complementary feeding programmes, including ours in rural Malawi. To support our projects on complementary feeding in Malawi and to facilitate similar studies in other developing countries, we assessed the accuracy of the mod 24-HR in infants/young children at the group level. Specifically, this study aimed to compare the mean intake value of energy, protein, fat, iron, zinc and vitamin A from foods and beverages other than breast milk among 15-month-old rural Malawian children obtained by WFR (an established criterion) with those obtained by mod 24-HR (test method), and to develop algorithm for adjusting mod 24-HR values so as to predict mean intake based on WFR.

Methods and materials

Study area

The study was carried out in Lungwena, Mangochi District, Southern Malawi. The area is rural; approximately 100 km^2 in size and the nearest town (Mangochi) lies some 30 km away. The population consists of about 27 000 people in approximately 5200 households in 26 rural villages (Humphreys *et al.* 2005). Farming and fishing are the main occupations. Maize is the main crop and staple food grown in the area. Other crops are cassava, sweet potatoes, groundnuts and millet. Very few families in Lungwena keep livestock (chicken and goats).

Participants

The children participating in a clinical trial called Lungwena Child Nutrition Intervention 3 (LCNI-3) were eligible to participate in this validation study. LCNI-3 aimed to test the growth and health benefits of daily supplementation with either lipid-based nutrient supplement (LNS) or corn-soy blend (CSB) between 6 and 18 months of age. LNS is an energy dense combination of peanut paste, sugar, milk powder, vitamin and mineral mix (Phuka *et al.* 2008).

Study design

This substudy took place between August 2005 and November 2005, about 9 months after the infants were enrolled into the clinical trial. Apart from this substudy, the trial did not involve any other WFR or 24-HR assessment of energy and nutrient intake. Food intake data was obtained by one interviewer using WFR and by a second interviewer using mod 24-HR on the following day. Breast milk intake was

Key messages

- To compare two food-intake research methods and to find a correction value that can be used in population level.
- The weighed-food record method is time- and personnel-consuming method compared with 24-HR.
- We found the correction values that can be used in rural low-income African setting.

not included in this comparison of the two methods. The data were collected in Chiyao (language) by five interviewers who were trained for 1 week in conducting dietary intake surveys using both methods. The effects of LCNI-3 food supplementation and breast milk intake have been reported elsewhere (Galpin *et al.* 2007; Phuka *et al.* 2008).

A list of commonly used local food items, composite dishes and their raw material contents and cooking methods were collected through focus group discussions with mothers of 14- to 15-month-old children. This information was used to create a local food picture calendar that was used during the recall. The picture calendar included colour photographs of the food shown in their natural sizes.

The day before conducting the WFR, the interviewer sought consent and made an appointment to spend 12 h (6 am–6 pm) with the participant's family to observe, weigh and record all foods and drinks that the participant consumed. The interviewer also delivered the local food picture calendar and asked the parent to tick all foods the child ate the following day. There is no electricity in the study area. After sunset, food preparation is difficult and food consumption is inconvenient. The norm is that breastfeeding may take place after sunset but food intake is unlikely.

Data collection using the WFR

The interviewer arrived at the participant's home before 6 am to capture the participant's very first food. At the participant's home, he or she observed the recipes of the prepared foods and weighed all raw ingredients before they were processed. The parent was asked to feed the child from a separate plate for easy weighing of the food.

The interviewers weighed every food and beverage before and after the participants consumed it to determine the actual consumption. Philips electronic child scales (SECA 834, Chasmors Ltd., London, UK) with a precision of 1 g were used. Scale accuracy was checked daily by calibrating with known weights. The interviewer recorded all the collected information and the exact time when these foods were taken on the WFR form. The results were not discussed with the guardians.

Data collection using the structured interactive 24-HR method

The mod 24-HR (Ferguson *et al.* 1989, 1995; Gibson & Ferguson 2008) was conducted a day after the WFR. On the data collection day, the parent was asked to give the marked food picture calendar to the interviewer. The food calendar was marked on the observation day, before the day of the recall interview. This is part of the mod 24-HR protocol developed and used by Gibson & Ferguson (2008). It is intended for measuring energy and nutrient intakes generally in nutrition research and programmes, not just for the present validation study. Then the parent was asked to list all foods eaten by the participant the preceding day. They were also asked to state the feeding times in the local language (Chiyao) – kundaw (morning), musi (early afternoon), ligulo (late afternoon) or chilo (after darkness starts), as they did not have clocks. The interviewer compared oral information with what was marked on the calendar. Any discrepancies were discussed. The parent was then asked to describe in detail each of the foods or beverage consumed by the child. This included ingredients, amounts and methods used to prepare the food or beverage. Some salted replicas (nsima, dry fish, boiled cassava and sweet potatoes, etc.), standard spoons, cups and plates were provided to help the guardian to estimate portions consumed. All the collected information was recorded on a specially developed mod 24-HR questionnaire.

Calculating energy and nutrient intake

The amounts of foods from the WFR and the mod 24-HR were converted into grams. The nutrient values were computed using a self-made computer program installed in Microsoft Excel 7.0 (Microsoft corp, Redmond, WA, USA) (Ndekha *et al.* 2000). The program has pre-entered data on raw food groups and is linked to a food composition database with a total of 212 foods that were coded into 14 food groups. These groups include starchy staples, legumes, fats and oils, sugars, roots, meats, dairy products, vegetables, beverages, fish, fruits, spices, LNS and CSB. For the food composition database, 118 foods were

taken from the Kenyan food composition table and 64 were from the international Minilist nutrient database (World Food Dietary Assessment System, version 2.0, Office of Technology Licensing, University of California Berkeley). The composition of remaining foods were analysed by Malawi Bureau of Standards.

Statistical analysis

Statistical analyses were performed using Stata 9.2 (StataCorp, College Station, TX, USA). Baseline anthropometric indices length-for-age z-scores, weight-for-age z-scores and weight-for-length z-scores were calculated using Epi Info 3.3.2. software [Centers for Disease Control and Prevention (CDC), Atlanta, USA], which were based on the CDC 2000 growth reference (Kuczmarski *et al*. 2000).

The mean and standard deviation were calculated for each of the total energy and nutrient intake and the significance of the difference between the intake values of the WFR and mod 24-HR was evaluated by paired *t*-test. The intraclass correlation coefficient was used to estimate the agreement between the two methods (Donner 1986). To further demonstrate the pattern of agreement, energy and nutrient intake measurements from the two methods were compared graphically. In case they disagreed, an adjustment model based on the regression-through-the-origin model (Montgomery *et al*. 2001; Eisenhauer 2003) was used for each of the nutrients. The model is based on the assumption that the regression line for two methods measuring the same observations should go through the origin (as will be shown in figures). It fits a regression line in the form of $Y = bX$, with Y and X as observed values from WFR and mod 24-HR, respectively, and b as the regression coefficient, or the estimated ratio between the observed WFR and mod 24-HR values. If both methods give the same value, the data points should go through the origin and lie on a 45-degree line showing '$Y = X$', i.e. $b = 1$. The estimated ratio can be used to adjust the mod 24-HR values to give a predicted WFR. In order to validate the regression models and adjusted predictions, we used the 'leave-out-one cross-validation' approach (Montgomery *et al*. 2001) in getting the predicted

values. That is, a regression model was run without subject i, the predicted value for subject i was obtained by using the model coefficient estimated from the $n-1$ other subjects, and this process was repeated n times to obtain the predicted value for each of the n subjects. We used these predicted values to calculate the prediction error sum of squares (PRESS) and R-squared for prediction based on PRESS (R-squared PRESS) to assess model predictive ability (Montgomery *et al*. 2001). A paired *t*-test compared the observed WFR values against the predicted values based on the leave-out-one cross-validation. As supplementary analyses, we also used the conventional $Y = a + bX$ regression model to predict the intake values (also based on leave-out-one cross-validation) and assessed their disagreement with the WFR by intraclass correlation coefficients. The results were compared with the results of the regression-through-the-origin models.

Furthermore, we calculated the prevalence of intake of iron, zinc and vitamin A below recommended nutrient intake (RNI) level, i.e. RNI = 7 mg, 3 mg and 400 μg, respectively (FAO/WHO 2004) according to WFR, mod 24-HR and the predicted values based on cross-validation. Because few children in this sample had intake reaching RNI level, we also estimated the prevalence of reaching 50% of RNI. We compared the estimates of the prevalence using the McNemar's test. We did not examine energy, protein and fat intake level in this regard because we did not measure intake from breastfeeding.

Ethics

The study was approved by the College of Medicine Research and Ethics Committee, Malawi, and Pirkanmaa Health Care Area Ethics Committee, Finland, and registered with clinical trials (Identifier: NCT00131209). Informed consent was sought from at least one guardian before enrolment into LCNI-3 trial.

Results

Out of the 182 eligible infants in the clinical trial, nine died and four moved out of the catchment area before this substudy started. All the remaining 169 agreed to

participate in this validation study. Table 1 shows the baseline characteristics of the 169 participants at the start of this validation study.

None of the intakes reported in the mod 24-HR were reported to occur after darkness started (chilo). The mean and standard deviation of the two methods concerning the energy, protein, fat, vitamin A, zinc and iron intakes were calculated. As shown in Table 2, the mod 24-HR methods gave higher mean intake estimates than the WFR for each energy or nutrient examined (each $P < 0.01$). The intraclass correlation coefficients ranged from 0.42 (zinc and iron) to 0.83 (vitamin A).

The energy and nutrient intake values from the weighing method (*y*-axis) were plotted against recalled intake values (*x*-axis). A 45-degree line and regression line were added to the scatter plot

(Fig. 1a–f) to facilitate visualizing the disagreement. Two individuals had very high zinc intake according to the weighing method (Fig. 1e). Because the plots for zinc and vitamin A were quite affected by a few large values, we have additionally included scatter-plots that zoom into the region of zinc < 10 mg and vitamin A < 200 μg area (Fig. 1e,f supplement). All the plots showed substantial overestimation by the mod 24-HR. We found that the regression coefficient values (WFR-to-mod 24-HR ratio) for energy was 0.86, protein 0.80, fat 0.68, iron 0.69, zinc 0.72 and for vitamin A 0.76 as shown in Table 3. All regression coefficients were highly significant (each $P < 0.001$). The R-squared PRESS was only 0.48 for zinc, but the values for the other nutrients ranged from 0.72 (fat) to 0.89 (energy).

The means of the predicted energy and nutrient intake values based on the cross-validation approach are shown in Table 3. A paired *t*-test compared the predicted values and the observed WFR values. The differences between the two series for all the nutrients were not statistically significant, indicating that predicted values could be used to approximate the mean WFR values. However, the intraclass correlation coefficients were similar to that between WFR and mod 24-HR, demonstrating disagreement between individual level data. Supplementary analyses using the with-intercept regression models gave intraclass correlation coefficients (against WFR) smaller than those using regression-through-the-origin models (details not shown in table).

Table 4 shows that the estimates of prevalence of iron intake and vitamin A intake below RNI level

Table 1. Participant characteristics at the beginning of the study (*n* = 169)

Parameters	Values
Mean (SD) age	13.7 (0.6) months
Mean (SD) weight	9.0 (1.1) kg
Mean (SD) length	72.0 (2.7) cm
Mean (SD) MUAC	14.2 (1.1) cm
Mean (SD) head circumference	45.9 (1.6) cm
Proportion of boys	84/169 (49.7%)
Mean (SD) LAZ	−1.56 (0.83)
Mean (SD) WAZ	−1.44 (1.10)
Mean (SD) WLZ	0.08 (0.99)

MUAC, Mid-upper-arm circumference; LAZ, length-for-age *z*-scores; WAZ, weight-for-age *z*-scores; WLZ, weight-for-length *z*-scores; SD, standard deviation.

Table 2. Mean (SD) energy and nutrient intake values measured by weighed food record (WFR) and modified 24-hour recall (mod 24-HR), mean difference assessed by paired *t*-test and agreement assessed by intraclass correlation

Nutrient	WFR	Mod 24-HR	Mean difference (95% CI)	Intraclass correlation (95% CI)
Energy (kcal)	379 (214)	429 (211)	50 (28–73)	0.73 (0.67–0.80)
Protein (g)	8.3 (5.1)	9.8 (3.9)	1.6 (1.0–2.2)	0.69 (0.61–0.77)
Fat (g)	5.3 (5.3)	7.1 (6.3)	1.8 (1.1–2.5)	0.65 (0.56–0.74)
Iron (mg)	2.3 (1.2)	3.2 (1.5)	0.8 (0.6–1.0)	0.42 (0.30–0.55)
Zinc (mg)	2.5 (3.2)	3.2 (2.3)	0.7 (0.3–1.2)	0.42 (0.29–0.54)
Vitamin A (μg)	134 (270)	179 (300)	45 (20–69)	0.83 (0.78–0.88)

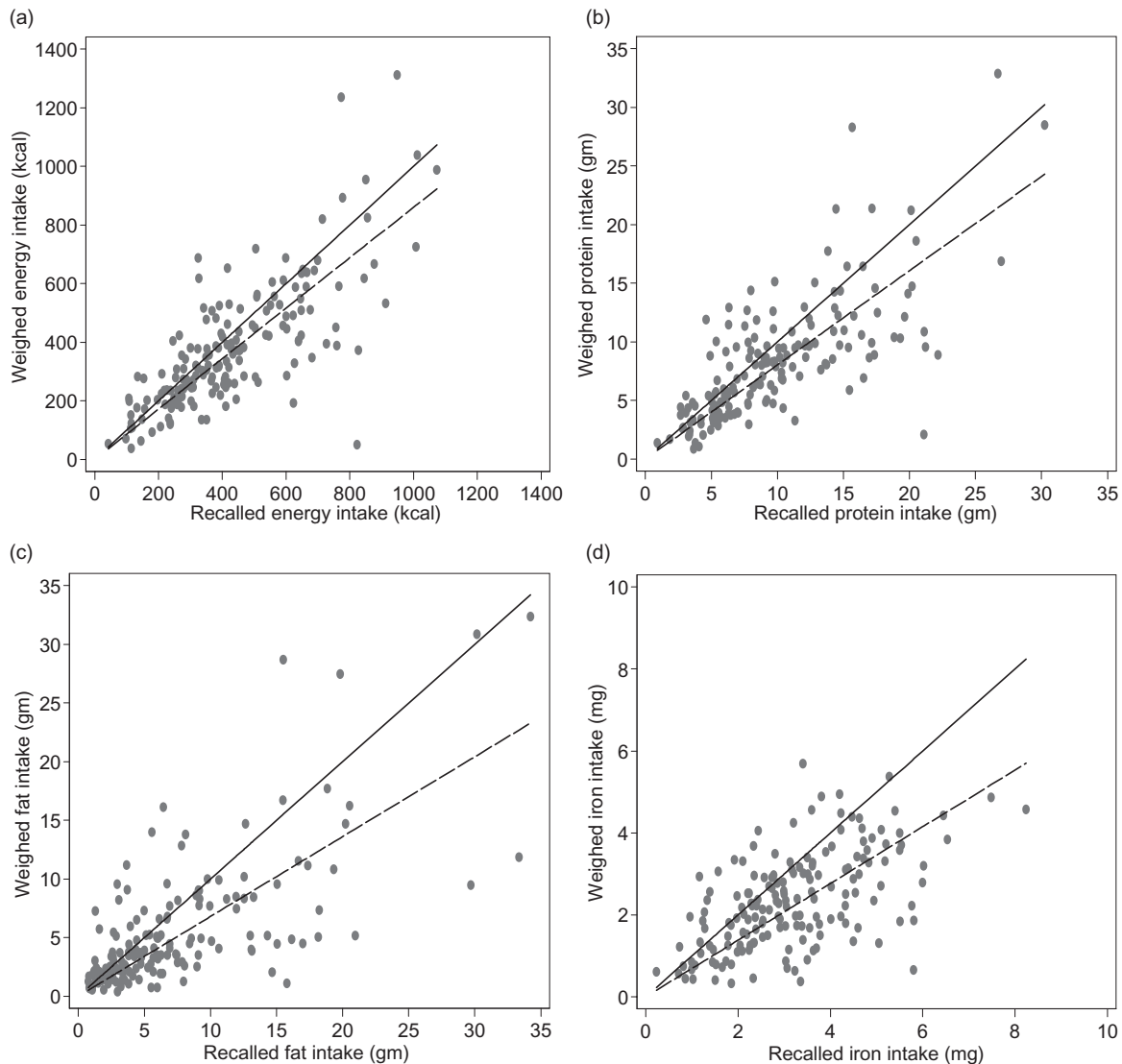CI, confidence interval; SD, standard deviation.

Fig. 1. Scatter-plots of weighed food records vs. modified 24-hour recalls, with 45-degree line (solid) and regression-through-the-origin line (broken). (a) Weighed vs. recalled energy intake (kcal), (b) weighed vs. recalled protein intake (g), (c) weighed vs. recalled fat intake (g), (d) weighed vs. recalled iron intake (mg).

were similar for WFR and mod 24-HR (each $P > 0.10$). The predicted values also gave similar estimates (each $P > 0.10$). However, according to the mod 24-HR, the prevalence of zinc intake below RNI was 56.8%, which was significantly ($P < 0.001$) lower than that estimated by WFR (78.1%). The estimated prevalence based on the predicted values was 76.9%, comparable to that of WFR ($P = 0.746$). Using 50% recommended daily intake as a yardstick of insuffi-

cient intake, the mod 24-HR underestimated the prevalence for each nutrient (each $P < 0.001$) and the prediction method gave prevalence estimates similar to those of WFR (each $P > 0.10$).

Discussion

Some previous studies suggested that the mod 24-HR method gives higher estimates for nutrient intake
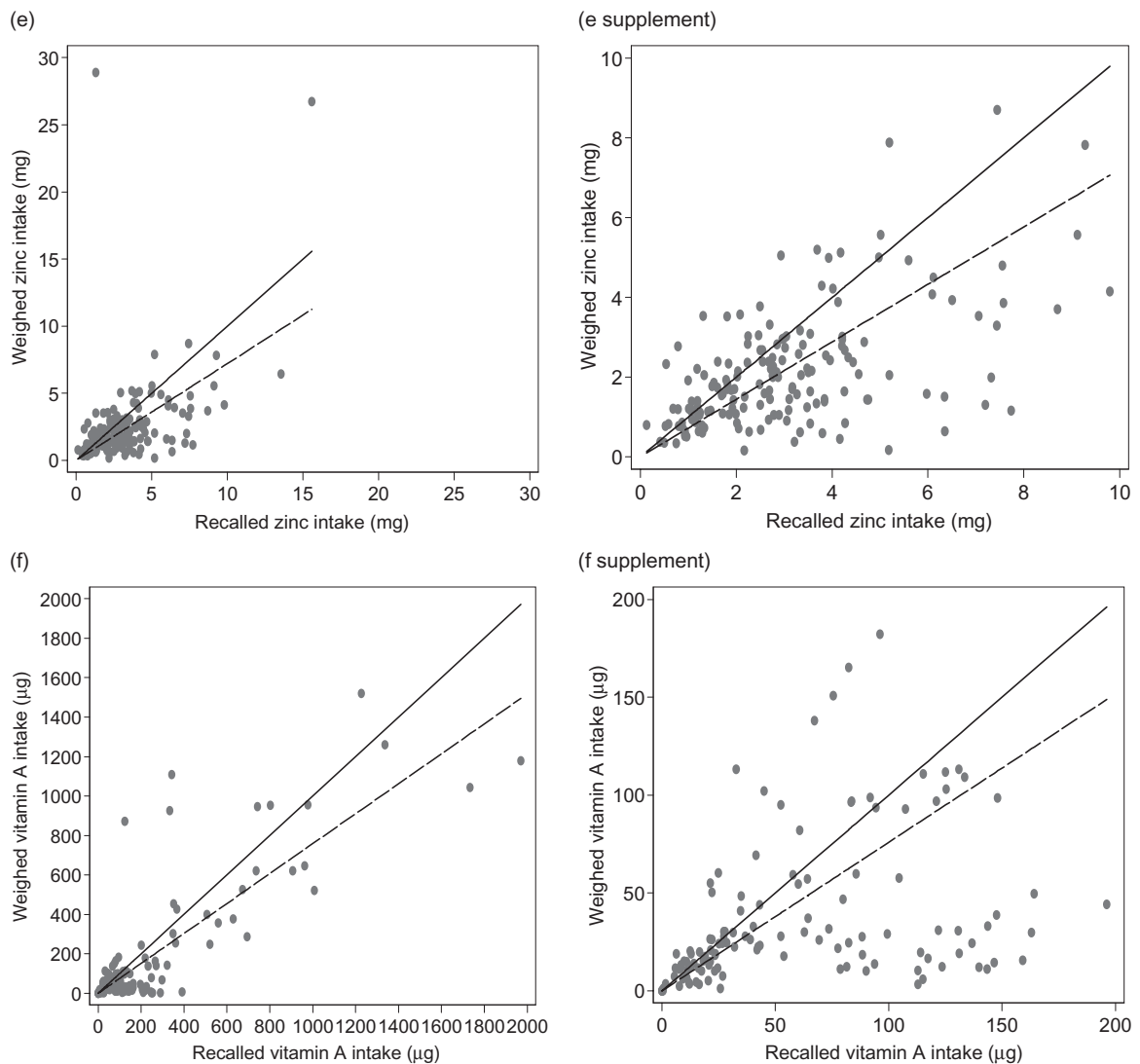
Fig. 1. (e) weighed vs. recalled zinc intake (mg), (e supplement) weighed vs. recalled zinc intake (mg). Zoom to subgroup with mg < 10, (f) weighed vs. recalled vitamin A intake (μg) and (f supplement) weighed vs. recalled vitamin A intake (μg). Zoom to subgroup with μg < 200.

values than the food weighing method (Ferguson *et al*. 1989, 1995; Fisher *et al*. 2008). The 24-HR method had also been found to overestimate energy and nutrient intake as compared with result of chemical analysis of duplicate portions (Horst *et al*. 1988). However, some other studies reported that 24-HRs resulted in underestimation of energy intake (Willett 1998; Jonnalagadda *et al*. 2000) or unbiased though imprecise estimates of energy and macronutrient

intakes (Dop *et al*. 1994). There is no conclusive explanation of how bias may arise. One possible explanation is social desirability or wishful thinking. For example, higher body fat content and relative weight were associated with under-reporting (Fisher *et al*. 2000), whereas low infant weight-for-length was found to be associated with over-reporting in a telephone-administered 24-HR (Fisher *et al*. 2008). Such patterns may reflect what people think they

Table 3. Ratios of weighed food record vs. modified 24-hour recall values estimated from regression-through-the-origin models, and comparison of predicted vs. observed weighed food records*

Nutrient	Ratio	(95% CI)	PRESS R-squared	Mean predicted value	P-value**	Intraclass correlation
Energy (kcal)	0.86	(0.82–0.91)	0.89	370	0.399	0.74 (0.67–0.81)
Proteins (g)	0.80	(0.75–0.85)	0.86	7.9	0.202	0.71 (0.63–0.78)
Fat (g)	0.68	(0.62–0.74)	0.72	4.8	0.102	0.66 (0.57–0.75)
Iron (mg)	0.69	(0.61–0.74)	0.85	2.2	0.060	0.56 (0.46–0.66)
Zinc (mg)	0.72	(0.61–0.83)	0.48	2.3	0.466	0.35 (0.22–0.48)
Vitamin A (μg)	0.76	(0.70–0.82)	0.76	136	0.873	0.83 (0.78–0.87)

CI, confidence interval; PRESS, prediction error sum of squares; *All predicted values are based on leave-out-one cross-validation; **P-value of paired *t*-test comparing predicted values and weighed food records.

Table 4. Estimates of prevalence of intake below recommended nutrient intake* (RNI) and intake below 50% of RNI according to weighed food records (WFR), modified 24-hour recalls (mod 24-HR) and predicted values**

Nutrient	WFR	Mod 24-HR	Predicted
RNI			
Iron	100%	98.8%	100%
Zinc	78.1%	56.8%***	76.9%
Vitamin A	87.0%	90.0%	91.1%
50% of RNI			
Iron	83.4%	62.7%***	88.2%
Zinc	39.6%	23.7%***	36.7%
Vitamin A	85.2%	75.7%***	82.8%

*7 mg for iron, 3 mg for zinc and 400 μg for vitamin A; **Predicted using leave-out-one cross validation and regression-through-the-origin models; ***McNemar test against WFR, $P < 0.001$; all other comparisons had $P > 0.10$.

should be or are supposed to be doing. Given these uncertainties and the possibility of cultural differences, local validation studies are important.

It is well known that the 24-HR has several limitations that were taken into consideration in this study. To help the mothers or guardians to remember what their children ate the previous 24 h, local food picture calendars were used. Standard measuring instruments were used to help in estimating food portion sizes.

In this study, the paired *t*-test and intraclass correlation were used to compare the nutrient intake values of the two methods (Table 2). The same trend was observed for each of the nutrients. We developed an adjustment model for the target nutrients to help

the mod 24-HR give better estimates of intake specifically with our target group (children under 2 years) living in a rural setting.

When the food intake values from both methods were shown graphically, the data points tended to lie below the 45-degree line, indicating that the recall method overestimated the WFR values. However, the scatter-plots show that the lower range of the data values pointed to the origin (0,0). We also observed that the difference between the two series of values appeared to increase as the nutritional intake values increased, but the ratio of the two values was roughly constant over the range of data. These justified the use of the regression-through-the-origin model with the formula $Y = bX$. This formula can be used to correct the recall estimate. For example, the WFR-to-recall ratio estimated by the model was 0.86. Each energy intake value based on the recall method can be multiplied by this ratio to approximate the WFR value. There were some extreme intake values for zinc (Fig. 1e) which originated from gourd (a type of pumpkin with high zinc value) intake estimates in the weighing method. This, we believe, comes from the 'super food effect', meaning that there are foods with high-nutrient density which are not valued by the respondents.

This study examined the estimation of energy and nutrient intakes in children in a rural setting and proposed an adjustment of values collected by mod 24-HR. The adjustment gives useful prediction of the true mean intake. Without the adjustment, the mean intakes would be overestimated by from about one-

seventh (energy) to about one-third (fat and iron). These differences seem practically substantial. Furthermore, the use of the mod 24-HR hugely underestimated the prevalence of insufficient nutritional intakes according to RNI values. The use of the adjusted prediction agreed well with the WFR in estimating the prevalence. Nevertheless, we caution against its use at the individual level: the difference between the true individual values and the predicted mean can be considerable. The adjustment method did not improve the intraclass correlation coefficient with respect to agreement with the WFR values, suggesting that the adjustment shifted the central position of the distribution towards the true position but did not reduce the disagreement with WFR at the individual level.

No dietary method is accurate for all age groups in different settings. The results presented here are for 15-month-old children in a rural African setting. The adjustment algorithms need to be evaluated for other age groups and settings to determine its validity. Growth stunting and malnutrition are common among infants and children aged below 24 months in developing countries (Maleta *et al.* 2003b; Phuka *et al.* 2008). The assessment of nutritional intakes and the prevalence of insufficient nutritional intakes are important for understanding the changing epidemiology and the impact of complementary feeding programmes. More accurate methods such as weighing food are difficult to implement on a large scale. The present findings will facilitate the conduct of epidemiological and clinical trial research in the setting of developing countries.

A limitation of the present study is that it assessed only one day per infant. It is possible that direct observation and food weighing may sensitize caretakers to their food preparation and feeding practice and therefore influence the subsequent recalls. We minimized this possibility by training the observers that included instructions on not revealing to the participants how much the food weighed and not commenting on their food preparation or consumption. Nevertheless, this potential influence cannot be totally ruled out. It is desirable that the present findings be verified by using other more resource-intensive research methodology, e.g. applying the two methods on two different sets of randomly selected days for each infant. Moreover, the

present study cannot be taken as demonstrating the level of nutritional intakes of individual Malawian infants as it was based on only one day's intake instead of a representative sample of days. However, the purpose of this work was solely to assess the validity of the mod 24-HR method, which is not affected by the use of only one day per participant. Another limitation is that we only applied two methods, without other measures such as doubly labelled water technique to corroborate the findings.

To conclude, the mod 24-HR method gives higher estimates than the WFR when assessing energy and nutrient intake values in rural Malawian infants aged 15 months. However, adjustment algorithms may be used to estimate the mean nutritional intake in a target group and the prevalence of insufficient intake as defined by RNI values.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Contributions

All authors except YBC designed the trial, PA wrote the protocol and designed funding and HMK raised funding for herself. JCP and CMT were responsible for data collection, YBC designed the details of statistical analysis and HMK and CMT wrote the first

draft of the manuscript under the supervision of PA, YBC and JCP. All authors commented on the analysis and participated in writing of the manuscript.

References

Donner A. (1986) A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model. *International Statistical Review* **54**, 67–82.

Dop M.C., Milan C.H., Milan C. & N'Diaye A.M. (1994) The 24-hr recall for Senegalese weanlings: a validation exercise. *European Journal of Clinical Nutrition* **48**, 643–653.

Eisenhauer J.G. (2003) Regression through the Origin. *Teaching Statistics* **25**, 76–80.

FAO/WHO (2004) *Human Vitamin and Mineral Requirements: Report of a Joint FAO/WHO Expert Consultation, Bangkok, Thailand*. Food and Agriculture Organisation: Rome.

Ferguson E.L., Gibson R.S., Ounpuu S. & Sabry J.H. (1989) The validity of the 24-hour recall for estimating the energy and selected nutrient intakes of a group of rural Malawian pre-school children. *Ecology of Food and Nutrition* **23**, 273–285.

Ferguson E.L., Gadowsky S.L., Huddle J.M., Cullinan T.R., Lehrfeld J. & Gibson R.S. (1995) An interactive 24 h recall technique for assessing the adequacy of trace mineral intakes of rural Malawian women, its advantages and limitations. *European Journal of Clinical Nutrition* **49**, 565–578.

Fisher J.O., Butte N.F., Mendoza P.M., Wilson T.A., Hodges E.A., Reidy K.C. *et al.* (2008) Overestimation of infant and toddler energy intake by 24-h recall compared with weighed food records. *American Journal of Clinical Nutrition* **88**, 407–415.

Fisher J.O., Johnson R.K., Lindquist C., Birch L.L. & Goran M.I. (2000) Influence of body composition on the accuracy of reported energy intake in children. *Obesity Research* **8**, 597–603.

Galpin L., Thakwalakwa C., Phuka J., Ashorn P., Maleta K., Wong W. *et al.* (2007) Breast milk intake is not reduced more by the introduction of energy dense complementary food than by typical infant porridge. *Journal of Nutrition* **137**, 1828–1833.

Gibson R.S. & Ferguson E.L. (2008) *An Interactive 24-hour Recall for Assessing the Adequacy of Iron and Zinc Intakes in Developing Countries. HarvestPlus Technical Monograph*. International Life Sciences Institute: Washington, DC.

Horst C.H., Obermann-De Boer G.L. & Kromhout D. (1988) Validity of the 24-hour recall method in infancy: the Leiden Pre-School Children Study. *International Journal of Epidemiology* **17**, 217–221.

Humphreys E., Misiri A. & Muula S. (2005) Under-reporting of gravidity in a rural Malawian population. *African Health Science* **5**, 335–337.

Jonnalagadda S.S., Mitchell D.C., Smiciklas-Wright H., Meaker K.B., Van Heel N., Karmally W. *et al.* (2000) Accuracy of energy intake data estimated by a multiple-pass, 24-hour dietary recall technique. *Journal of the American Dietetic Association* **100**, 303–308.

Kuczmarski R.J., Grummer-Strawn L.M., Flegal K.M., Guo S.S., Wei R., Mei Z. *et al.* (2000) CDC growth charts for the United States: methods and development. National Center for Health Statistics. *Vital Health Statistics* **11** (246), 1–190.

Maleta K., Virtanen S.M., Espo M., Kulmala T. & Ashorn P. (2003a) Childhood malnutrition and its predictors in rural Malawi. *Paediatrics and Perinatal Epidemiology* **17**, 384–390.

Maleta K., Virtanen S., Espo M., Kulmala T. & Ashorn P. (2003b) Timing of growth faltering in rural Malawi. *Archives of Diseases in Childhood* **88**, 574–578.

Montgomery D.C., Peck E.A. & Vining G.G. (2001) *Introduction to Linear Regression Analysis*. Wiley: New York.

Ndekha M., Kulmala T., Vaahtera M., Cullinan T., Salin M.-L. & Ashorn P. (2000) Seasonal variation in the dietary sources of energy for pregnant women in Lungwena, rural Malawi. *Ecology of Food and Nutrition* **38**, 605–622.

Phuka J.C., Maleta K., Thakwalakwa C., Cheung Y. & Manary M. (2008) Complementary feeding with fortified spread and incidence of stunting in 6- to18-month-old rural Malawians. *Archives of Pediatric and Adolescence Medicine* **162**, 619–626.

Willett W. (1998) *Nutritional Epidemiology*, 2nd edn, Oxford University Press: Oxford.