A New Index to Measure Healthy Food Diversity Better Reflects a Healthy Diet Than Traditional Measures1,2

Larissa S. Drescher,3 Silke Thiele,3* and Gert B. M. Mensink4

1Supported by the German Science Foundation (DFG).
2Supplemental Table 1 and Supplemental Figures 1 and 2 are available with the online posting of this paper at jn.nutrition.org.
3Department of Food Economics and Consumption Studies, Christian-Albrechts-University of Kiel, D-24098 Kiel, Germany
4Robert Koch-Institute, D-13353 Berlin, Germany

Abstract

The recommendation to eat diverse types of foodstuffs is an internationally accepted recommendation for a healthy diet. The importance of dietary variety is based on several studies that have shown that diverse diets are accompanied by positive health outcomes. However, the definition and measurement of healthy food diversity are often criticized in the literature. Nutritional studies generally use count indices to quantify food diversity. As these measures have considerable disadvantages, several nutritionists have called for a precise definition and measurement of food diversity. This study aimed to develop a new healthy food diversity indicator. This index is based on a distribution measure mainly applied in economic and ecological studies. It considers 3 aspects important for healthy food diversity: number, distribution, and health value of consumed foods. We have validated the new index using energy-adjusted correlations with diet quality indicators. A comparison with selected traditional diversity indices revealed that the new indicator more appropriately reflected healthy food diversity. J. Nutr. 137: 647–651, 2007.

Introduction

Eating a large diversity of foods is an internationally accepted recommendation for a healthy diet, because it is associated with positive health outcomes such as reduced incidence of cancer or mortality (1–5). For an assessment of the diversity of an individual’s consumption, appropriate measures are necessary. In the nutritional literature, count measures are frequently applied, whereby the number of consumed food items and food groups is recorded (1,6–8). These indices have 2 crucial disadvantages. First, they do not distinguish whether the observed variety is the result of different healthy or unhealthy products (fruits and sweets, respectively). Previous studies dealt with this obstacle by merely counting desired foods while omitting undesired foods (e.g. 3,7–9). Because it is questionable to categorize foods such as sugared cereals, this method is normative.

Second, count measures do not account for the distribution of individual food quantities. An individual with equal shares of food products has a greater food diversity than an individual who consumes 90% of 1 product and 10% of the others. A food diversity index must reflect this. Altogether, a healthy food diversity index must consider 3 aspects simultaneously: number, distribution, and health value of a consumed food basket. All aspects are emphasized in newer recommendations for healthy eating. The German diversity guideline, for example, underlines: “Enjoy the great variety of food. There is no ‘healthy’, ‘unhealthy’ or even ‘forbidden’ food. It is the quantity, selection, and combination of food that matters” (10).

Particularly in economic studies, there is a growing application of distribution measures that consider number as well as distribution of different (food) products to quantify diversity (11,12). We wanted to determine whether these indices are suitable measures of healthy food diversity. This answers the demand of many nutritionists for a more precise definition of food diversity and to develop a suitable indicator (13,14,8,15).

Materials and Methods

Index construction. There are several different methods to measure dietary diversity. For example, Kant et al. (1) constructed the Dietary Diversity Score. This index counts the number of food groups consumed daily: dairy, meat, grain, fruits, and vegetables. The maximum score is 5; 1 point is counted for each group consumed. Drewnowski et al. (6) developed the Dietary Variety Score, which is based on the cumulative number of 164 different foods consumed over a 15-d period. In the Diet Quality Index-International, Kim et al. (16) integrated a dietary variety component where variety is evaluated in 2 ways, i.e. the overall variety is recorded by the number of 5 different food groups consumed daily and the variety within protein sources is measured by the number of different protein sources. All variety indices focus on counting different food groups and subgroups; the distribution of consumed food quantities is not taken into account.

A measure to evaluate diversity in terms of number as well as distribution of different food items is the Berry-Index. This index was applied mainly in economic food diversity studies (11,12,17–19).
Recently, the Berry-Index, which is also known as the Simpson-Index, 
has been applied in a nutritional study by Katanoeda et al. (20) to 
measure dietary diversity and its annual changes in Japan. The Berry-Index (BI) is 
defined as: $BI = 1 - \sum s^2$ (21), where $s_i$ is the share of product $i$ in the 
total amount of food consumed. The index is bounded between 0 and 1–1/n, 
whose limit value approximates 1 if the number of foods ($n$) increases. 
$BI = 0$ indicates that an individual consumes only 1 food product, $BI = 1–1/n$ refers to a situation where the individual consumes equal shares of 
all products considered.

From a nutritional perspective, the latter assumption is not desired. 
According to food guide recommendations, healthy foods should be 
consumed in higher shares than unhealthy ones. Hence, the highest index 
value has to be assigned to a situation where an individual consumes 
recommended food group shares. The basic idea of the new healthy food 
diversity indicator was to modify the Berry-Index so that the index rises 
if the distribution of foods moves in favor of healthier products. 
Therefore, we incorporated a component into the Berry-Index that is 
able to reflect the health value of consumed foods. In this analysis, we 
derived health values from actual food guidelines of the German 
Nutrition Society (DGE),3 but other food guidelines can also be used as a 
basis for the modification of the Berry-Index.

The visual representation of the DGE food guidelines are illustrated 
by a nutrition circle and a food pyramid. The nutrition circle (Fig. 1) 
illustrates the shares of food groups that should be consumed in terms of weight. 
These shares are calculated on the basis of the DGE reference 
values for nutrient intake. Exemplary diets are constructed with the aim 
that the reference values for nutrient intake are achieved on average over 
a 7-d period (22). In accordance with the sides of the food pyramid, these 
shares can be summed up in 3 groups: 73% plant foods, 25% animal 
foods, 2% fats and oils. The pyramid also includes a 4th dimension that 
refers to beverages. Because we are primarily interested in caloric foods, 
we exclude all noncaloric beverages such as mineral water, coffee, and tea, 
whereas caloric beverages are assigned to other food groups. For 
example, 100% juices are assigned to fruits and vegetables, sugar-
containing beverages (e.g. lemonades) are assigned to sweets, etc.

The food pyramid illustrates the qualitative dimension within these 3 
food groups by the hierarchy of foods. To take the plant food group as an 
example, most valuable (healthy) subgroups such as vegetables and fruits 
are placed at the bottom, less valuable (unhealthy) subgroups such as 
cakes and sweets at the top (Fig. 2).

We made use of these guidelines to create health values for foods. 
This explicit valuation of foods is our own interpretation of the DGE 
guidelines. The constructed health values are dependent on the position 
of foods in the pyramid as well as on the affiliation to a main food group 
(plant foods, animal foods, fats, and oils). The health value of the main 
food groups ($G_w$) was assessed according to the percentage consumption 
recommendation shown in the nutrition circle. For the assessment of 
food subgroups ($G_s$), the qualitative dimension of the pyramid sides had 
to be quantified. It was assumed that each side is divided into 5 different 
subgroups. Because the pyramid is an isosceles triangle, it was further 
supposed that the 5 subgroups have the same heights on all sides. With 
geometric calculations, the percentage of each subgroup within the 
upper group was calculated. Combination of the subgroup shares ($G_w$) 
with the main group shares ($G_s$) yield health factors ($hf = G_w \cdot G_s$) for 
15 different subgroups (Table 1).

Using these health factors, the health value of an individual’s food 
basket was assessed by multiplying the quantitative shares of single foods 
in terms of weight ($s_i$) on total quantities with their corresponding health factor ($hf$). The output is called health value $hv = \sum hf_i \cdot s_i$. The maximum health value that can be achieved is 0.26. Thus, division of 
hv by its maximum ensures that $hv$ is bounded between 1 and nearly 0.

The final construction of the healthy food diversity index was achieved by combining health value ($hv$) and diversity of the food basket $BI = 1 – \sum s^2$. The resulting Healthy Food Diversity (HFD)-Index is 
defined as $HFD = (1 – \sum s^2) \cdot hv$. Multiplication of the Berry-Index with the 
health value ensures that neither a high $hv$ nor a high $BI$ alone yield a 
high HFD-Index. The HFD-Index, which is bounded between 0 and 1–1/n, has the following desired properties: 1) If the distribution between 
$h/f$ groups of the pyramid does not change, it increases with the growing 
number of food items; it increases the more equally the food items are 
distributed within the $h/f$-groups; and 2) If distribution between $h/f$ groups of the pyramid does change in favor of healthy (unhealthy) food groups, 
it increases (decreases).

Therefore, the HFD-Index is able to differentiate between healthy 
and unhealthy food diversity over all food groups based on real 
observable diets without omitting unhealthy foods.

**Data and statistical analyses.** To verify if the developed HFD-Index 
was able to reflect a healthy diet, we conducted Pearson’s correlation 
analyses, where individuals’ HFD were correlated with nutrient supply 
and biochemical parameters. All statistical analyses were conducted 
using SPSS version 12.0. The significance level was set at 0.05 in all 
analyses.

The empirical analyses were based on data from the German 
Nutrition Survey (GeNuS) of 1998, which is representative for 
noninstitutional German adults (23). (GeNuS) was part of the German 
National Health Interview and Examination Survey 1998. The survey 
was approved by the Federal Data Protection officials. Survey partici-
pants were informed in detail about the study goals, interview, and 
examination procedures as well as anonymous data record keeping and analyses. They were able to refuse any part of the examination program. 
Participants provided written informed consent prior to the interview 
and examination.) A number of 4030 participants were comprehensively 
interviewed concerning their diet of the preceding 4 wk by trained

---

3 Abbreviations used: DGE, German Nutrition Society; GeNuS, German Nutrition Survey; HFD, healthy food diversity; NAR, nutrient adequacy ratio.
Results and Discussion

Correlations were conducted between nutrient supply (NAR) and HFD-Index (Table 3). For comparison reasons, correlation results of 2 other indices, the Count-Index as well as the Berry-Index, are listed. The Count-Index used in this analysis counted the number of 133 food items consumed in a 1-mo period. As both dietary variety and nutrient adequacy were strongly correlated with energy intake, energy-adjusted partial correlations are presented. Fooe et al. (27) already illustrated the necessity of applying multivariate models that adjust for energy intake when analyzing the relation between dietary diversity and nutrient adequacy.

All nutrients at risk of deficient supply, except for vitamin B-12, showed significant and positive correlations with the HFD-Index. The highest correlation was observed for folate. The coefficient of 0.644 revealed that a high variety necessity of applying multivariate models that adjust for energy intake when analyzing the relation between dietary diversity and nutrient adequacy.

All nutrients at risk of deficient supply, except for vitamin B-12, showed significant and positive correlations with the HFD-Index. The highest correlation was observed for folate. The coefficient of 0.644 revealed that a high variety necessity of applying multivariate models that adjust for energy intake when analyzing the relation between dietary diversity and nutrient adequacy.

All nutrients at risk of deficient supply, except for vitamin B-12, showed significant and positive correlations with the HFD-Index. The highest correlation was observed for folate. The coefficient of 0.644 revealed that a high variety necessity of applying multivariate models that adjust for energy intake when analyzing the relation between dietary diversity and nutrient adequacy.

All nutrients at risk of deficient supply, except for vitamin B-12, showed significant and positive correlations with the HFD-Index. The highest correlation was observed for folate. The coefficient of 0.644 revealed that a high variety necessity of applying multivariate models that adjust for energy intake when analyzing the relation between dietary diversity and nutrient adequacy.
A comparison of the HFD-Index with the Berry-Index and Count-Index for selected nutrients with notable risk of deficient supply (Fig. 3) revealed that nearly all of them were more strongly correlated with the HFD-Index than with other food diversity indices. Only the supply of nutrients mainly occurring in animal-based foods (thiamin, riboflavin, and calcium) was better reflected in the Berry-Index and Count-Index. Again, this reflects the low valuation of animal-based foods in the German food guidelines and hence in the HFD-Index.

In the case of vitamin B-12, the correlation coefficient for the HFD-Index was not significant (P = 0.11), i.e. higher healthy food diversity was not associated with vitamin B-12 supply. This finding can be explained by the positioning of foods in the German food pyramid. The main suppliers of vitamin B-12 are animal-based foods (thiamin, riboflavin, and calcium) was better reflected in the Berry-Index and Count-Index. Again, this reflects the low valuation of animal-based foods in the German food guidelines and hence in the HFD-Index.

Altogether, the HFD-Index showed for both the supply with nutrients and serum biochemical variables predominantly better correlation results and hence was more suitable for measuring healthy food diversity compared with the Count- and Berry-Index. The inclusion of both diversity and health

---

**TABLE 3** Partial energy-adjusted correlations between nutrient supply (NAR) and different food diversity indices among GeNuS 1998 participants.1,2,3

<table>
<thead>
<tr>
<th>Nutrients with risk of deficient supply</th>
<th>HFD-Index</th>
<th>Berry-Index</th>
<th>Count-Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAR vitamin A (retinol equivalents)</td>
<td>0.157**</td>
<td>0.098**</td>
<td>0.193**</td>
</tr>
<tr>
<td>NAR Vitamin D</td>
<td>0.105**</td>
<td>0.220**</td>
<td>0.222**</td>
</tr>
<tr>
<td>NAR vitamin E</td>
<td>0.435**</td>
<td>0.302**</td>
<td>0.240**</td>
</tr>
<tr>
<td>NAR vitamin K</td>
<td>0.085**</td>
<td>0.169**</td>
<td>0.023**</td>
</tr>
<tr>
<td>NAR thiamin</td>
<td>0.128**</td>
<td>0.154**</td>
<td>0.231**</td>
</tr>
<tr>
<td>NAR riboflavin</td>
<td>0.070**</td>
<td>0.160**</td>
<td>0.247**</td>
</tr>
<tr>
<td>NAR niacin equivalents</td>
<td>0.067**</td>
<td>0.139**</td>
<td>0.158**</td>
</tr>
<tr>
<td>NAR vitamin B-6</td>
<td>0.232**</td>
<td>0.257**</td>
<td>0.243**</td>
</tr>
<tr>
<td>NAR folate</td>
<td>0.644**</td>
<td>0.267**</td>
<td>0.113**</td>
</tr>
<tr>
<td>NAR pantothenic acid</td>
<td>0.313**</td>
<td>0.299**</td>
<td>0.285**</td>
</tr>
<tr>
<td>NAR biotin</td>
<td>0.259**</td>
<td>0.292**</td>
<td>0.257**</td>
</tr>
<tr>
<td>NAR vitamin B-12</td>
<td>−0.027</td>
<td>0.172**</td>
<td>0.286**</td>
</tr>
<tr>
<td>NAR vitamin C</td>
<td>0.344**</td>
<td>0.105**</td>
<td>0.201**</td>
</tr>
<tr>
<td>NAR chloride</td>
<td>0.075**</td>
<td>0.173**</td>
<td>0.051**</td>
</tr>
<tr>
<td>NAR potassium</td>
<td>0.271**</td>
<td>0.238**</td>
<td>0.213**</td>
</tr>
<tr>
<td>NAR calcium</td>
<td>0.142**</td>
<td>0.096**</td>
<td>0.195**</td>
</tr>
<tr>
<td>NAR phosphorus</td>
<td>0.049**</td>
<td>0.106**</td>
<td>0.131**</td>
</tr>
<tr>
<td>NAR magnesium</td>
<td>0.331**</td>
<td>0.242**</td>
<td>0.248**</td>
</tr>
<tr>
<td>NAR iron</td>
<td>0.236**</td>
<td>0.171**</td>
<td>0.224**</td>
</tr>
<tr>
<td>NAR iodine</td>
<td>0.214**</td>
<td>0.057**</td>
<td>0.127**</td>
</tr>
<tr>
<td>NAR fluoride</td>
<td>0.392**</td>
<td>0.191**</td>
<td>0.105**</td>
</tr>
<tr>
<td>NAR zinc</td>
<td>0.117**</td>
<td>0.204**</td>
<td>0.233**</td>
</tr>
<tr>
<td>NAR copper</td>
<td>0.160**</td>
<td>0.115**</td>
<td>0.192**</td>
</tr>
<tr>
<td>NAR dietary fiber</td>
<td>0.484**</td>
<td>0.291**</td>
<td>0.051**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nutrients with risk of excess supply</th>
<th>HFD-Index</th>
<th>Berry-Index</th>
<th>Count-Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAR fat</td>
<td>−0.008</td>
<td>0.038*</td>
<td>0.079**</td>
</tr>
<tr>
<td>NAR cholesterol</td>
<td>0.138</td>
<td>−0.006</td>
<td>0.068**</td>
</tr>
<tr>
<td>NAR saturated/unsaturated fatty acids</td>
<td>0.108**</td>
<td>−0.019</td>
<td>−0.109**</td>
</tr>
<tr>
<td>NAR sugar</td>
<td>0.041*</td>
<td>0.050*</td>
<td>0.031*</td>
</tr>
<tr>
<td>NAR sodium</td>
<td>−0.026</td>
<td>0.120**</td>
<td>−0.170**</td>
</tr>
</tbody>
</table>

1 ** Significance level 1%, * significance level 5%.
2 Correlations based on n = 4030 sample size.
3 All analyses were adjusted for sample weights.
4 Scatter plots of energy-adjusted HFD-Index on selected energy-adjusted NAR are available in Supplemental Figure 2.

---

1 ** Significance level 1%, * significance level 5%.
2 Correlations based on n = 4030 sample size.
3 All analyses were adjusted for sample weights.
4 Scatter plots of energy-adjusted HFD-Index on selected energy-adjusted NAR are available in Supplemental Figure 2.

---

**Figure 3** Deviations of correlation units of the Berry-Index and the Count-Index from the HFD-Index for selected nutrients with notable risk of deficient supply. Deviations are calculated on the basis of partial energy-adjusted correlations of nutrient supply (NAR) with different food diversity indices among GeNuS 1998 participants (Table 4). All shown correlation coefficients were significantly different from 0 at the 5% level. Correlations based on n = 4030 sample size. All analyses were adjusted for sample weights.
recommendation aspects seems to be an important advantage of the HFD-Index. However, the development of a healthy food diversity indicator is dependent on nutritional guidelines for optimal food distribution. The unequal assessment of animal-based foods in different guidelines (e.g. American vs. German) reveals the demand for future research on the valuation of foods. The call for improved healthy food diversity indicators stated in previous studies (15,8,14,13) provided the impetus to develop a new index that considered 3 important aspects of a varied diet simultaneously: number, distribution, and health value of foods. The health value was derived from the recently published German food pyramid. The incorporation of this value into an existing diversity indicator, the Berry-Index, was implemented. The correlation results showed that this new HFD-Index was able to reflect a healthy diet. Nutrient supply variables as well as serum biochemical parameters were significantly correlated with the HFD-Index and the signs were as expected. The comparison with previous indices, the Count-Index and the Berry-Index, showed that the HFD-Index was the most suitable indicator to measure healthy food diversity. Particularly, nutrients with notable risk of deficient supply such as folate, fluoride, dietary fiber, vitamin E, and iodine were better reflected in the new HFD-Index than in traditional variety indices. The highest correlation was detected for folate, which can be found in small amounts in numerous foods. Nutrients mainly (exclusively) occurring in animal foods were not adequately reflected in the HFD-Index. This was explained by the low valuation of animal foods in the German food guidelines.

Taking other food guidelines as a basis for the HFD-Index for analyzing the association between healthy food diversity and nutrient supply is a promising area of future research. The HFD-Index seems to be a suitable foundation to review the performance of different food guidelines in terms of achieving nutrient intake recommendations.

**Literature Cited**


**TABLE 4** Correlations between serum biochemical variables and different food diversity indices among GeNuS 1998 participants.

<table>
<thead>
<tr>
<th></th>
<th>HFD-Index</th>
<th>Berry-Index</th>
<th>Count-Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum HDL-cholesterol, mmol/L</td>
<td>0.169**</td>
<td>0.082**</td>
<td>0.026</td>
</tr>
<tr>
<td>Serum triacylglycerol, mmol/L</td>
<td>-0.076*</td>
<td>-0.025</td>
<td>-0.046*</td>
</tr>
<tr>
<td>Serum uric acid, μmol/L</td>
<td>-0.098*</td>
<td>-0.026</td>
<td>-0.028</td>
</tr>
<tr>
<td>Serum homocysteine, μmol/L</td>
<td>-0.065*</td>
<td>-0.034*</td>
<td>-0.105**</td>
</tr>
</tbody>
</table>

1. **Significance level 1%**, *significance level 5%.
2. Correlations based on n = 4030 sample size.
3. All analyses were adjusted for sample weights.