ORIGINAL ARTICLE

Dietary diversity scores can be improved through the use of portion requirements: an analysis in young Filipino children

MC Daniels[1], LS Adair[1,2], BM Popkin[1,2] and YK Truong[3]

[1]Department of Nutrition, University of North Carolina at Chapel Hill, NC, USA; [2]Carolina Population Center, University of North Carolina at Chapel Hill, NC, USA and [3]Department of Biostatistics, University of North Carolina at Chapel Hill, NC, USA

Objectives: Early childhood malnutrition is a pressing international concern which dietary diversity scores (summary scores of food groups in the diet) may be helpful in addressing. We explored three current research needs surrounding diversity scores: the impact of portion size on score function, the relationship of scores to nutrient adequacy and density and the ability of scores to function as screening tools.
Subjects/Methods: 1810 children, age 24 months. Cross sectional study of a birth cohort.
Results: We evaluated two nine food group dietary diversity scores based on 0 and 10 g minimum food group requirements for their relationship to nutrient adequacy and nutrient density. Both scores were significantly correlated with nutrient adequacy and density and predicted statistically significant increases ($P < 0.05$) in the probability of adequacy for all nutrients. However, correlations and predicted increases were somewhat larger for the 10 g score. We also considered the sensitivity and specificity of each score for detecting low and high nutrient adequacy in the population. The 10 g cutoff improved score ability to predict low nutrient adequacy, and reduced the misclassification of subjects for all comparisons.
Conclusions: This research suggests that the score without portion requirements reflects dietary adequacy, but when feasible, further refinement of diversity scores is desirable through the application of minimum portion requirements.
European Journal of Clinical Nutrition advance online publication, 31 October 2007; doi:10.1038/sj.ejcn.1602927

Introduction

Chronic malnutrition affects at least 150 million children under age 5 worldwide (UN, 2004). Nutrient deficiencies affect physiologic systems broadly leading to potentially severe developmental consequences such as reduced immune function, stunted growth, blindness, retardation and physical deformities. At least half of childhood deaths under 5 years are malnutrition related (UNICEF, 1998).

Methods for rapidly assessing dietary adequacy and identifying children at risk of malnutrition in developing country settings are needed. Toward these efforts, various dietary variety indicators have been developed. These indicators consider the number of different food items or food groups contributing to the diet in a given time period. They are useful because they are correlated with nutrient adequacy (Hatloy *et al.*, 1998; Ruel, 2003; Torheim *et al.*, 2004; Steyn *et al.*, 2006; Kennedy *et al.*, 2007) and various anthropometric measures in children (Onyango *et al.*, 1998; Ruel and Menon, 2002; Arimond and Ruel, 2004; Sawadogo *et al.*, 2006). Measurements are simple to collect, easily adapted to diet in various settings and have been used to study diet in early childhood and adulthood. In addition, they are able to help identify key components of diet (foods and food groups), which can be clearly linked to nutritional needs and translated into population-specific nutritional guidelines.

Indicators based on food groups (dietary diversity scores) have been shown to be more valuable in predicting nutrient adequacy than those based on individual foods (Hatloy *et al.*, 1998). But further research is needed to refine and determine

the full utility of these indicators. A recent review of dietary diversity studies suggested that scores might be improved by inclusion of portion size requirements (Ruel, 2003). This is because diversity scores counting any amount of intake from food groups may overemphasize very small amounts of food, and not place proportionate value on larger, more nutritionally meaningful intakes. Another problem is that dietary diversity scores have been shown to correlate with energy intake (McCrory et al., 1999) so that promoting dietary diversity might increase obesity risks as well as nutrient adequacy. To address this, studies evaluating the correlation of diversity indicators with dietary adequacy should also report energy adjusted correlations, or correlations with dietary nutrient density which is independent of energy intake (Ruel, 2003). Studies are also needed to further explore the screening capability of diversity scores. Analyses testing the sensitivity and specificity of diversity scores to the presence of nutrient inadequacy in various contexts have been recommended for this purpose (Hatloy et al., 1998; Ruel, 2003).

This study addresses the above research needs by comparing two scores (one with and one without portion requirements) to determine the strength of their relationships with regular and energy-adjusted nutrient adequacy as well as nutrient density in children aged 2 years.

Subjects and methods

Subjects
The Cebu Longitudinal Health and Nutrition Survey collected data from all singleton offspring ($n = 3080$) of Filipino women who resided in 33 randomly selected barangays (communities) of Metropolitan Cebu and gave birth between May 1, 1983 and April 30, 1984. Dietary data were collected from mothers bimonthly until 24 months after delivery. (Cebu Study Team, 1991; Adair and Popkin, 2001). All Cebu Longitudinal Health and Nutrition Survey research was approved by the University of North Carolina School of Public Health Institutional Review Board for the Protection of Human Subjects.

This analysis uses data from children 24 months of age. Absences at 24 months were primarily due to death ($n = 155$) or migration or refusal ($n = 477$). Children were excluded from this analysis if at the time of data collection they were ill (fever, diarrhea or poor appetite; $n = 403$), or were still being breastfed ($n = 282$). Excluded children meeting the same criteria at 22 months were included, creating a total sample of 1810 children. Forty-nine children fed commercial formulas were excluded in nutrient intake calculations because of the potential for fortified foods to obscure the true relationship of dietary diversity and nutrient intakes.

Dietary data
Dietary data were collected via 24 h recall by trained study personnel. Information was collected on liquids and semisolid/solid foods. Further details are published elsewhere (Bisgrove et al., 1991; Adair et al., 1993).

Food composition data were obtained from the Food and Nutrition Research Institute (FNRI) of the Philippines 1980 Food Composition Table (FNRI, 1980) which did not yet include fortified foods. For foods with partial nutrient data, nutrient values were obtained by hand-matching to similar foods from FNRI (1990), World food Dietary Assessment System (FAO, 1996), or the USDA (2005) food composition tables. Missing nutrient values for mixed foods were calculated after consulting Filipino colleagues and recipes online to obtain listings of major ingredients.

Values for energy, carbohydrates, protein, total fat, calcium, iron, retinol, beta-carotene, thiamin, riboflavin and niacin were compiled for all foods. Vitamin C and zinc intakes were not assessed due to excessive incomplete food composition data. Values for absorbed calcium were computed according to recommendations from a recently developed validation protocol (Kennedy and Nantel, 2006) for international testing of diversity scores in children $\geqslant 2$ years. Absorption levels were assigned as follows: 25% for roots, tubers, grains and legumes; 45% for fruits and vegetables; 5% for high oxalate foods and 32% for all other food groups. We used a list of high oxalate foods provided by FAO (Kennedy and Nantel, 2006), and oxalate values from FNRI (1990) to identify fruits and vegetables with high oxalate content which was defined as > 5 g oxalate per 100 g. Absorbed iron was assumed to be 14% from animal products and 6% for plant foods. This assumes animal products to consist of 40% heme iron (Monsen et al., 1978) with an average absorption of 25 and 60% non-heme iron (average absorption of 6%) (FAO/WHO, 2001). A similar method was recommended in the FAO/WHO expert consultation on nutrient requirements in which the validation protocol was developed (Cohen et al., 2005). A lower value of 11% for animal products was suggested in the protocol because of the low bioavailability of iron in dairy products and their frequent consumption in most countries among infants and young children. We chose to use a higher value because dairy intake was limited in our sample: only 35% of our sample consumed dairy, and consumer intake amounted to a mean of 74.3 g per day.

Dietary diversity scores
Dietary diversity scores were calculated by a tally of food groups consumed. Since varied instrumentation had hindered comparison between studies (Kant, 1996; Ruel, 2003) we chose to base our study on a diversity score which has been recommended for international comparisons. Nine food groups representing major nutritionally important components of the diet and emphasizing a balance between plant and animal foods (Table 1) were evaluated, as recommended by the FAO protocol (Kennedy and Nantel, 2006). Children were awarded one point for consuming a food at least once from each unique food group. A modified

Table 1 Diversity score food groups and their abbreviations

Cereals, roots, tubers	CRT
β-carotene rich fruits and vegetables	BCFV
Other fruits	OFRT
Other vegetables	OVEG
Legumes, pulses, nuts	LPN
Meat, poultry, fish	MPF
Fats, oils	Fats
Dairy	Dairy
Eggs	Eggs

score was also created according to protocol guidelines (Kennedy and Nantel, 2006) and other previous research (Dewey et al., 2005) which required children to eat at least 10 g from a food group before a point was awarded. A 10 g cutoff was not recommended for oil due to its high energy density; nor have guidelines recommended differentiating between smaller fat and oil intakes given their low nutrient content. Therefore, a score point for oil was awarded for either score if at least 1 g of oil was consumed, as suggested previously for children through age 24 months (Dewey et al., 2005). Both scores had a potential range of 1–9 points.

Recommended nutrient intakes
International nutrient intake recommendations from the World Health Organization (FAO/WHO, 2001) were used since the applicability of diversity scores is of global interest. Estimated Average Requirement (EAR) values were back-calculated from Recommended Nutrient Intakes (RNI) values for thiamin, riboflavin, niacin and calcium using the formula $EAR = RNI/(1 + 2 \times CV)$ where CV is the coefficient of variation for individual nutrient requirement distributions recommended by the Institute of Medicine (IOM, 2001). As recommended we used a CV of 10% for all nutrients other than vitamin A (CV = 20%) and niacin (CV = 15%). No RNI values were listed for vitamin A or absorbed calcium, but approximate EAR equivalents were given by WHO, from which RNIs were calculated. (Note: FNRI retinol equivalents included only β-carotene (1 RE = 6 μg β-carotene) (FNRI, 1990)). Percent RNI (%RNI) was calculated for each child as individual nutrient intake divided by the nutrient's RNI, times 100%. Percentages were then averaged for children at each diversity score level. To allow visualization of total increases across score levels, percentages were not truncated at 100%.

Probabilities of adequacy
A method for calculating probabilities of nutrient adequacy for a single day was set forth by Foote et al. (2004) and uses requirement distributions for each nutrient which are defined by their mean (that is EAR), and s.d. which is the EAR multiplied by the distribution CV.

The probability of adequacy (PA) equals the probability that a child's intake is adequate on a single day and

corresponds to the proportion of the requirement distribution that is below the child's measured intake. For example, if a child consumes 15 mg of thiamin in a day, we compare this number to the requirement distribution for thiamin intakes (mean or EAR = 12.31 and s.d. = 1.85). Assuming a normal distribution, we can calculate that 92.7% of this distribution is below 15 mg. Therefore the child's PA for thiamin on the particular day is 0.93. PA in this respect does not reflect the adequacy of a person's usual diet, but is useful in epidemiologic research for assessing adequacy for larger groups of people. The requirement distribution for iron is not normally distributed; therefore probabilities of adequacy were derived from table I-5 in the IOM manual for iron (IOM, 2001) as suggested by the protocol (Kennedy and Nantel, 2006).

Mean probability of adequacy (MPA) was calculated as the average of individual nutrient probabilities of adequacy (vitamin A, thiamin, riboflavin, niacin, absorbed calcium and absorbed iron) the distributions of which were truncated at a probability of 1.

Nutrient densities
Nutrient densities per 100 kcal were computed for each child over the 24 h period as 100 times the ratio of each nutrient intake divided by total caloric intake (kcal).

Sensitivity/specificity analysis
We tested the ability of both diversity scores to detect the prevalence of low and high mean nutrient adequacy (MPA < 50 or ⩾ 75%, respectively) using a sensitivity/specificity analysis. Increasing sensitivity always results in a corresponding decline in specificity; therefore we depict the relationship between the two through the use of ROC curves to facilitate choosing a score cut point which will optimize both. To reflect the full extent of misclassification we also investigated the proportions of all children tested who would be classified as 'false positives' and as 'false negatives' (children who are misclassified by the score cutoffs), and the 'positive predictive value' which can be interpreted as the true likelihood of having the outcome among those testing positive.

Statistical analysis
Sociodemographic differences between included and excluded children were evaluated using one-way analysis of variance for continuous data (birth weight, parental age and education, maternal height and parity, household education and assets). Tabulations and exact P-values were used for dichotomous sociodemographic variables (children's gender, urban location and the presence of electricity, piped water, flush toilets). Bivariate methods were used to evaluate relationships between dietary diversity and food group nutrient intake. Probabilities of adequacy were calculated

using the normprob function in Stata. Spearman's rank correlations were used to test relationships of probability of adequacy and nutrient density with diversity because probabilities of adequacy were not normally distributed. Linear regression analysis was used to model crude predicted increases in grams consumed per food group and %RNI per nutrient with increasing dietary diversity. Sensitivity/specificity analyses to determine the ability of diversity scores to predict high and low probabilities of nutrient adequacy were performed using Stata's 'roctab' function and hand-checked with tabular calculations. All statistical analysis were performed using Stata 9 (StataCorp, 2005). Statistical significance was indicated by $P < 0.05$ for all analyses.

Results

Sample characteristics
Baseline characteristics of our sample ($n = 1810$) differed from those of the original birth cohort. Included children were slightly heavier at birth (3007.7 vs 2962.6 g, $P = 0.005$) and had parents who were slightly younger (28.3 vs 29.6 years (fathers), $P < 0.001$; 25.7 vs 26.4 years (mothers), $P = 0.001$) and better educated (7.5 vs 7.1 years (fathers), $P = 0.002$; 7.3 vs 6.8 years (mothers), $P < 0.001$) with smaller families (2.2 vs 2.4 live births, $P = 0.002$), than those excluded. There was no significant difference in mean weekly income, mean value of household assets, gender, percent urban, percent with electricity and piped water among included compared to excluded children.

Relationship of diversity scores to dietary patterns
Before evaluating score performance, we looked for differences in how the two scores represent dietary patterns. Patterns relevant to adequacy (that is how many children consumed each food group, and how much they consumed) were evaluated.

Frequency of food group consumption. The distribution of children by the 10 g score (DDS10 g) is presented in Table 2, along with the proportion of children at each score level who consumed specific food groups. Cereals/roots/tubers (CRT) were consumed by all but one child. Meat/poultry/fish (MPF) were second most consumed (by half of children with a score of 2, and over ninety percent of children with a score ⩾5). β-Carotene rich fruits and vegetables (BCFV), other fruits and dairy were next most common, consumed by about 10–25% of children with scores of 2 or 3. However, among children with more diverse diets (scores of 6+), dairy was strongly favored over BCFV, while other fruits were in between. Legumes/pulses and nuts (LPN), other vegetables and eggs were seldom consumed in diets with lower diversity (scores <4), but were consumed more frequently than BCFV among children with higher diversity.

Similar patterns were seen for the 0 g diversity score (DDS) (data not shown). As anticipated, the distribution of scores was shifted slightly to the right. Patterns of food selection were similar to those described above, except selection of other vegetables and BCFV was greater than LPN and eggs for all scores.

Average food group intake among consumers. The graphs in Figure 1 represent mean grams of intake of each food group among children who consumed the food group by diversity score. Intakes of CRT are not shown since they were similarly high (∼400 g) and flat across both scores. Children with a diversity score of 1 consumed mainly CRT. Five times as many children received a score of 1 with DDS10 g (131 vs 26), because the other foods they consumed were in such small amounts (<10 g). No children who scored 2 on DDS consumed other vegetables.

Linear regressions were used to quantify increases in the different food group intakes with increasing diversity. For DDS only consumption of MPF and fats and oils increased significantly with diversity increases (3.6 and 0.5 g per diversity unit respectively, $P < 0.01$ for both). Other apparent trends were non-significant including a slight decrease in

Table 2 The distribution of children by the 10 g diversity score

Diversity score	n	Proportion of children with diversity score consuming each food group								
		CRT	BCFV	OFRT	OVEG	LPN	Fats	MPF	Dairy	Eggs
1	132	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
2	372	1.00	0.16	0.09	0.01	0.03	0.03	0.51	0.15	0.01
3	469	1.00	0.23	0.25	0.09	0.09	0.30	0.75	0.26	0.02
4	426	1.00	0.29	0.35	0.15	0.21	0.60	0.88	0.41	0.10
5	268	1.00	0.34	0.54	0.24	0.33	0.78	0.94	0.60	0.23
6	114	1.00	0.47	0.68	0.39	0.43	0.89	0.92	0.80	0.43
7 +	29	1.00	0.52	0.86	0.59	0.66	0.97	1.00	0.93	0.72
Total	1810	1.00	0.25	0.30	0.13	0.16	0.41	0.72	0.35	0.11

Abbreviations: BCFV: beta-carotene rich fruits and vegetables; CRT: cereals, roots, tubers; LPN: legumes, pulses and nuts; MPF: meat, poultry, fish; OFRT: other fruits and vegetables; OVEG: other vegetables.
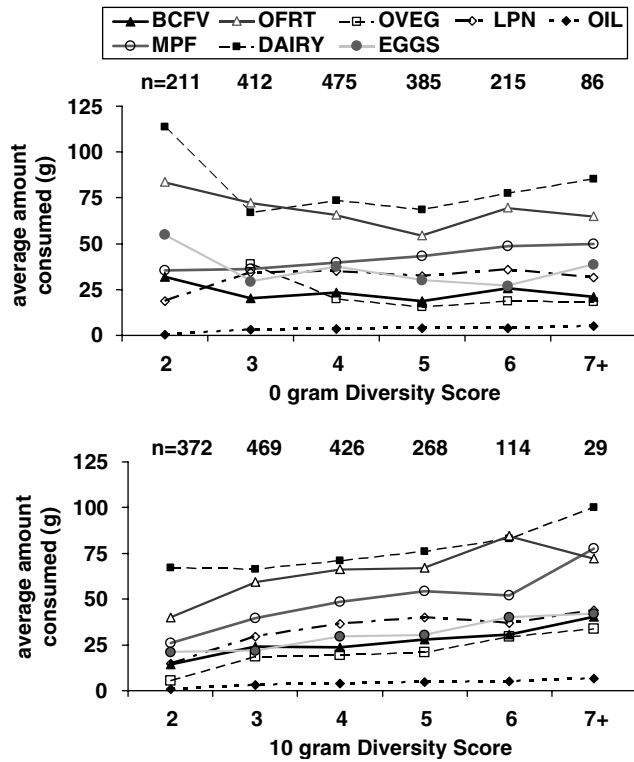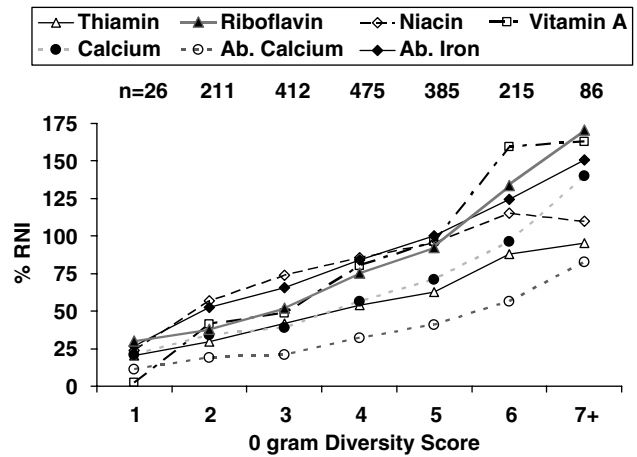
Figure 1 Average grams per food group consumer by diversity score. Minimum intakes in each food category must be >0 and >10 g for assignment into the 0 and 10 g score categories, respectively. However, for calculation of averages all intake >0 g was assessed. Children with score values 7 or greater were combined. BCFV: β carotene-rich fruits and vegetables; OFRT: other fruit; OVEG: other vegetables; LPN: legumes, pulses and nuts; MPF: meat, poultry and fish.

Figure 2 Mean percentages of recommended nutrient intakes achieved across dietary diversity. Based on the 0 g dietary diversity score, %RNI = percent of recommended nutrient intake.

vegetable consumption. For DDS10 g, intakes increased significantly for all food groups other than CRT as diversity increased. Increases in other fruits and MPF were nearly double those of other foods (8.3 and 9.4 g vs <5.5 g, respectively). Among consumers of dairy, LPN, eggs, other vegetables and BCFV, increases in grams consumed were similar (5.4, 5.1, 4.8, 4.7 and 4.5 per score point, respectively).

Relationship of diversity scores to adequacy
We evaluated how well each diversity score predicted nutrient adequacy represented both as a percent of recommended nutrient intakes (%RNI), and as a PA.

Relationship with % recommended nutrient intakes. Mean %RNI increased monotonically across both diversity scores for all nutrients, except niacin declined slightly between levels 6 and 7+ of DDS. Trends for DDS are depicted in Figure 2. Initial percentages were similar for both scores, but increased more dramatically across DDS10 g ending up 26–74% higher in the 7+ category. Differences were as follows:

25.6 (thiamin), 29.5 (niacin), 32.9 (absorbed calcium), 52.2 (riboflavin), 56.1 (calcium), 59.3 (vitamin A), 73.8 (absorbed iron). These sharper increases across DDS10 g reflect the scores greater sensitivity to adequacy.

Correlation with probability of adequacy. Spearman's correlation between diversity scores and PA are shown in Table 3, along with mean nutrient intakes and EAR values used. Unadjusted correlations reflect the strength of the linear relationship between ranked diversity scores and ranked adequacy. Energy adjusted correlations (and alternatively, correlations with nutrient density) reflect the strength of this relationship once energy intake is held constant.

DDS10 g performed more favorably than DDS, achieving higher correlations with PA (adjusted and unadjusted) and nutrient density for most nutrients. Coefficients were at least 45% greater in all three comparisons for niacin, 30–45% percent greater for absorbed iron comparisons and 14–21% greater in all comparisons for thiamin and riboflavin. Coefficients for vitamin A, calcium, and absorbed calcium were 12–15% higher for DDS10 g in unadjusted correlations, but the difference between scores was less clear after energy adjustments.

Energy adjusted correlations were consistently the lowest. This was expected given the known correlations of energy with nutrient intakes and diversity scores. Adjusted coefficients were reduced most dramatically for niacin (−76% for DDS and −55% for DDS10 g) followed by absorbed iron (−51 and −45%, respectively). Since these reductions were greater for DDS, mean probability of adequacy was also more reduced (−45 vs −38%). This seems to suggest greater resilience to energy adjustment in DDS10 g for these two nutrients. Percent reductions for other nutrients were similar.

Table 3 Correlations of dietary diversity with probabilities of nutrient adequacy and with nutrient density

Variable	Intake (s.d.)	EAR[a]	Probability of adequacy				Nutrient density	
			Correlation	Adjusted correlation[b]	Correlation	Adjusted correlation[b]	Correlation	Correlation
Diversity score cutoff, g			0	0	10	10	0	10
Vitamin A, μg RE[c]	233.82 (478.64)	200	0.4763	0.3484	0.5348	0.3717	0.3535	0.3617
Thiamin, mg	0.28 (0.26)	0.4	0.4832	0.2594	0.5823	0.2945	0.3736	0.4258
Riboflavin, mg	0.40 (0.51)	0.4	0.5092	0.3128	0.6172	0.3785	0.4335	0.5062
Niacin, mg	5.11 (3.91)	4.6	0.3447	0.0844	0.5073	0.2298	0.0574	0.1691
Calcium, mg	303.48 (362.10)	417	0.5196	0.3453	0.5998	0.3776	0.3766	0.3945
Absorbed calcium, mg	91.48 (116.00)	220	0.5343	0.3684	0.6126	0.3985	0.4179	0.4382
Absorbed iron, mg	0.50 (0.49)	na	0.4188	0.2074	0.5461	0.3025	0.1462	0.1989
Mean PA[d]			0.484	0.2655	0.6268	0.3917		

Correlations are Spearman's Rank for nonparametric analysis. $P<0.001$ for all correlations.
[a]Estimated average requirement.
[b]Energy adjusted correlations.
[c]Retinol equivalents.
[d]Mean probability of adequacy for all nutrients (uses absorbed calcium).
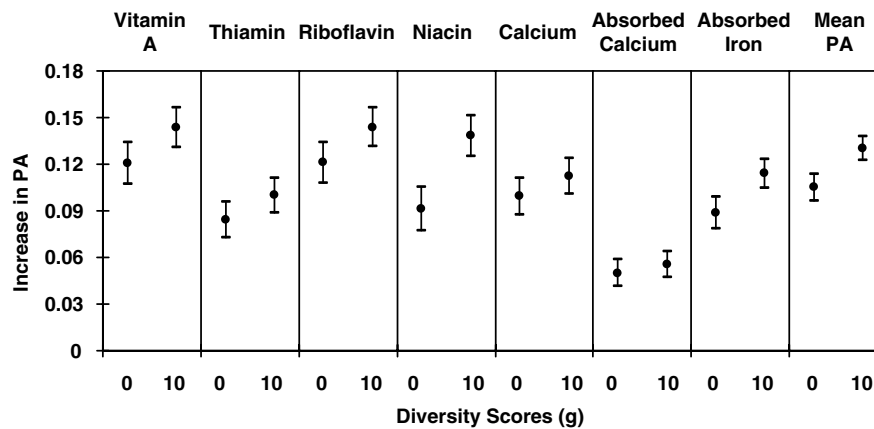
Figure 3 Increases in probability of nutrient adequacy per one unit change in diversity scores. PA = probability of adequacy. Mean PA = mean probability of adequacy for all nutrients (uses absorbed calcium). Data points represent coefficients from linear regressions using diversity scores to predict probability of adequacy for each nutrient. Error bars correspond to 95% confidence intervals. $n=1761$ non-breastfed, well children who were not receiving fortified formula.

Linear increases in probabilities of adequacy. We evaluated the relationship between adequacy and the dietary diversity scores with linear regressions (see Figure 3). Relative increases in PA ranged between 5.0% (absorbed calcium) to 12.1% (riboflavin) per point for DDS, and from 5.6% (absorbed calcium) to 14.4% (riboflavin and vitamin A) per point for DDS10 g. Predicted increases in PA were greater for every nutrient using DDS10 g, reflecting a better ability to differentiate children with differences in nutrient adequacy. In energy adjusted models (not shown) diversity remained significantly correlated with PA for all nutrients, though coefficients were attenuated. Coefficients for DDS10 g remained higher for all nutrients except absorbed calcium.

Note that since the diversity scores are correlated, the confidence intervals in Figure 3 describe individual coefficients only and not the relationship between the scores.

We evaluated whether this correlation might affect the pattern of regression coefficients by randomly selecting a 50% sample and comparing models using each score on opposite sample halves. Similar patterns were observed.

Sensitivity/specificity analysis
We used sensitivity/specificity analysis to evaluate score performance in screening for low and high mean nutrient adequacy.

Roc curves. Receiver operator curves (Figure 4) are a plot of the sensitivity (*y* axis) and specificity (*x* axis) of diversity score cutoffs for detecting adequacy. Each point along the curve represents a single score cutoff, and greater perpendicular distance from a point on the curve to the diagonal
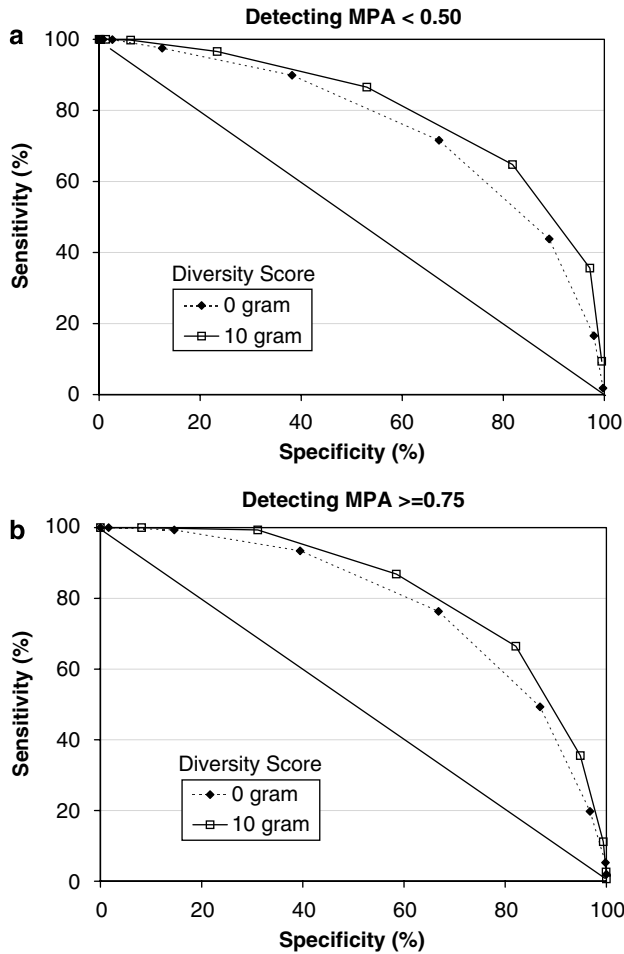
Figure 4 Roc Curves depicting sensitivity and specificity of 0 and 10 g diversity scores for detecting MPA. MPA = mean probability of adequacy. Panel a evaluates scores for detecting low MPA (<0.50); panel b evaluates scores for detecting high MPA ($\geqslant$0.75).

indicates greater utility of the indicator at that cutoff. Likewise, a greater total area between the curve and the diagonal suggests a better overall indicator. From the figure, we selected a range of cutoffs for each score which maximized sensitivity and specificity to nutrient adequacy: for determining low intake (panel a) we selected cutoffs ranging from $\leqslant$2 to $\leqslant$5 for further evaluation; for determining high intake (panel b) we selected cutoffs of $\geqslant$4 to $\geqslant$6. We also evaluated total area under the curve and found that it was greater for DDS10 g in both outcomes tested (Area = 0.754 vs 0.802 testing MPA <0.50, and 0.781 vs 0.822 testing MPA $\geqslant$0.75).

Selecting cutoffs and a preferred score. The cutoffs we selected above for each score are presented in Table 4, with corresponding sensitivity and specificity values. Comparing scores for detecting low adequacy (MPA <0.50) DDS10 g had greater sensitivity at all cutoffs, but lower specificity. A cutoff of four maximized sensitivity and specificity for both scores,

and both had high positive predictive value (those identified as 'at risk' truly were at risk). A cutoff of three also resulted in relatively high sensitivity and specificity for DDS10 g, but false negatives and overall misclassification were increased. Screening projects generally prefer to identify more of those 'at risk', which means reducing false negatives through false positives may be increased. This is especially true if the intervention is affordable, and safe if given to those not at risk. Given this, we preferred cutoffs of four for both scores, but preferred DDS10 g over DDS since DDS10 g resulted in many fewer false negatives.

Comparing scores for detecting high adequacy (MPA $\geqslant$0.75), DDS10 g had lower sensitivity at all cutoffs, but higher specificity. For DDS, a cutoff of five maximized sensitivity while maintaining acceptable specificity. Only a small proportion of children with highly adequate diets were not identified (2% false negatives). However, only 18% of those identified as 'high' were actually high (positive predictive value). The other 82% misclassified comprised about 30% of the total sample (false positives). DDS10 g had similar difficulties identifying children with 'high' adequacy. A score of four identified 87% of these (see sensitivity) with few false negatives, but many false positives (38%). A score of five had fewer false positives, but only identified 66% of those with 'high' adequacy. Given these tradeoffs, a 'preferred' score for identifying high adequacy was not clear.

Discussion

This study focused on three current research needs surrounding diversity scores: the impact of minimum portion size requirements on score function, the relationship of scores to adequacy before and after energy adjustment, and the ability of scores to function as screening tools.

Both scores we tested were functional—each correlated significantly with nutrient adequacy and predicted significant increases (P<0.05) in the probability of adequacy for all nutrients. Previous studies have demonstrated the relationship between similar scores and adequacy in several contexts (Hatloy *et al.*, 1998; Ruel, 2003; Arimond and Ruel, 2004; Kennedy *et al.*, 2007). Kennedy's recent work in a slightly older Filipino population also compared 0 and 10 g diversity scores and found slightly higher correlations with adequacy using a 10 g cutoff. The obvious reason for this is that without a minimum requirement, scoring is easily inflated by intakes of very small amounts of food, which don't contribute appreciably to nutrient intakes. By requiring children to consume 10 g of every food group (except oil) before receiving credit, the modified score reassigned a portion of these misclassified children to appropriately lower score categories and strengthened the scores relationship with adequacy. We demonstrated a second reason for improvements in the 10 g score. Diversity was positively associated with intake for all food groups, a trend that has been cited elsewhere (McCrory *et al.*, 1999). This trend was

Table 4 Sensitivity and specificity analysis evaluating diversity scores for detecting high and low MPA

n	Cutoff	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	False positives (%)	False negatives (%)	Total misclassified (%)
Ability to detect low MPA (MPA < 0.50)							
0 g DDS							
n ⩽ cutoff	⩽						
236	2	16.57	97.92	96.61	0.45	65.19	65.64
645	3	43.82	89.09	93.49	2.39	43.90	46.28
1111	4	71.58	67.27	88.66	7.16	22.20	29.36
1475	5	89.90	38.18	83.86	13.52	7.89	21.41
10 g DDS							
501	2	35.61	97.14	97.80	0.62	50.31	50.94
969	3	64.75	81.82	92.72	3.98	27.54	31.52
1372	4	86.56	52.99	86.81	10.28	10.51	20.78
1624	5	96.58	23.38	81.83	16.75	2.67	19.42
Ability to detect high MPA (MPA ⩾ 0.75)							
0 g DDS							
n ⩾ cutoff	⩾						
1116	4	93.42	39.47	12.72	55.31	0.57	55.88
650	5	76.32	66.81	17.85	30.32	2.04	32.37
286	6	49.34	86.89	26.22	11.98	4.37	16.35
10 g DDS							
800	4	86.84	58.48	16.50	37.93	1.14	39.07
389	5	66.45	82.10	25.96	16.35	2.90	19.25
137	6	35.53	94.84	39.42	4.71	5.57	10.28

Abbreviations: DDS, dietary diversity score; MPA, mean probability of adequacy. Total $n = 1761$.
Top panel: Sensitivity = % at or below cutoff, among children with low MPA. Specificity = % above cutoff, among children without low MPA; positive predictive value = % with low MPA, among children at or below cutoff.
Bottom panel: Sensitivity = % at or above cutoff, among children with high MPA. Specificity = % below cutoff among children without high MPA. Positive predictive value = % with high MPA among children at or above cutoff.

not observed with DDS, possibly due to the inflation from very small gram intakes. That DDS10 g improved correlations with niacin and iron may relate to its cleaner differentiation of increasing intakes of meat, poultry or fish (Figure 1).

We used energy adjustment to better understand how scores were related to adequacy. A positive correlation can mean two things: either children with higher diversity are getting more nutrient rich diets, or they are simply getting absolute nutrient increases through more food. Further studies differentiating these relationships are needed (Ruel, 2003). We found that energy adjustment reduced but did not remove the significant positive relationships between dietary diversity and adequacy for either score. This suggests that diversity driven increases in adequacy are at least partially due to increased nutrient density. Also, energy adjusted correlations were somewhat higher for DDS10 g, indicating the 10 g requirement improved score performance.

Some nutrient correlations were less robust to energy adjustment (that is thiamin, niacin, absorbed iron). This may be due to these nutrients being more concentrated in commonly consumed, energy dense foods such as meats and cereals. Calcium and vitamin A also occur in these foods, but are distributed more broadly in fruits, vegetables and dairy as well.

The sensitivity/specificity analysis indicated that both scores could be used to detect low mean probability of adequacy (MPA < 0.50). DDS10 g performed most favorably— a cutoff of four identified 87% of low MPA children with only a modest number of false positives and negatives (10 and 11%, respectively). Both scores had difficulty detecting high MPA, largely because high MPA was less common and many false positives occurred in identifying them. Excessive 'false positives' generally inflate the cost of interventions. However, in our case 'testing positive' for high MPA was favorable and populations with large numbers of children thus identified may need less intervention. While our numbers are highly context-specific, the potential for improved screening capabilities invites further research.

While these analyses show greater score sensitivity with the use of minimum portion requirements, many questions remain to be answered. Initially, it is important that these findings are confirmed in other populations. Also, the benefits gained must be balanced against the impracticality of minimum requirements. No-cutoff diversity scores are derived only from food types consumed. Introducing a cutoff into the score will require that field-workers ask questions about amounts consumed as well. This could be relatively simple if interviewers presented models for typical requirement volumes to interviewees, but this approach would need to be validated. In addition, work is needed to determine whether minimum requirements should vary

across food groups or by age of children. Minimums should be high enough to screen out noise in the score, but low enough to retain sensitivity to nutritionally significant intakes. Some misclassifications will be inevitable: for example, organ meat is very nutrient dense and could appropriately be counted at low levels, impractical for the majority of foods. Studies on within-food group variety, to determine whether more than one cutoff should be applied within food groups for foods of varying nutrient densities, may be useful.

Beyond addressing our main questions, the analysis shows the functionality of diversity scores in elucidating dietary patterns. Both scores revealed similar patterns in food intake: a cereal- or root-based diet most commonly complemented with meat, fish or poultry. Consumption of dairy, fruits and BCFV was less frequent. Other vegetables, LPN and eggs were consumed even less. Also since consumption of BCFV and other vegetables was measured higher than LPN and eggs only by DDS, consumption of BCFV and other vegetables must often be in miniscule amounts. By revealing broad dietary needs (for example increased intakes of fruits and vegetables) diversity scores can inform interventions.

Other patterns visible specifically in this analysis may be useful in context. Among consumers, mean grams of BCFV, other vegetables, LPN and eggs eaten did not exceed 45 g at any diversity score level. Increasing both the frequency and amount consumed for these foods will likely improve diet quality. Consumers of other fruits tended to eat larger amounts, but these other fruits were almost exclusively different types of bananas, which are grown commercially in the Philippines. Green papaya and lemon were also eaten frequently, but in very small amounts (mean of ~10 g). Foote *et al*. (2004) recently demonstrated the importance of within-group variety of fruits, vegetables, dairy and grain contributing to probability of adequacy in a US population. Therefore, increasing the within-group variety of fruits consumed in this population is likely to yield improved nutritional adequacy.

Children in this analysis were slightly better off at baseline in a few respects (higher birth weight, younger and better educated parents) than excluded children. Loss to follow-up came from both extremes of the socioeconomic spectrum (that is poorer children had higher mortality, wealthier children more often migrated). At 2 years, children excluded for illness did not differ significantly, while those excluded for breastfeeding were from poorer circumstances with lower income households in spite of larger families, older and less educated parents and poorer housing conditions (less access to toilets, running water, electricity). However, the slight differences between our analysis sample and the original cohort are unlikely to reduce the applicability of our findings. Children from our sample were somewhat shorter than those of similar age from a nationally representative sample taken in the same year (-2.8 cm girls, -3.9 cm boys) (Florentino *et al*., 1992), and are clearly nutritionally disadvantaged. Our sample is typical of the nutritionally challenged populations which dietary diversity scores are being developed to help identify and assist.

The study has several strengths. The sample is of respectable size and located in an area where nutritional improvements are needed. The Cebu data is extensive with multiple follow-up surveys in later years, which presents a unique opportunity to evaluate dietary diversity at successive ages in the same individuals. This preparatory work in infant diversity will allow further studies of diversity trends in this population. The study uses a proven instrument and has been informed by recent FAO guidelines (Kennedy and Nantel, 2006) in an effort to be comparable with forthcoming analyses. The study compares in detail two versions of the diversity score, which have been recommended for comparison. Sensitivity/specificity analysis has been used in only a few studies to date (Hatloy *et al*., 1998; Kennedy *et al*., 2007), but was recommended by Ruel (2003) for future use in evaluating the functionality of dietary diversity indicators.

Study weaknesses include those inherent in dietary data research. Nutrient variation within foods, nutrient analysis error and food substitution errors lead to imperfections in food composition information. In spite of precautions in data collection and analysis, some error in reported intakes inevitably occurs due to error in recipes, portion sizes, and dish ingredients. In addition, uncertainty about the actual shape of nutrient requirement distribution remains.

Dietary diversity scores are a promising method for identifying populations at increased risk of malnutrition and related advances are valuable. We found that diversity scores for small children can be improved by applying a 10 g minimum portion requirement. Improvements included stronger correlations with nutrient adequacy as well as energy adjusted nutrient adequacy and nutrient density, greater increases in overall nutrient intakes across score levels and better functionality in screening tests. However, further refinement of these indicators is needed. Strengthening the relationship of scores to nutrient adequacy and density remains a priority, as does further testing of scores as screening tools in a variety of populations. Care should be taken that scores remain practical for field use, that is as affordable and simple to administer as possible. Also, research is needed assessing the potential of scores to detect inadequacy in older children and teens. As dietary diversity scores are improved, studies evaluating predictors of dietary diversity will become important guidance for targeting diversity interventions. We encourage further research in these and related areas.

10

References

Adair LS, Popkin BP (2001). The Cebu longitudinal health and nutrition survey: history and major contributions of the project. *Philipp Q Cult Soc* **29**, 5–37.

Adair LS, Popkin BM, VanDerslice J, Akin JS, Guilkey DK, Black R *et al.* (1993). Growth dynamics during the first two years: A Prospective Study in the Philippines. *Eur J Clin Nutr* **47**, 42–51.

Arimond M, Ruel MT (2004). Dietary diversity is associated with child nutritional status: evidence from 11 demographic and health surveys. *J Nutr* **134**, 2579–2585.

Bisgrove E, Popkin BM, Barba C (1991). Profiling infant-feeding patterns in the Philippines. *Ecol Food Nutr* **26**, 91–302.

Cebu Study Team (1991). Underlying and proximate determinants of child health: The Cebu Longitudinal Health and Nutrition Study. *Am J Epidemiol* **133**, 2 185–201.

Cohen R, Arimond M, Dewey K, Ruel M (2005). *Developing and Validating Simple Indicators of Complementary Food Intake and Nutrient Density for Infants and Young Children in Developing Countries: Protocol for Data Analysis. Final Report*. Food and Nutrition Technical Assistance (FANTA) Project, Academy for Educational Development (AED): Washington, DC.

Dewey KG, Cohen RJ, Arimond M, Ruel MT (2005). *Developing and Validating Simple Indicators of Complementary Food Intake and Nutrient Density for Breastfed Children in Developing Countries. Final Report*. The Food and Nutrition Technical Assistance (FANTA) Project, Academy for Educational Development (AED):Washington DC.

Florentino RF, Santos-Ocampo PD, Magbitang JA, Mendoza TS, Flores EG, Madrid BD (1992). *FNRI-PPS Anthropometric Tables and Charts for Filipino Childres*. Food and Nutrition Research Institute (FNRI), Department of Science and Technology (DOST): Manilla, Philippines.

Food and Agricultural Organization of the United Nations (FAO) (1996). *Worldfood Dietary Assessment System 2 Database*. Available online at:http://www.fao.org/infoods.Accessed 8 July 2006.

Food and Agriculture Organization of the United Nations (FAO) and World Health Organization (WHO) (2001). *Human Vitamin and Mineral Requirements*. Food and Nutrition Division, Food and Agriculture Organization of the United Nations (FAO): Rome, Italy.

Food and Nutrition Board, Institute of Medicine (IOM) (2001). *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silver, Vanadium, and Zinc*. The National Academies Press: Washington DC.

Food and Nutrition Research Institute of the Philippines (FNRI) (1980). *Food Composition Tables*. Department of Science and Technology (DOST): Manilla, Philippines.

Food and Nutrition Research Institute of the Philippines (FNRI) (1990). *Food Composition Tables: Recommended for use in the Philippines, Handbook 1*. Department of Science and Technology (DOST): Manilla, Philippines.

Foote JA, Murphy SP, Wilkens LR, Basiotis PP, Carlson A (2004). Dietary variety increases the probability of nutrient adequacy among adults. *J Nutr* **134**, 1779–1785.

Hatloy A, Torheim LE, Oshaug A (1998). Food variety—a good indicator of nutritional adequacy of the diet? A case study from an urban area in Mali, West Africa. *Eur J Clin Nutr* **52**, 891–898.

Kant AK (1996). Indexes of overall diet quality: a review. *J Am Diet Assoc* **96**, 785–791.

Kennedy G, Nantel G (2006). *Basic Guidelines for Validation of a Simple Dietary Diversity Score as an Indicator of Dietary Nutrient Adequacy for Non-Breastfeeding Children 2–6 Years*. Food and Agriculture Organization of the United Nations: Rome, Italy. Available online at: http://www.fao.org/ag/agn/nutrition/assessment_en.stm. Accessed 1 July 2006.

Kennedy GL, Pedra MR, Seghieri C, Nantel G, Brouwer I (2007). Dietary diversity score is a useful indicator of micronutrient intake in Non-breastfeeding Filipino children. *J Nutr* **137**, 472–477.

McCrory MA, Fuss PJ, McCallum JE, Yao M, Vinken AG, Hays NP *et al.* (1999). Dietary variety within food groups: association with energy intake and body fatness in men and women. *Am J Clin Nutr* **69**, 440–447.

Monsen ER, Hallberg L, Layrisse M, Hegsted DM, Cook JD, Mertz W (1978). Estimation of available dietary iron. *Am J Clin Nutr* **31**, 134–141.

Onyango A, Koski KG, Tucker KL (1998). Food diversity versus breastfeeding choice in determining anthopometric status in rural Kenyan toddlers. *Int J Epi* **27**, 484–489.

Ruel MT (2003). Operationalizing dietary diversity: a review of measurement issues and research priorities. *J Nutr* **133**, 3911S–3926S.

Ruel MT, Menon P (2002). Child feeding practices are associated with child nutritional status in Latin America: innovative uses of the demographic and health surveys. *J Nutr* **132**, 1180–1187.

Sawadogo PS, Martin-Prével Y, Savy M, Kameli Y, Traissac P, Traoré AS *et al.* (2006). An infant and child feeding index is associated with nutritional status of 6- to 23-month-old children in rural Burkina Faso. *J Nutr* **136**, 656–663.

StataCorp (2005). *Stata Statistical Software Release 9.0*. Stata Corporation: College Station, TX.

Steyn NP, Nel JH, Nantel G, Kennedy G, Labadarios D (2006). Food variety and dietary diversity scores in children: are they good indicators of dietary adequacy? *Public Health Nutr* **9**, 644–650.

Torheim LE, Ouattara F, Diarra MM, Thiam FD, Barikmo I, Hatloy A *et al.* (2004). Nutrient adequacy and dietary diversity in rural Mali: associations and determinants. *EJCN* **58**, 594–604.

United Nations Children's Fund (UNICEF) (1998). *The State of the Worlds Children: Focus on Nutrition*. Oxford University Press: New York.

UN Standing Committee on Nutrition 5th Annual Report on the World Nutrition Situation: nutrition for improved development outcomes. March 2004.

United States Department of Agriculture (2005). *National Nutrient Database for Standard Reference, Nutrient Data Laboratory Home Page*. Agriculture Research Service (ARS), USDA: Washington DC. Available online at: http://www.ars.usda.gov/main/site_main.htm?modecode = 12354500Accessed 8 July 2006.