Public Health Nutrition: 10(11), 1247–1256 doi: 10.1017/S136898000768714X

Conventional analyses of data from dietary validation studies may misestimate reporting accuracy: illustration from a study of the effect of interview modality on children's reporting accuracy

Albert F Smith[1,*], Suzanne Domel Baxter[2], James W Hardin[3] and Michele D Nichols[4]

[1]Department of Psychology, Cleveland State University, 2121 Euclid Avenue, Cleveland, OH 44115, USA: [2]Institute for Families in Society, University of South Carolina, Columbia, SC, USA: [3]Center for Health Services and Policy Research & Department of Epidemiology and Biostatistics, University of South Carolina, Columbia, SC, USA: [4]Center for Research in Nutrition and Health Disparities, University of South Carolina, Columbia, SC, USA

Submitted 30 March 2006: Accepted 4 January 2007: First published online 19 March 2007

Abstract

Objective: To compare two approaches to analysing energy- and nutrient-converted data from dietary validation (and relative validation) studies – conventional analyses, in which the accuracy of reported items is not ascertained, and reporting-error-sensitive analyses, in which reported items are classified as matches (items actually eaten) or intrusions (items not actually eaten), and reported amounts are classified as corresponding or overreported.

Design: Subjects were observed eating school breakfast and lunch, and interviewed that evening about that day's intake. For conventional analyses, reference and reported information were converted to energy and macronutrients; then *t*-tests, correlation coefficients and report rates (reported/reference) were calculated. For reporting error-sensitive analyses, reported items were classified as matches or intrusions, reported amounts were classified as corresponding or overreported, and correspondence rates (corresponding amount/reference amount) and inflation ratios (overreported amount/reference amount) were calculated.

Subjects: Sixty-nine fourth-grade children (35 girls) from 10 elementary schools in Georgia (USA).

Results: For energy and each macronutrient, conventional analyses found that reported amounts were significantly less than reference amounts (every $P < 0.021$; paired *t*-tests); correlations between reported and reference amounts exceeded 0.52 (every $P < 0.001$); and median report rates ranged from 76% to 95%. Analyses sensitive to reporting errors found median correspondence rates between 67% and 79%, and that median inflation ratios, which ranged from 7% to 17%, differed significantly from 0 (every $P < 0.0001$; sign tests).

Conclusions: Conventional analyses of energy and nutrient data from dietary reporting validation (and relative validation) studies may overestimate accuracy and mask the complexity of dietary reporting error.

Keywords
Validation
Relative validation
Epidemiological methods
Dietary assessment

Validation and relative validation studies of dietary reporting methods are conducted to evaluate the extent to which those methods elicit accurate reports; this involves comparing reports obtained by such methods as 24-hour recalls and food-frequency questionnaires (FFQs) with reference information obtained by such methods as direct observation, duplicate portion collection or food records. The reported and reference information each consist of a set of food items and their respective amounts. (Table 1 shows terms used in this article and their definitions.)

Assessment of reporting accuracy for a group of participants often begins with converting both reported and reference information to energy and nutrients. After cumulating, for each participant, the energy and nutrients within each set, analyses involve several descriptive measures and inferential tests (see e.g. references 1–14). First, paired *t*-tests may be used to compare the mean of the participant-specific differences between reported and reference energy and nutrients with 0. Second, the correlation, over participants, between reported and reference energy and nutrients may be calculated, with correlations near 1 presumed to indicate higher reporting accuracy. Third, the ratio of reported to reference energy and nutrients, multiplied

**Corresponding author:* Email a.f.smith@csuohio.edu

Table 1 Terminology

Correspondence rate	The percentage of the reference amount to which the reported amount corresponds; it is the ratio of the corresponding amount of energy or a nutrient to the reference amount of energy or that nutrient, multiplied by 100. The correspondence rate is a genuine measure of accuracy with values between 0 and 100%, with higher values indicating higher accuracy.
Corresponding amount	The amount of a match that overlaps between the reported amount and the reference amount, measured either in servings or in servings converted to energy or nutrients.
Inflation ratio	The ratio of the overreported amount of energy or a nutrient to the reference amount of energy or that nutrient, multiplied by 100. The inflation ratio for energy or a nutrient quantifies inaccurate reporting. It has a lower bound of 0%, but no upper bound, because there is no limit on what may be reported. Lower ratios reflect better reporting accuracy.
Intrusion	A food item reported eaten by the participant that is not in the reference set.
Match	A food item in the reference set that is reported eaten by the participant.
Omission	A food item in the reference set not reported eaten by the participant.
Overreported amount	The amount by which the reported amount of a match exceeds the reference amount, or the amount of an intrusion, measured either in servings or in servings converted to energy or nutrients.
Reference amount	The amount of a food item in the reference set, measured either in servings or in servings converted to energy or nutrients.
Reference information	The set of food items (reference set) and their amounts that were actually eaten by a study participant.
Report rate	The ratio of the reported to reference amounts of energy or a nutrient, multiplied by 100. Conventionally, values close to 100% are interpreted as indicating high accuracy. For energy or a nutrient, the report rate of a participant is the sum of the correspondence rate and the inflation ratio for energy or that nutrient.
Reported amount	The amount of a reported food item, measured either in servings or in servings converted to energy or nutrients.
Reported information	The set of food items and their amounts reported by a study participant.
Unreported amount	The amount by which the reported amount of a match falls short of the reference amount, or the amount of an omission, measured either in servings or in servings converted to energy or nutrients.

by 100, may be calculated. We call this the report rate; others have labelled it 'percentage recalled' and 'percentage over (or under) reported' (see e.g. references 3, 5 and 6) Values close to 100%, found when reported and reference energy and nutrients are approximately equal, are presumed to represent high reporting accuracy[12–14]. We call these analyses, taken together, the 'conventional approaches' to evaluating reporting accuracy.

These conventional approaches are generally indifferent to whether food items and amounts are reported correctly – reported information is converted to energy and nutrients regardless of whether items were actually eaten. In this article, we demonstrate that conventional analyses of data from dietary reporting validation studies may mask the complexity of dietary reporting error and overestimate reporting accuracy.

Consider, for example, mean differences between reported and reference information. For any dietary measure, the mean difference would be 0 in a validation study in which all participants reported items and amounts with perfect accuracy. However, the mean difference could also be 0 if, on average, participants reported items and amounts that they did not eat that offset their failures to report items and amounts that they did eat. Suppose that an individual was observed eating a serving of hash browns, and subsequently reported eating a serving of apple sauce. Converted to energy and macronutrients, the individual was observed eating 100 kcal, 1 g of protein, 12 g of carbohydrate and 5 g of fat; and reported eating 96 kcal, 0.2 g of protein, 25.4 g of carbohydrate and 0.2 g of fat. (Composition information for a 57 g serving of hash browns is from the product label, and for a 120 ml serving of sweetened apple sauce is from the Nutrition Data System for Research [NDSR, version 4.03; Nutrition Coordinating Center, University of Minnesota] database.) The report rates for energy (96%), protein (20%), carbohydrate (212%) and fat (4%) would be interpreted as essentially perfect reporting of energy, underreporting of protein and fat, and overreporting of carbohydrate. However, all reported energy and macronutrients were from a reported item that was, in fact, not eaten by that individual. The conventional approaches, by their indifference to whether reported items and amounts were actually eaten, may misrepresent dietary reporting accuracy.

We introduce some terminology to facilitate our discussion of this problem and of our recommendations concerning how to address it. We then present an illustration using data from a validation study of dietary recall, and so frame the discussion in terms of dietary recall. However, the concerns also pertain to relative validation of FFQs.

In dietary recall, the reported information is a set of item–amount pairs (e.g. [apple, 1 medium]; [whole milk, 1 serving]) reported by a participant. The reference information is also a set of item–amount pairs.

Reporting accuracy may be evaluated by assessing the congruence between the reported and reference sets of

https://doi.org/10.1017/S136898000768714X Published online by Cambridge University Press

information, which involves partitioning the sets as illustrated in Fig. 1 and discussed here.

For any participant, reported items are partitioned into matches and intrusions, and reference items are partitioned into (the same) matches and omissions[15,16]. A *match* is an item reported eaten that is in the reference set. An *intrusion* is an item reported eaten that is not in the reference set. An *omission* is an item in the reference set that is not reported eaten.

Amounts – whether of servings or of such quantities as kilocalories and grams of protein – are classified as *corresponding*, *overreported* or *unreported*. For a match, the reported amount equals, exceeds or falls short of the reference amount. If the reported amount equals the reference amount, the reported amount is a corresponding amount. If the reported amount exceeds the reference amount, the amount by which it does so is an overreported amount; if the reported amount falls short of the reference amount, the amount by which it falls short is an unreported amount. In each case, the overlap between the reported and reference amounts is a corresponding amount. (For example, if an individual was observed consuming a 1/4 serving of milk, and reported consuming a whole serving, the reported serving consists of a corresponding 1/4 serving and an overreported 3/4 serving. If an individual was observed consuming 1 serving of milk and reported consuming 1/4 serving, the reported 1/4 serving is corresponding, and 3/4 serving is unreported.) For an omission, the entire amount is unreported. For an intrusion, the entire amount is overreported.

From this, it follows that:

$$\text{Reported amount} = \text{Corresponding amount} + \text{Overreported amount}$$

and

$$\text{Reference amount} = \text{Corresponding amount} + \text{Unreported amount},$$

in which each term is an arithmetic quantity and + represents addition. Letting *rep*, *ref*, *c* and *o* represent reported, reference, corresponding and overreported amounts, respectively, we can abbreviate the former expression to $rep = c + o$.

The *report rate* of energy or of a nutrient, a conventional measure of reporting accuracy, is the reported amount divided by the reference amount, multiplied by 100. A report rate has a lower bound of 0%, which would indicate that nothing was reported, and no upper bound, because there is no limit on what an individual might report. Algebraic expression of the report rate, as

$$(rep/ref) \times 100 = [(c + o)/ref] \times 100$$
$$= [(c/ref) \times 100] + [(o/ref) \times 100],$$

shows that it consists of two parts. We call $[(c/ref) \times 100]$ the *correspondence rate*: it is the percentage of the reference amount to which the reported amount corresponds. It is a genuine measure of accuracy with a value between 0%, which would indicate that no reference

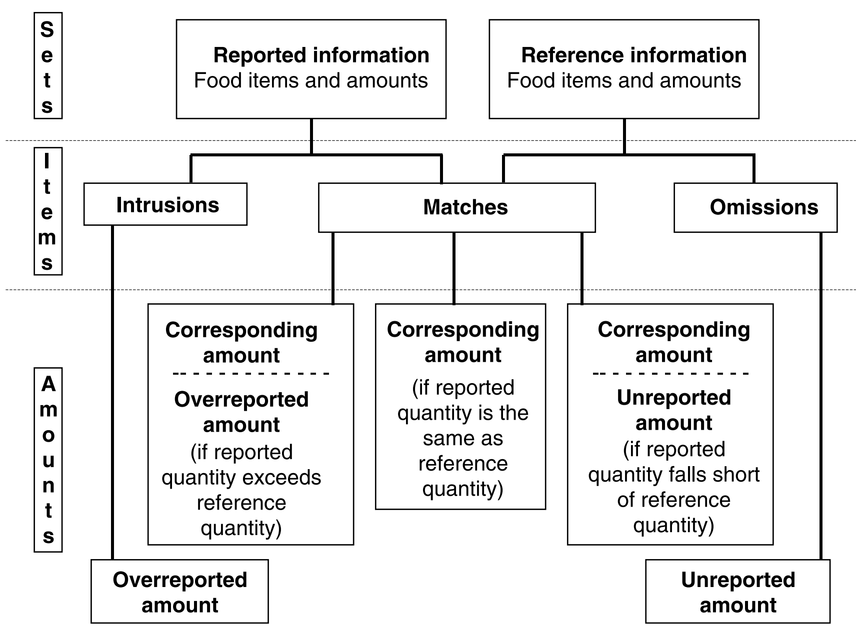

Fig. 1 Dietary reporting accuracy in a validation (or relative validation) study should be assessed by evaluating the congruence between reported and reference information. In this approach, reported and reference items are classified as intrusions (reported items not in the reference set), matches (reported items in the reference set) and omissions (reference set items not reported). Amounts – of servings or of energy and nutrients – are classified as overreported (reported but not in the reference information), corresponding (reported and in the reference information) and unreported (in the reference information, but not reported)

https://doi.org/10.1017/S136898000768714X Published online by Cambridge University Press

item had been reported, and 100%, which would indicate that all reference items and their amounts had been reported correctly. We call [(*o*/*ref*) × 100] the *inflation ratio*: it is a non-negative augmentation to correctly reported information based on inaccurate reporting; it has no upper bound, because there is no limit on what can be reported.

Although the report rate is treated conventionally as a measure of accuracy, it includes a component based on reporting error: because the inflation ratio is non-negative, any overreporting – whether it is reporting intrusions or amounts of matches that exceed reference values – inflates the report rate.

Consider again the individual who was observed eating one serving of hash browns and reported eating one serving of apple sauce, and assume that this individual ate only hash browns and reported only apple sauce. This individual's report rates for energy, protein, carbohydrate and fat, are 96%, 20%, 212% and 4%, respectively, but the correspondence rate for each is 0%, because none of what was reported was from a match. The inflation ratios for energy and each macronutrient are equal to the respective report rates, because all of what was reported was from an intrusion.

In this article, we demonstrate that conventional analyses of data from dietary reporting validation studies that convert reported and reference information to energy and nutrients, and that do not distinguish between reported matches and intrusions or between corresponding and overreported amounts, may mask the complexity of dietary reporting error and overestimate reporting accuracy. We use data from a study conducted to determine whether children's dietary recall accuracy depends on interview modality (in-person or by telephone)[17].

Methods

Our objective was to compare conclusions that might be drawn from different analytical approaches to dietary validation data, so the particularities of the original study are of minimal interest. We described the sample and methodology elsewhere[17], so summarise them here.

The Institutional Review Board of the Medical College of Georgia approved the procedures for recruiting participants and collecting data, and that of the University of South Carolina approved using the data for the analyses described in this article.

Participants

During the 2001–2002 school year, all fourth graders ($n = 799$) from 10 public schools in one school district were invited to participate. From the 451 children who provided written child assent and parental consent, children were selected randomly, subject to the constraint that there be roughly equal numbers in each of four race (black, white) × sex groups. Sixty-nine children (35 girls) were interviewed; their mean age at the time of their interviews was 10.15 years (standard deviation = 0.55 years).

Procedures

To compile reference information, one of three dietitians observed each child eating school breakfast and school lunch on a school day. Following procedures used in previous studies[18,19], observers recorded items eaten and their amounts in servings. Only children who obtained breakfast and lunch at school were observed, because it is difficult to identify contents of meals brought from home[20]. Children were observed for their entire meal periods so that any food trading could be noted[7,21,22]. Although children knew in general that they were being observed, they did not know who, if anyone, would be interviewed or whether any interview would be in person or by telephone. Weekly assessments indicated that interobserver reliability was satisfactory[17,23].

Within each race × sex group, the children were divided randomly between the in-person and telephone interview modalities. One of two dietitians interviewed each child after 6.30 p.m. on the day that the child had been observed; the interviewer had not observed either of the child's meals. During the interview, the child was to describe and quantify everything consumed that day; this was that child's reported information. Interviewers followed a protocol modelled on the NDSR computerised interview, but wrote information on forms. Interviews were audiorecorded. Interview quality was monitored by daily assessment of a randomly selected interview on a structured list of criteria, including agreement among the interview form, the audiotape and the typed transcript[24]; these assessments indicated that interviewers adhered to protocol[17].

Construction of analytic variables

Reference information was available only for school breakfast and school lunch, so analyses were restricted to reports of those meals.

We prepared variables for two analytic approaches; each involved converting reported and reference information to energy and macronutrients. (1) For conventional analyses, we ignored whether reported items were matches or intrusions (and therefore whether reported amounts were corresponding or overreported). (2) For analyses sensitive to reporting errors, we classified reported items as matches or intrusions, and reported amounts as corresponding or overreported.

We published results concerning reporting accuracy at the food-item level elsewhere[17]. These showed that none of omission rate (percentage of observed items not reported; overall mean = 33%), intrusion rate (percentage of reported items not observed; overall mean = 17%) and total inaccuracy (sum of itemwise absolute difference between observed and reported numbers of servings;

overall mean $=4.4$ servings) depended on whether interviews were conducted in person or by telephone[17].

As in our previous studies, reported meals were treated as reports about school meals only if the child reported eating that meal at school; referred to breakfast as 'school breakfast' or 'breakfast' and to lunch as 'school lunch' or 'lunch'; and reported the mealtime to within an hour of the observed mealtime[17–19].

Numeric values of servings were assigned, as follows, to the qualitative labels used during observations and interviews: none (0.00), taste (0.10), little bit (0.25), half (0.50), most (0.75), all (1.00) or the actual number of servings if more than one was observed or reported[17–19]. For each reference item and each reported item, we obtained per serving information about energy and macronutrient content from the NDSR database or, for items not in that database, from product information or recipes provided by the school district's nutrition programme.

Conventional variables

Reference and reported amounts of energy and macronutrients were calculated from reference and reported information, respectively, and food composition information by multiplying the quantified servings of each item by the per-serving energy and macronutrient values. For each child, for energy and each macronutrient, these values were summed across the reference (i.e. observed) items for the two school meals, and also across the items reported for the two school meals. Table 2 illustrates, for one child, how we generated values of reported kilocalories and reference kilocalories – the variables required for conventional analyses of energy reporting accuracy. For each child, for energy and each macronutrient, we calculated the report rate from values of these variables.

Variables sensitive to reporting errors

For each child, following classification of each reported and reference item as a match, an intrusion or an omission, the constituent energy and macronutrients of each item were classified as corresponding, overreported or unreported.

An item reported by a child was treated as matching a reference item unless the reported item clearly did not describe an item that the child was observed eating; this might overestimate accuracy. For example, any reported white milk (e.g. skimmed, whole) was matched to any observed white milk, and any reported pizza (e.g. cheese, pepperoni) was matched to any observed pizza. However, reported fruit juices (e.g. apple, orange), vegetables (e.g. green beans, broccoli) and milk flavours (e.g. chocolate, strawberry) that differed from what was observed were not classified as matches[17–19].

Each corresponding, overreported and unreported number of servings was multiplied by the appropriate per-serving values of kilocalories and grams of macronutrients to obtain corresponding, overreported and unreported amounts of energy and macronutrients. For each child, summing the values for each category across items yielded total kilocalories and total grams of each macronutrient that were corresponding, overreported and unreported. Table 2 illustrates, for energy for one child, how we generated these variables. For each child, for energy and each macronutrient, we used appropriate values of these variables to calculate the correspondence rate and the inflation ratio.

Analyses

Analyses were conducted with SAS (version 8.2; SAS Institute).

Effect of interview modality

Whether interviews were in person or by telephone was irrelevant to the concerns of this article but, before conducting analyses for this article, we compared the interview-modality groups on correspondence rates and report rates for energy and each macronutrient (data not shown). Using Wilcoxon rank-sum tests, we found no significant effect of interview modality on either measure for energy or any macronutrient (every $P > 0.14$), so, for subsequent analyses, we combined data from the two groups.

Conventional variables

We used *t*-tests to compare mean differences between reported and reference values of energy and of each macronutrient with 0. We calculated Pearson correlations, over subjects, between reported and reference energy and macronutrient values. (Neither the *t*-tests nor the tests of correlation coefficients were independent, because the energy and macronutrient variables were not independent – all variables were calculated from single sets of food items. However, treating such variables separately appears to be customary practice.) For each macronutrient, we also calculated correlations between energy-normalised reference values (grams observed/kilocalories observed) and energy-normalised reported values (grams reported/kilocalories reported).

Variables sensitive to reporting errors

Because the distributions of inflation ratios were right skewed, we used sign tests to compare inflation ratios for energy and for each macronutrient with 0.

Results

Conventional variables

The first two data columns of Table 3 show descriptive statistics for reported and reference amounts of energy and each macronutrient. For energy and every macronutrient, the mean reported amount was significantly less than the

Table 2 Classifications and computations used to assess accuracy of reported energy compared with reference energy for one child*

Food item†	Reported amount (servings)	Reference amount (servings)	kcal per serving‡	Reported kcal§	Reference kcal§	Overreported kcal from intrusions¶	Overreported kcal from matches‖	Corresponding kcal from matches**	Unreported kcal from matches††	Unreported kcal from omissions‡‡
Breakfast										
Chocolate milk	0.10	0.75	146	15	110	0	0	15	95	0
Biscuit§§	0.00	0.00	–	–	–	–	–	–	–	–
Sausage	0.00	1.00	80	0	80	0	0	0	0	80
Peaches	1.00	0.00	68	68	0	68	0	0	0	0
Grits	0.25	0.50	90	23	45	0	0	23	22	0
Lunch										
Honey mustard	0.00	1.50	29	0	44	0	0	0	0	44
Chocolate milk	1.50	0.75	146	219	110	0	109	110	0	0
Mixed fruit¶¶	–	0.00	–	–	–	–	–	–	–	–
Rice pilaf	0.10	0.10	69	7	7	0	0	7	0	0
Egg roll	0.00	1.00	250	0	250	0	0	0	0	250
Broccoli	0.50	0.00	15	8	0	8	0	0	0	0
Total				340	646	76	109	155	117	374

*Reference information was obtained by observation. This child was observed eating chocolate milk, sausage and grits at school breakfast, and honey mustard, chocolate milk, rice pilaf and egg roll at school lunch. This child reported eating chocolate milk, peaches and grits at school breakfast, and chocolate milk, rice pilaf and broccoli at school lunch.
† Every item that was reported eaten or observed eaten or both is listed.
‡ From the Nutrition Data System for Research database (NDSR, version 4.03; Nutrition Coordinating Center, University of Minnesota) or from product information or recipes obtained from the school district's nutrition programme.
§ Values are products of serving and kcal per serving information. Reported kcal are either corresponding kcal or are overreported, so that in each row, corresponding kcal and overreported kcal sum to reported kcal. Reference kcal are either corresponding or unreported, so that in each row, corresponding kcal and unreported kcal sum to reference kcal.
¶ Overreported kcal from intrusions = kcal from intrusions (i.e. food items reported eaten but not in reference set).
‖ Overreported kcal from matches = for matches for which reported kcal > reference kcal, absolute differences between reported kcal and reference kcal.
**Corresponding kcal from matches = for matches, overlap between reported kcal and reference kcal. (It is the minimum of reported kcal and reference kcal. Corresponding amounts are necessarily from matches.)
†† Unreported kcal from matches = for matches for which reported kcal < reference kcal, absolute differences between reported kcal and reference kcal.
‡‡ Unreported kcal from omissions = kcal from omissions (i.e. food items in reference set but not reported eaten).
§§ This item was observed on this child's tray but none was eaten. The child mentioned this item during the interview but reported that none of the item was eaten. This item was not included in calculating reporting accuracy.
¶¶ This item was observed on this child's tray but none was eaten and it was not reported in the interview. This item was not included in calculating reporting accuracy.

Table 3 Descriptive statistics for amounts of energy and macronutrients according to reported, reference, and five categories of amounts given as mean (standard deviation) ($n = 69$)

	Reported	Reference*	Overreported amount from intrusions	Overreported amount from matches	Corresponding amount from matches	Unreported amount from matches	Unreported amount from omissions
Kilocalories	645 (256)	789 (299)	57 (88)	89 (119)	499 (255)	72 (106)	219 (191)
Protein (g)	29 (11)	32 (12)	2 (4)	4 (6)	23 (12)	3 (5)	6 (7)
Carbohydrate (g)	95 (39)	116 (43)	10 (14)	14 (19)	70 (36)	11 (16)	34 (30)
Fat (g)	18 (9)	24 (12)	1 (3)	2 (3)	15 (9)	2 (4)	7 (7)

*Reference information was obtained by observation. See Fig. 1 for an illustration of the five categories of amounts.

mean reference amount (for energy, protein, carbohydrate and fat, paired $t(68)$s = 4.57, 2.37, 4.35 and 5.00, respectively; every $P < 0.021$). However, for energy and every macronutrient, the correlation, over children, between reported and reference amounts was statistically significant (for energy, protein, carbohydrate and fat, $r = 0.56$, 0.63, 0.52 and 0.61, respectively; every $P < 0.001$). Correlations between energy-normalised reported and reference values of protein, carbohydrate and fat were similar in magnitude (every $r > 0.48$). The first data column of Table 4 shows that for energy and the macronutrients, median report rates ranged from 76% to 95%.

Variables sensitive to reporting errors

The last five data columns of Table 3 show descriptive statistics for the decomposition of reported and reference amounts of energy and each macronutrient into the five categories of amounts shown in Fig. 1. These values clarify that corresponding amounts from matches constituted only part of reported amounts, and that overreported amounts and unreported amounts were not in balance.

The last two data columns of Table 4 show descriptive statistics for correspondence rates and inflation ratios for energy and each macronutrient. Median correspondence rates were between 67% and 79%, and median inflation ratios were between 7% and 17%. Not surprisingly, median correspondence rates were lower than median report rates. Although minimum inflation ratios were 0% for energy and every macronutrient, all first quartiles were non-zero, indicating that many children reported intrusions, overreported amounts of matches, or both. Sign tests showed, for energy and for every macronutrient, that inflation ratios significantly exceeded 0 (every $P < 0.0001$).

Discussion

Conventional analyses provided a mixed picture of children's reports: although t-tests showed that children underreported their dietary intake, reported and reference information was significantly associated. Further, from median report rates for energy, protein and carbohydrate that were close to 100%, one might conclude that children, on average, reported these accurately. Although report rates of 100% have been interpreted as indicating perfect reporting accuracy[12–14], given that report rates have no upper limit, and that their numerator quantities are not necessarily subsets of their denominator quantities, report rates have little logical connection to reporting accuracy.

Median energy and macronutrient correspondence rates for these children – from reported items that actually matched reference items – were lower, ranging from 67% to 79%. Reporting accuracy was worse than the conventional report rates suggested. The differences came from intrusions and overreported amounts of matches. These are quantified by inflation ratios, the median values of which ranged from 7% to 17%.

Table 4 Descriptive statistics for report rates, correspondence rates and inflation ratios for energy and macronutrients, given as median % (minimum %; first quartile %; third quartile %; maximum %) ($n = 69$)

	Report rate*	Correspondence rate†	Inflation ratio‡
Kilocalories	84 (0; 68; 100; 470)	67 (0; 50; 83; 100)	14 (0; 6; 26; 389)
Protein (g)	95 (0; 83; 107; 11 210)	79 (0; 51; 92; 100)	14 (0; 6; 27; 11 210)
Carbohydrate (g)	87 (0; 68; 100; 359)	67 (0; 46; 79; 100)	17 (0; 5; 28; 259)
Fat (g)	76 (0; 62; 98; 2107)	67 (0; 48; 86; 100)	7 (0; 3; 20; 2007)

*For an individual, report rate = (Reported amount/Reference amount) × 100. A report rate has a lower bound of 0% (which indicates nothing was reported), but no upper bound (because there is no limit on what an individual can report).
† For an individual, correspondence rate = (Corresponding amount/Reference amount) × 100; this is the proportion of the reference amount to which the reported amount corresponds, and is a genuine measure of accuracy. A correspondence rate is between 0% (which indicates that nothing reported eaten had been observed eaten) and 100% (which indicates that all reported items were observed, with the number of servings being reported correctly). Corresponding amounts are necessarily from matches.
‡ For an individual, inflation ratio = (Overreported amount/Reference amount) × 100. Note that the numerator quantity is not a subset of the denominator quantity. The inflation ratio is non-negative and can exceed 100%. Overreported amounts may come from matches or from intrusions.

https://doi.org/10.1017/S136898000768714X Published online by Cambridge University Press

Overreported amounts – from intrusions and from matches – were not balanced by unreported amounts. Conventional analyses of validation study data that do not distinguish between reported matches and intrusions or between corresponding and overreported amounts provide an overly simplified and possibly misleading picture of dietary reporting accuracy.

Mixed findings have been yielded by validation studies that have obtained reference information by observation and used conventional approaches to assess the dietary-reporting accuracy for energy and macronutrients of 7- to 13-year-old children who reported without help from parents[4,7–10,12–14]. Some studies have not found significant differences between reported and observed intake for energy[7,13,14] (13-year-olds[13]; Chinese, Hispanic[14]), protein[7,13,14] (8-year-olds[13]; Chinese, Filipino, Hispanic[14]), carbohydrate[7,14] (Chinese, Hispanic[14]) and fat[8,13,14] (13-year-olds[13]; Chinese, Hispanic[14]), which could be interpreted as high reporting accuracy. Others have found reported intake to be significantly greater than observed intake for energy[10,13,14] (8-year-olds[13]; Filipino[14]), carbohydrate (Filipino[14]) and fat[13,14] (8-year-olds[13]; Filipino[14]), which is conventionally interpreted as overreporting. Yet other studies have found reported intake to be significantly less than observed intake for energy[4,12,14] (Cambodian[14]), protein[4,13,14] (13-year-olds[13]; Cambodian[14]) and carbohydrate and fat[12,14] (Cambodian[14]), which is conventionally interpreted as underreporting. One study found that although reported energy intake significantly exceeded reference intake, the percentages of energy from protein, carbohydrate, and fat in reported and observed items did not differ significantly[9]. Without distinguishing between matches and intrusions, and between corresponding and overreported amounts of matches, it is difficult, if not impossible, to assess the accuracy of children's reports in these studies. Regardless of whether energy and macronutrient values from reported information were greater than, less than, or not significantly different from values from reference information, reports could include intrusions and overreported amounts.

The correlation, over individuals, between energy and macronutrient values converted from reported and reference information is another conventional metric of reporting accuracy[7–10,13,14]. Correlations close to 1 are interpreted conventionally as indicating high reporting accuracy. In the data described in this article, correlations between reported and reference energy and macronutrients ranged from 0.52 to 0.63 (every $P < 0.001$). These correlations are similar to those found in many validation studies of dietary reports by 8- to 13-year-old children who reported without help from parents and in which reference information was obtained by observation[7–10,13,14]. However, correlations give no information about actual reporting accuracy: although, over a set of individuals, reported energy and macronutrients might covary with reference energy and macronutrients, respectively, there is no necessary connection, within individuals, between information obtained by aggregating over reports, on the one hand, and information obtained by aggregating over observations, on the other.

Epidemiological interest may often be in nutrients, not foods, so that failing to distinguish intrusions from matches, or overreported amounts from corresponding amounts, might not seem problematic. However, there is increasing interest in analysis of dietary patterns and their relationship to health outcomes (see e.g. references 25–27) increasing appreciation that sources of nutrients are important and that nutrients interact (see e.g. references 28–31) and recognition that differential reporting accuracy for different foods on FFQs should be taken into account in applying measurement error correction methods[32]. These scientific trends suggest that assessment of validity should be concerned with reporting accuracy of food items and their quantities.

In addition, as shown in Table 3, overreported amounts do not necessarily balance unreported amounts. The degree to which these amounts are out of balance likely depends on specific aspects of the interview. For example, we have found, with both adults[16] and children[33], that certain reporting instructions lead research participants to report more intrusions than do other instructions without affecting reports of matches. In conventional analyses, this would be manifested as higher report rates with the former instructions than with the latter, whereas analyses of variables sensitive to reporting errors would show that correspondence rates would be the same with the two instructions. How incorrect reports of amounts are distributed across the categories illustrated in Fig. 1 is an aspect of reporting complexity that is lost in conventional analyses of validation-study data, but that would be revealed by analyses of variables sensitive to reporting errors.

This investigation has certain limitations inherited from the design of the original study: these include that only fourth-grade children from one school district were studied; that observation was restricted to children who obtained their breakfast and lunch from the school food service; and that analysis was restricted to the school meal portions of children's reports.

Conclusions

Conventionally, reported information and reference information in dietary validation (and relative validation) studies are converted to energy and nutrients; *t*-tests are used to examine differences, and correlations are used to describe relationships, between reported and reference energy and nutrient values; and report rates are calculated. But these approaches do not take into account that reports of food items include intrusions as well as matches, that reported amounts of matches may be incorrect, and that reported amounts of intrusions are necessarily incorrect.

Analyses that do not distinguish between matched and intruded food items, or between corresponding and overreported amounts, may provide a misleading picture of reporting accuracy.

It is often implicit – and sometimes explicit – in discussions of the results of validation studies that significant differences between reported and reference energy and nutrients indicate overreporting or underreporting of amounts, and that the absence of differences indicates excellent reporting accuracy. However, in the absence of knowledge of the correspondence between reported and reference information, it is impossible to estimate genuine reporting accuracy.

We suggest that conclusions about reporting accuracy from dietary validation (and relative validation) studies should not be drawn from energy and nutrient variables based on total reported and total reference information. Instead, we recommend analysis of energy and nutrient variables sensitive to reporting errors, constructed following classification of reported items as matches and intrusions, and of reported amounts as corresponding and overreported.

Acknowledgements

Sources of funding: This research was supported by grants 43-3AEM-2-80101 from the United States Department of Agriculture, Economic Research Service, Food Assistance and Nutrition Research Program, and R01 HL63189 from the National Heart, Lung, and Blood Institute of the National Institutes of Health. S.D.B. was Principal Investigator of both grants. A.F.S. was Principal Investigator of a subcontract of 43-3AEM-2-80101.

Conflict of interest declaration: No author has any conflict of interest.

Authorship responsibilities: Study concept–A.F.S., S.D.B.; acquisition of data, recruitment of schools and participants, data coding and entry, quality control–S.D.B.; statistical analysis and interpretation of data–A.F.S., S.D.B., J.W.H., M.D.N.; drafting and revising the manuscript–A.F.S., S.D.B., J.W.H., M.D.N.

Acknowledgements: Some of these results were presented at the Sixth International Conference on Dietary Assessment Methods (Copenhagen, 2006). We appreciate the cooperation of children, faculty and staff at the schools at which we collected data, and of the school district's nutrition programme and board of education. We thank Ram Chandran (Economic Research Service, USDA), Katherine Flegal (National Center for Health Statistics) and Caroline H Guinn (University of South Carolina) for their comments on earlier versions of this article, and J Whitney Keener (Cleveland State University) for her assistance in preparing the manuscript. We thank each named individual for his or her permission to acknowledge him or her.

References

1 Masson IF, McNeill G, Tomany JO, Simpson JA, Pearce HS, Wei L, *et al*. Statistical approaches for assessing the relative validity of a food frequency questionnaire: use of correlation coefficients and the kappa statistic. *Public Health Nutrition* 2003; **6**: 313–21.

2 Achterberg C, Pugh MA, Collins S, Getty VM, Shannon B. Feasibility of telephone interviews to collect dietary recall information from children. *Journal of the Canadian Dietetic Association* 1991; **52**: 226–8.

3 Bransby ER, Daubney CG, King J. Comparison of results obtained by different methods of individual dietary survey. *British Journal of Nutrition* 1948; **2**: 89–110.

4 Carter RL, Sharbaugh CO, Stapell CA. Reliability and validity of the 24-hour recall. *Journal of the American Dietetic Association* 1981; **79**: 542–7.

5 Conway JM, Ingwersen LA, Vinyard BT, Moshfegh AJ. Effectiveness of the US Department of Agriculture 5-step multiple-pass method in assessing food intake in obese and nonobese women. *American Journal of Clinical Nutrition* 2003; **77**: 1171–8.

6 Conway J, Ingwersen L, Moshfegh AJ. Accuracy of dietary recall using the USDA five-step multiple-pass method in men: an observational validation study. *Journal of the American Dietetic Association* 2004; **104**: 595–603.

7 Crawford PB, Obarzanek E, Morrison J, Sabry ZI. Comparative advantage of 3-day food records over 24-hour recall and 5-day food frequency validated by observation of 9- and 10-year-old girls. *Journal of the American Dietetic Association* 1994; **94**: 626–30.

8 Emmons L, Hayes M. Accuracy of 24-hr recalls of young children. *Journal of the American Dietetic Association* 1973; **62**: 409–15.

9 Lytle LA, Nichaman MZ, Obarzanek E, Glovsky E, Montgomery D, Nicklas T, *et al*. Validation of 24-hour recalls assisted by food records in third-grade children. *Journal of the American Dietetic Association* 1993; **93**: 1431–6.

10 Lytle LA, Murray DM, Perry CL, Eldridge AL. Validating fourth-grade students' self-report of dietary intake: results from the 5-A-Day Power Plus program. *Journal of the American Dietetic Association* 1998; **98**: 570–2.

11 Weber JL, Lytle L, Gittelsohn J, Cunningham-Sabo L, Heller K, Anliker JA, *et al*. Validity of self-reported dietary intake at school meals by American Indian children: The Pathways Study. *Journal of the American Dietetic Association* 2004; **104**: 746–52.

12 Reynolds LA, Johnson SB, Silverstein J. Assessing daily diabetes management by 24-hour recall interview: the validity of children's reports. *Journal of Pediatric Psychology* 1990; **15**: 493–509.

13 Samuelson G. An epidemiological study of child health and nutrition in a northern Swedish county. II. Methodological study of the recall technique. *Nutrition & Metabolism* 1970; **12**: 321–40.

14 Todd KS, Kretsch MJ. Accuracy of the self-reported dietary recall of new immigrant and refugee children. *Nutrition Research* 1986; **6**: 1031–43.

15 Smith AF. Cognitive processes in long-term dietary recall. *Vital and Health Statistics* 1991; **6**(4). Also available at http://www.cdc.gov/nchs/data/series/sr_06/sr06_004.pdf. Accessed 31 January 2007.

16 Smith AF, Jobe JB, Mingay DJ. Retrieval from memory of dietary information. *Applied Cognitive Psychology* 1991; **5**: 269–96.

17 Baxter SD, Thompson WO, Litaker MS, Guinn CH, Frye FHA, Baglio ML, *et al*. Accuracy of fourth-graders' dietary recalls of school breakfast and school lunch validated with observations: in-person versus telephone interviews. *Journal of Nutrition Education and Behavior* 2003; **35**: 124–34.

18 Baxter SD, Thompson WO, Litaker MS, Frye FHA, Guinn CH. Low accuracy and low consistency of fourth-graders' school breakfast and school lunch recalls. *Journal of the American Dietetic Association* 2002; **102**: 386–95.

19 Baxter SD, Thompson WO, Smith AF, Litaker MS, Yin Z, Frye FHA, *et al*. Reverse versus forward order reporting and the accuracy of fourth-graders' recalls of school breakfast and school lunch. *Preventive Medicine* 2003; 36: 601–14.

20 Simons-Morton BG, Forthofer R, Huang IW, Baranowski T, Reed DB, Fleishman R. Reliability of direct observation of schoolchildren's consumption of bag lunches. *Journal of the American Dietetic Association* 1992; **92**: 219–21.

21 Simons-Morton BG, Baranowski T. Observation in assessment of children's dietary practices. *Journal of School Health* 1991; **61**: 204–7.

22 Baxter SD, Thompson WO, Davis HC. Trading of food during school lunch by first- and fourth-grade children. *Nutrition Research* 2001; **21**: 499–503.

23 Baglio ML, Baxter SD, Guinn CH, Thompson WO, Shaffer NM, Frye FHA. Assessment of inter-observer reliability in nutrition studies that use direct observation of school meals. *Journal of the American Dietetic Association* 2004; 104: 1385–93.

24 Shaffer NM, Baxter SD, Thompson WO, Baglio ML, Guinn CH, Frye FHA. Quality control for interviews to obtain dietary recalls from children for research studies. *Journal of the American Dietetic Association* 2004; **104**: 1577–85.

25 Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Current Opinion in Lipidology* 2002; **13**: 3–9.

26 Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, *et al*. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *American Journal of Clinical Nutrition* 1999; **69**: 243–9.

27 Kant AH. Dietary patterns and health outcomes. *Journal of the American Dietetic Association* 2004; **104**: 615–35.

28 Choi HK, Willett WC, Stampfer MJ, Rimm E, Hu FB. Dairy consumption and risk of type 2 diabetes mellitus in men. *Archives of Internal Medicine* 2005; **165**: 997–1003.

29 Shin MH, Holmes MD, Hankinson SE, Wu K, Colditz GA, Willett WC. Intake of diary products, calcium, and vitamin D and risk of breast cancer. *Journal of the National Cancer Institute* 2002; **94**: 1301–11.

30 Pereira MA, Jacobs DR, Van Horn L, Slattery ML, Kartashov AI, Ludwig DS. Dairy consumption, obesity, and the insulin resistance syndrome in young adults: The CARDIA study. *JAMA: Journal of the American Medical Association* 2002; **287**: 2081–9.

31 Baech SB, Hansen M, Bukhave K, Jensen M, Sorensen SS, Kristensen L, *et al*. Nonheme-iron absorption from a phytate-rich meal is increased by the addition of small amounts of pork meat. *American Journal of Clinical Nutrition* 2003; 77: 173–9.

32 Rosner B, Gore R. Measurement error correction in nutritional epidemiology based on individual foods, with application to the relation of diet to breast cancer. *American Journal of Epidemiology* 2001; **154**: 827–35.

33 Baxter SD, Smith AF, Guinn CH, Thompson WO, Litaker MS, Baglio ML, *et al*. Interview format influences the accuracy of children's dietary recalls validated with observations. *Nutrition Research* 2003; **23**: 1537–46.