Clinical Nutrition ESPEN 33 (2019) 207–212

Contents lists available at ScienceDirect

Clinical Nutrition ESPEN

journal homepage: http://www.clinicalnutritionespen.com

Original article

Compositional data analysis as an alternative paradigm for nutritional studies

Maria Léa Corrêa Leite[1]

National Research Council, Institute of Biomedical Technologies, Milan, Italy

ARTICLE INFO

Article history:
Received 3 August 2018
Accepted 21 May 2019

Keywords:
Nutrient balances
Compositional data
Log-ratio transformation
Energy adjustment
Obesity

SUMMARY

Background & aim: Although the compositional nature of dietary data is well recognized, little attention has been given to the methods specifically developed for the statistical analysis of compositional data. The use of standard statistical procedures that ignore the *relative* nature of compositional elements can lead to spurious results when applied to crude data. This note proposes using a compositional data approach for the statistical analysis of nutritional data.

Methods: The analyses of data from an Italian population study of 673 non-diabetic women aged 40–74 years were based on isometric log-ratio (ilr) transformation applied to three dietary compositions: seven-part macronutrients, nine-part vitamins and six-part minerals. The ilr transformation produces new variables that represent specific contrasts (*balances*) between the compositional parts. Different sequential binary partitions have been described and used as a means of flexibly defining balances on the basis of a researcher's interest, and the new variables (that are suitable for undergoing standard statistical procedures) have been included as covariates in linear regression models in order to examine the isocaloric associations between specific dietary balances and waist circumference (WC).

Results: Regardless of the dietary compositions, total energy was positively and fibre intake was negatively associated with WC. Net of these effects and keeping constant the proportional relationships between the other dietary components, WC was negatively associated with the relative increase in starches, and positively associated with the relative niacin content. Vitamin C was negatively associated with WC, whereas folate content was positively related.

Discussion: The proposed approach allows a clear interpretation of the relative roles of different dietary components within a holistic overview of a diet. The analyses involving WC provide some useful insights.

1. Introduction

Dietary data are typical compositional data: dietary nutrients are closely inter-related and it can be intuitively understood that a variation in one entails variations in one or more of the others. This happens because nutrients are not consumed in isolation, but as parts of complex food matrices. The overall nutrient composition of a diet results from a weighted mixing of the eaten foods' nutrient compositions where the weights are modulated by the dietary pattern. These dynamics produce multi-collinearity, a well-known pitfall of the usual regression techniques including explanatory dietary data.

Collinearity occurs when independent variables are so closely correlated that it is very difficult to isolate and measure their contributions to variations in response. It leads to unstable and unreliable coefficients, and may therefore underlie the inconsistencies between the results of nutritional studies based on the effects of individual nutrients. These studies have then been discouraged in favour of dietary pattern approaches, although it is acknowledged that analyses based on individual dietary elements are the only means of elucidating the biological mechanisms underlying the associations between diet and disease risk.

When investigating diet–disease associations, nutrient interrelationships (particularly between minerals and vitamins) seem to be even more complex. The increased intake of one nutrient can influence the absorption or the use of another, thus creating

[1] Epidemiology and Biostatistics Unit, Institute of Biomedical Technologies, National Research Council, Via Fratelli Cervi 93, 20090 Segrate, MI, Italy. Fax: +39 0226422770.

E-mail address: lea.correa@itb.cnr.it.

https://doi.org/10.1016/j.clnesp.2019.05.011

network of antagonisms and synergies at both metabolic and absorptive levels. It is also known that metabolic dysfunctions may not only be due to a deficiency or excess of a particular nutrient, but are also often due to a loss of homeostatic equilibrium between nutrients.

The natural consequence is the idea that "proportionalities" (ratios) between nutritional elements may be more informative than their absolute values when evaluating the role of nutrients in disease risk. This is not a new concept, but it is only recently that it has been suggested to use it in nutritional studies, and the thought that working with ratios "may bridge the gap between single nutrient-based information and dietary patterns" has also been stressed [1]. However, little attention has been given to the methodological development of techniques for ratio-based compositional data analysis, and their actual use in nutritional research.

The aim of this note is to extend our previous discussion concerning the use of a compositional data approach to nutritional studies involving the macronutrient sources of energy [2,3] by including the composition of micronutrients (vitamins and minerals), and illustrating the approach in the context of the relationship between diet and abdominal obesity.

2. Methods

2.1. Compositional data, the isometric log-ratio (ilr) transformation

Compositional data refer to any set of components that constitute parts of a whole, and their relative structure represents the essential basis for the analysis [4]. Such data are often constrained by a constant sum and are expressed as proportions or percentages, but their definition has now been extended to emphasize the key attribute characterizing these data: any relevant compositional information involves the ratios between its parts rather than its isolated absolute values [5].

The interdependence of the compositional parts and their *relative* nature has important implications for statistical analysis because the use of standard statistical procedures originally conceived for unconstrained data and absolute scale values may lead to misinterpretations and spurious inferences when applied to crude compositional data. The concept that compositions provide information about relative rather than absolute values has led to the development of particular methods based on logarithms of ratios between parts (or groups of parts), which are a suitable means of transforming compositional data in such a way as to allow the use of standard statistical methods [4,5].

The family of log-ratio transformations includes the ilr transformation, which leads to the definition of new variables that have the property of representing groups of parts and the relationships between them, and can therefore be interpreted as particular balances. One of the procedures used to construct balances is sequential binary partition (SBP), which consists of successively and hierarchically splitting the parts of a composition. First of all, the full composition is divided into two non-overlapping groups of parts, each of which is then similarly divided until all of the groups have only a single part. Given a composition of D parts, one balance is defined for each of the D−1 steps required to complete the partition as follows:

$$\mathrm{Ilr}_i = \frac{\sqrt{rs}}{\sqrt{r+s}} \ln \frac{\left(y_{j1} \cdot y_{j2} \cdot \ldots \cdot y_{jr}\right)^{1/r}}{\left(y_{k1} \cdot y_{k2} \cdot \ldots \cdot y_{ks}\right)^{1/s}}, \quad \text{for } i = 1, \ldots, D-1$$

where r and s are respectively the number of parts in the first (y_j) and second group (y_k) at each order of partition *i* [4].

The balances represent the ratios of the central values (the geometric mean) of groups of parts, the logarithm provides the appropriate scale, and the square root coefficient is a standardising constant that allows different balances to be compared.

There are various ways of creating SBPs, and particular system of balances can be defined to answer specific research questions. For example, Table 1 shows two alternative partitions applied to a six-part dietary mineral composition. Each partition produces five balances that can be included as covariates in a regression equation. The balances correspond to specific contrasts between the compositional parts, and they represent the decomposition of the total variance of the full composition. Although interest may be restricted to just some of them (as may be the case of the three last balances of SBP_1), they are all simultaneously included in the regression equation in order to assure that all of the variability attributable to the original parts is represented in the model.

One important characteristic of SPB_2 in Table 1 is that the first balance (ilr_1) captures all of the relevant information concerning the first part (calcium) compared to the rest and, exploiting this property, a procedure based on such a partition has been proposed as a means of addressing distinct questions relating to the role of each of the parts of a composition [6].

On the basis of this choice, the formula above can be rewritten more simply as:

$$\mathrm{Ilr}_i = c_i^* \ln (r_i),$$

$$\text{where } c_i = \frac{\sqrt{(D-i)}}{\sqrt{(D-i+1)}}, r_i = \frac{y_i}{g(y_{i+1}, \ldots, y_D)}, \quad \text{for } i = 1, \ldots, D-1$$

and g(.) is the geometric mean.

The procedure [6] consists in constructing D different partitions, each with D−1 balances, by permutating the compositional parts in such a way that the first balances explain one of the parts at a time. The D−1 balances are included as covariates in a regression model that is run D times and, at each run, inferences are focused on the parameters corresponding only to ilr_1. As the different partitions are orthogonal (i.e. they are linearly independent), all of the other covariate terms remain exactly the same across the different runs [6].

Although it requires a more work, the procedure allows full inferences to be made about the role of each dietary component. Table 2 shows the formulae for computing the balances in the case of a six-part dietary mineral composition.

The examples shown in the next section use data from the Italian Bollate Eye Study [7], a population-based study of subjects aged 40–74 years that was carried out in 1992–1993. The analyses involved 673 non-diabetic women whose usual food intake was assessed by means of a food frequency questionnaire, and whose mean daily nutrient intakes were calculated using the food compositional database compiled for epidemiological studies in Italy [8]. Linear regression models were used to examine the associations between the dietary macro- and micronutrients and waist circumference (WC), which was a mean (standard deviation) of 87.17 cm (12.18).

Data from the following three dietary compositions were available:

1) a seven-part macronutrient composition (mean daily kcal) consisting of animal proteins (281.55), vegetable proteins (156.40), starches (769.55), simple sugars (691.58), saturated fats (268.92), monounsaturated fats (331.92), and polyunsaturated fats (119.03).

Table 1
Sequential binary partitions (SBP) of a composition of six dietary minerals. At each order of partition, +1 means that the part is assigned to the first group, −1 means that it is assigned to the second group, and 0 means that it is not involved in the partition at this order.

Order	Minerals						r	s	Balances (ilr_i)
	Calcium (Ca)	Phosphorous (P)	Sodium (Na)	Potassium (K)	Iron (Fe)	Zinc (Z)			
SBP₁									
1	+1	+1	−1	−1	−1	−1	2	4	$(8/6)^{1/2}*\ln[(Ca*P)^{1/2}/(Na*K*Fe*Z)^{1/4}]$
2	0	0	+1	+1	−1	−1	2	2	$(4/4)^{1/2}*\ln[(Na \times K)^{1/2}/(Fe \times Z)^{1/2}]$
3	+1	−1	0	0	0	0	1	1	$(1/2)^{1/2}*\ln(Ca/P)$
4	0	0	+1	−1	0	0	1	1	$(1/2)^{1/2}*\ln(Na/K)$
5	0	0	0	0	+1	−1	1	1	$(1/2)^{1/2}*\ln(Fe/Z)$
SBP₂									
1	+1	−1	−1	−1	−1	−1	1	5	$(5/6)^{1/2}*\ln[Ca/(P*Na*K*Fe*Z)^{1/5}]$
2	0	+1	−1	−1	−1	−1	1	4	$(4/5)^{1/2}*\ln[P/(Na*K*Fe*Z)^{1/4}]$
3	0	0	+1	−1	−1	−1	1	3	$(3/4)^{1/2}*\ln[Na/(K*Fe*Z)^{1/3}]$
4	0	0	0	+1	−1	−1	1	2	$(2/3)^{1/2}*\ln[K/(Fe*Z)^{1/2}]$
5	0	0	0	0	+1	−1	1	1	$(1/2)^{1/2}*\ln(Fe/Z)$

Table 2
Formulae for computing the isometric log-ratios (ilr) of a six-part dietary mineral composition. Five (i) balances are computed within each of the six (j) permutations of the compositional parts.

$r_i^{(j)}$	$Ilr_1^{(j)} = (5/6)^{1/2}*\ln(r_1^{(j)})$	$Ilr_2^{(j)} = (4/5)^{1/2}*\ln(r_2^{(j)})$	$Ilr_3^{(j)} = (3/4)^{1/2}*\ln(r_3^{(j)})$	$Ilr_4^{(j)} = (2/3)^{1/2}*\ln(r_4^{(j)})$	$Ilr_5^{(j)} = (1/2)^{1/2}*\ln(r_5^{(j)})$
$r_i^{(1)}$	$r_1^{(1)} = Ca/(P*Na*K*Fe*Z)^{1/5}$	$r_2^{(1)} = P/(Na*K*Fe*Z)^{1/4}$	$r_3^{(1)} = Na/(K*Fe*Z)^{1/3}$	$r_4^{(1)} = K/(Fe*Z)^{1/2}$	$r_5^{(1)} = (Fe/Z)$
$r_i^{(2)}$	$r_1^{(2)} = P/(Ca*Na*K*Fe*Z)^{1/5}$	$r_2^{(2)} = Ca/(Na*K*Fe*Z)^{1/4}$	$r_3^{(2)} = Na/(K*Fe*Z)^{1/3}$	$r_4^{(2)} = K/(Fe*Z)^{1/2}$	$r_5^{(2)} = (Fe/Z)$
$r_i^{(3)}$	$r_1^{(3)} = Na/(Ca*P*K*Fe*Z)^{1/5}$	$r_2^{(3)} = Ca/(P*K*Fe*Z)^{1/4}$	$r_3^{(3)} = P/(K*Fe*Z)^{1/3}$	$r_4^{(3)} = K/(Fe*Z)^{1/2}$	$r_5^{(3)} = (Fe/Z)$
$r_i^{(4)}$	$r_1^{(4)} = K/(Ca*P*Na*Fe*Z)^{1/5}$	$r_2^{(4)} = Ca/(P*Na*Fe*Z)^{1/4}$	$r_3^{(4)} = P/(Na*Fe*Z)^{1/3}$	$r_4^{(4)} = Na/(Fe*Z)^{1/2}$	$r_5^{(4)} = (Fe/Z)$
$r_i^{(5)}$	$r_1^{(5)} = Fe/(Ca*P*Na*K*Z)^{1/5}$	$r_2^{(5)} = Ca/(P*Na*K*Z)^{1/4}$	$r_3^{(5)} = P/(Na*K*Z)^{1/3}$	$r_4^{(5)} = Na/(K*Z)^{1/2}$	$r_5^{(5)} = (K/Z)$
$r_i^{(6)}$	$r_1^{(6)} = Z/(Ca*P*Na*K*Fe)^{1/5}$	$r_2^{(6)} = Ca/(P*Na*K*Fe)^{1/4}$	$r_3^{(6)} = P/(Na*K*Fe)^{1/3}$	$r_4^{(6)} = Na/(K*Fe)^{1/2}$	$r_5^{(6)} = (K/Fe)$

Ca (calcium), P (phosphorus), Na (sodium), K (potassium), Fe (iron), Z (zinc).

2) a nine-part vitamin composition (mean daily mg) consisting of thiamin (1.49), riboflavin (2.27), niacin (25.75), vitamin B_6 (2.97), vitamin C (321.88), folate (0.471), beta-carotene (8.00), vitamin D (0.0036), and vitamin E (12.72).
3) a six-part mineral composition (mean daily mg) consisting of calcium (1300.35), phosphorus (1825.75), sodium (5298.85), potassium (5900.63), iron (18.63), and zinc (15.56).

A total of 19 dietary balances (the ilr-transformed variables) were computed (6 for macronutrients, 8 for vitamins and 5 for the mineral composition). In addition to total daily energy (kcal) and fibre intakes, the analyses also took into account demographic variables: age (continuous), education (above high school, high school and primary school or less), lifestyles: smoking (no, ex-smokers and currently smokers), alcohol consumption (abstainers, <30.0 g/day, ≥30.0 g/day) and physical activity: sport practice (yes/no), walking and/or bicycling at least 15 min/day (yes/no), tv watching, hours/day (−2, 2−3, 4+). No significant heterogeneity in the relationships between dietary balances and WC was observed across the strata of the covariates.

Thus, the generic model for WC including kcal, fibre, 19(I) dietary balances and 11(K) control variables terms for the jth subject is

$$WC_j = A + B_{kcal}\,kcal_j + B_{fibre}fibre_j + \sum\nolimits_i B_i ilr_{ij} + \sum\nolimits_k Z_k covariate_{kj} + e_j,$$

where A is the constant and e_j is the error term.

3. Results

3.1. Macro- and micronutrient balances and abdominal obesity

Table 3 shows the coefficients expressing the strength of the association between WC and increases in the specific dietary balances defined within the three compositions of macronutrients, vitamins and minerals. These partitions are examples of particular choices from among the many other possibilities (for the mineral composition, the partition corresponds to SBP_1 in Table 1). What

Table 3
Linear regression coefficients (B) and standard errors (ES) of waist circumference (cm) in relation to dietary components: total daily energy and fibre intakes, and the dietary balances derived by applying the isometric log-ratio transformation to the macro- and micronutrient compositions. Women of the Bollate Eye Study (1992–93).

Dietary components and balances	B (ES)
Total energy (100 kcal)	0.20 (0.09)*
Fibre (10 g)	−1.38 (0.65)*
Macronutrient balances	
Proteins & Carbohydrates vs. Fats	0.47 (1.99)
Proteins vs. Carbohydrates	8.95 (3.98)*
Animal proteins vs. Vegetable proteins	−4.54 (4.04)
Starches vs. Simple sugars	−6.15 (2.37)**
Saturated fats vs. Unsaturated fats	−0.99 (2.85)
Monounsaturated fats vs. Polyunsaturated fats	−5.15 (2.50)*
Vitamin balances	
B-vitamins vs. Vit D & Antioxidant vitamins	4.86 (2.12)*
B_1, B_2, B_3 vs. B_6 & Folates	3.70 (4.79)
B_1 & B_2 vs. B_3	−11.47 (3.26)***
B_1 vs. B_2	2.72 (4.58)
B_6 vs. Folates	−11.30 (4.36)**
Vit D vs. Antioxidants (Vit E, Vit C, Beta-carotene)	2.18 (1.28)
Vit E vs. Vit C & Beta-carotene	−0.13 (2.82)
Vit C vs. Beta-carotene	−3.33 (1.54)*
Mineral balances	
Ca & P vs. Na & K & Fe & Z	−3.46 (4.33)
Na & K vs. Fe & Z	2.78 (4.32)
Ca vs. P	7.53 (5.63)
Na vs. K	−2.83 (3.41)
Fe vs Z	−6.36 (4.32)

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. $F(32,640) = 5.16$, $p = 0.000$; $R^2 = 0.205$.
Ca (calcium), P (phosphorus), Na (sodium), K (potassium), Fe (iron), Z (zinc).
The model also includes terms for age, education, practising sports, walking/cycling, television watching, smoking, and alcohol consumption.

matters is that their hierarchical nested property was respected. The estimates were mutually adjusted, which means that they were made while keeping the proportional relations between all the other dietary elements unchanged. It can therefore be seen that, net of the effect of the variation in the proportions of the macro- and micronutrients (i.e., the qualitative aspect of the diet), total energy intake was positively and total fibre intake was negatively associated with WC. The "proteins *vs* carbohydrates" balance was positively related to WC, and the "starches *vs* simple sugars" balance (within carbohydrates) was negatively associated; there was no association with the "saturated *vs* unsaturated fats" balance, whereas that of "monounsaturated *vs* polyunsaturated fats" was inversely related; and the balance between the set of B vitamins and the other vitamins was positively related, but there were strong negative associations with the balances "thiamin (B_1) and riboflavin (B_2) *vs* niacin (B_3)" and "vitamin B6 *vs* folates".

Table 4 shows the coefficients relating the variations in WC to the increases in the first balances of the SBPs shown in Table 2 for the compositions of minerals, and made in similar way for the macronutrient and vitamin compositions. Each coefficient therefore estimates the effect of increases in the "weight" of each dietary component within its composition, while assuring that the estimate relates to all of the information concerning that compositional part in the data. This information maintains its *relative* nature, and it is related to a specific compositional context, since the balances correspond to the contrasts of each compositional element versus all the other parts in the composition. Model 1 included total energy, fibre intake and the macronutrient balances, and model 2 further included the micronutrient balances. Note that the protein coefficients in model 2 lost their statistical significance, and that the inverse relationship between WC and the content of simple sugars (mono- and disaccharides, the main sources of which were fruit and vegetables) practically disappeared, thus indicating that these effects were confounded by the dietary vitamin and mineral contents in model 1. As expected, in the complete model, the total energy and fibre intake estimates were exactly the same as those shown in Table 3. Keeping total energy and fibre intake constant, and maintaining the proportional relationships between the other dietary components, abdominal obesity was negatively associated with the relative increase in starches, and positively associated with the niacin content of the diet. Furthermore, vitamin C was negatively associated with abdominal obesity, whereas folate content seemed to be positively related to it.

Technically, nine runs were required to obtain the estimates of the complete model of which, given their orthogonal property, those regarding one composition remain unchanged regardless of the partition applied to the other compositions. The permutations of the compositional parts were therefore simultaneously made in all three compositions. Also as a result of the orthogonal nature of the different partitions, the constant and other covariates terms, as well as the quality of fit of the regression models, are exactly the same, regardless of the choice made. The examination of the residual plots did not show any particular trend in their distribution and did not reveal the presence of any outliers, thus indicating that the model is adequately specified (data not shown).

Table 4
Linear regression coefficients (B) and standard errors (ES) of waist circumference (cm) in relation to dietary components: total daily energy and fibre intakes, and the dietary balances[a] derived by applying the isometric log-ratio transformation to the macro- and micronutrient compositions. Women of the Bollate Eye Study (1992–93).

Dietary components and balances	Macronutrient models[b]	Complete models[c]
	B (ES)	B (ES)
Total energy (100 kcal)	0.19 (0.08)*	0.20 (0.09)*
Fibre (10 g)	−1.71 (0.53)**	−1.38 (0.65)*
Animal proteins	3.22 (1.47)*	1.53 (2.70)
Vegetable proteins	13.83 (3.82)***	8.47 (4.95)
Starches	−10.08 (2.68)***	−9.37 (3.21)**
Simple sugars	−4.20 (1.34)**	0.03 (1.98)
Saturated fats	−2.18 (2.42)	−1.09 (2.63)
Monounsaturated fats	−3.12 (2.44)	−3.72 (2.51)
Polyunsaturated fats	2.53 (1.60)	4.15 (2.49)
Vitamin B_1 (Thiamin)		0.04 (3.98)
Vitamin B_2 (Riboflavin)		−4.04 (4.03)
Vitamin B_3 (Niacin)		12.90 (3.62)***
Vitamin B_6		−9.09 (5.04)
Folates		7.86 (3.79)*
Vitamin D		0.08 (0.98)
Vitamin E		−2.70 (3.11)
Vitamin C		−5.03 (1.84)**
Beta-carotene		−0.03 (1.51)
Calcium		3.65 (2.92)
Phosphorus		−8.02 (6.67)
Sodium		0.42 (1.20)
Potassium		4.81 (5.13)
Iron		−5.36 (4.20)
Zinc		4.50 (4.57)

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.
The models also include terms for age, education, practising sports, walking/cycling, television watching, smoking, and alcohol consumption.
[a] The balances represent the contrasts of each nutrient in the first column versus the other nutrients within their respective compositions.
[b] $F(19,653) = 6.83$, $p = 0.000$; $R^2 = 0.166$ (the same in the 7 runs).
[c] $F(32,640) = 5.16$, $p = 0.000$; $R^2 = 0.205$ (the same in the 9 runs).

4. Discussion

This note describes an alternative approach to the statistical analysis of dietary data that makes it possible to evaluate the effect of specific dietary balances within an overview of the diet as a whole. Moreover, the effects of macro- and micronutrient balances can be estimated in isocaloric analyses, and meaningful estimates can be made of the effect of total energy intake. When interpreting the results, it is important to bear in mind that the regression coefficients reported in Tables 3 and 4 represent the relationship of WC with specific dietary balances, each of which expresses peculiar dynamics. For example, an increase of the balance between the dietary content of proteins against that of carbohydrates (positively associated with WC, Table 3) can be due to either the increase of the protein contents, or the decrease of those of carbohydrates, or due to both. Furthermore, a possible decrease of carbohydrate contents could be due – in different degrees – to the reduction of the starch or simple sugar contents, and this would work as a potential confounding factor. The inclusion of terms for the "starches vs. sugars" and the "animal proteins vs. vegetable proteins" balances in the model assures that the evaluation of the effect related to variations in the "proteins vs. carbohydrates" balance is done in a *balanced* way. Indeed, the mutual adjustment of the balance effect estimates implies that each of them is estimated while keeping constant all the other dietary ratios included in the model, thus maintaining unchanged the proportional relationships between the other nutrients. Balances as those indicated in Table 3 can be flexibly constructed in order to answer the desired research question. This flexibility can be exploited by an analyst as a means to improve the interpretability of the balance effects. When considering the macronutrient composition, the balances can be easily defined insofar they follow the natural clustering of the subgroups of parts (e.g. starches and sugars, saturated and unsaturated fats) in groups of parts (e.g. carbohydrates and fats). In the case of the micronutrient compositions, the most appropriate way of performing the hierarchical splitting into subgroups of parts may appear less clear. The choice in this case will mostly depend on the researcher's expertise and specific study question. On the other hand, the results shown in Table 4 come from an alternative way of defining balances, and each one of the coefficients is related to the first ilr of sequential partitions as the SBP_2 in Table 1. This approach, which involves the permutation of the compositional parts and repeated runs, has been proposed as a means to deal with explanatory compositional variables when the interest is to address distinct questions about each of the compositional parts [6]. However, the procedure may turn out to be somewhat complicated. The coefficients represent the contrasts between each nutrient and the remaining parts of its composition. They therefore are not related to the isolated variation of single elements, but rather reflect the effects of the increase in the contents of each nutrient in relation to the other compositional parts whose contents proportionately decrease.

In our data, total fibre was expressed as the sum of cellulose, hemicellulose, pectins, gum and lignin, and its intake (like that of starches) inversely correlated with abdominal obesity in women. It has long been widely agreed that fibre-rich diets play a role in preventing obesity and metabolic diseases [9,10], but the relationship between the consumption of resistant starch and reduced abdominal fat and improved insulin sensitivity is a more recent discovery [11]. Fibre and resistant starches are poorly digestible carbohydrates as they escape enzymatic degradation in the small intestine and are fermented by resident micro-organisms in the large intestine. Interest in the role of intestinal microbiota has increased over recent years because it is now known that they mediate the effect of diet on obesity by regulating host metabolism [12,13], and there is growing evidence that this mediation involves various metabolites, particularly the organic acids released as a result of microbial fermentative activity in the gut [14]. It has also been shown that short-chain fatty acids (SCFAs), the principal products of carbohydrate fermentation in the large intestine, have beneficial effects on host metabolism and protect against diet-induced obesity [15,16].

We observed a strong positive relationship between the relative dietary content of niacin (nicotinic acid and nicotinamide) and abdominal obesity. This is an important finding because it has been suggested that the anti-lipidemic effect of niacin could help to prevent cardiovascular disease (CVD) [17], although, a recent review of randomised controlled trials has concluded that this is unlikely to be the case [18] and, as nicotinic acid is associated with an increased risk of impaired fasting glucose tolerance and can cause hyperuricemia, there is controversy concerning its role in lipid management and CVD prevention [19]. Moreover, a large population study has shown that N(1)-methylnicotinamide (meNAM), a direct metabolite of nicotinamide, is closely associated with obesity and diabetes [20], and a population-based ecological study has warned about the possible role of excess niacin consumption in the increased prevalence of both [21].

Vitamin C consumption negatively related to abdominal obesity, which is in line with the findings of other studies [22]. However, somewhat surprisingly, the relative content of dietary folates was positively associated with WC, whereas vitamin B_6 was marginally inversely related ($p = 0.072$). It has been previously hypothesised that these two vitamins may play a role in preventing CVD because of their homocysteine lowering effect, but the results of clinical trials have been inconsistent [23]. Folate intake and folate status are involved in lipid metabolism, obesity and the metabolic syndrome [24], and animal studies have shown that an excess of folic acid may exacerbate weight gain [25]. However, recent studies have paradoxically revealed that obese subjects often have low serum folate levels, whereas their red blood cell folate concentrations (which are known to reflect long-term average consumption) are significantly higher than those of non-obese subjects [26].

In conclusion, the proposed analytical approach allows a clear and intuitive interpretation of the different effects of dietary components on variations in WC, and seems to provide some insights that may be useful in metabolomic studies. In terms of abdominal obesity in women, our findings suggest that plant-based diets owe their good reputation to their fibre, starch and vitamin C content, and that the potentially deleterious effects of animal-based diets may particularly be attributable to their niacin content.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The author declares that she has no conflict of interest.

References

[1] Kelly OJ, Gilman JC, Ilich JZ. Utilizing dietary micronutrient ratios in nutritional research may be more informative than focusing on single nutrients. Nutrients 2018. https://doi.org/10.3390/nu10010107 [published online Jan 19].
[2] Correa Leite ML. Applying compositional data methodology to nutritional epidemiology. Stat Methods Med Res 2016;25:3057–65.
[3] Correa Leite ML, Prinelli F. A compositional data perspective on studying the associations between macronutrient balances and diseases. Eur J Clin Nutr 2017;71:1365–9.
[4] Pawlowsky-Glahn V, Egozcue JJ, Tolosana-Delgado R. Modeling and analysis of compositional data. Chichester: Wiley; 2015.
[5] Pawlowsky-Glahn V, Egozcue JJ. Compositional data and their analysis: an introduction. In: Buccianti A, Mateu-Figueras G, Pawlowsky-Glahn V, editors. Compositional data analysis in the geosciences: from theory to practice. London: Geological Society (Special Publications 264); 2006. p. 1–10.
[6] Hron K, Filzmoser P, Thompson K. Linear regression with compositional explanatory variables. J Appl Stat 2012;39:1115–28.
[7] Correa Leite ML, Nicolosi A. Dietary patterns and metabolic syndrome factors in a non-diabetic Italian population. Public Health Nutr 2009;12:1494–503.
[8] Salvini S, Parpinel M, Gnagnarella P, Maisonneuve P, Turrini A. Banca dati di composizione degli alimenti per studi epidemiologici in Italia. Milan: Istituto Europeo di Oncologia; 1998.
[9] Slavin JL. Dietary fiber and body weight. Nutrition 2005;21:411–8.
[10] Du HD, van der ADL, Boshuizen HC, Forouhi NG, Wareham NJ, Halkjaer J, et al. Dietary fiber and subsequent changes in body weight and waist circumference in European men and women. Am J Clin Nutr 2010;91:329–36.
[11] Keenan MJ, Zhou J, Hegsted M, Pelkman C, Durham HA, Coulon DB, et al. Role of resistant starch in improving gut health, adiposity, and insulin resistance. Adv Nutr 2015;6:198–205.
[12] Delzenne NM, Cani PD. Interaction between obesity and the gut microbiota: relevance in nutrition. Annu Rev Nutr 2011;31:15–31.
[13] Shen J, Obin MS, Zhao L. The gut microbiota, obesity and insulin resistance. Mol Aspects Med 2013;34:39–58.
[14] Koh A, De Vadder F, Kovatcheva-Datchary P, Backhed F. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. Cell 2016;165:1332–45.
[15] Kasabuchi M, Hasegawa S, Hiramatsu T, Ichimura A, Kimura I. Dietary gut microbial metabolites, short-chain fatty acids, and host metabolic regulation. Nutrients 2015;7:2839–49.
[16] Rios-Covian D, Ruas-Madiedo P, Margolles A, Gueimonde M, de los Reyes-Gavilan CG, Salazar N. Intestinal short chain fatty acids and their link with diet and human health. Front Microbiol 2016. https://doi.org/10.3389/fmicb.2016.00185 [published online Feb 17].
[17] Lavigne PM, Karas RH. The current state of niacin in cardiovascular disease prevention: a systematic review and meta-regression. J Am Coll Cardiol 2013;61:440–6.
[18] Schandelmaier S, Briel M, Saccilotto R, Olu KK, Arpagaus A, Hemkens LG, et al. Niacin for primary and secondary prevention of cardiovascular events. Cochrane Database Syst Rev 2017. https://doi.org/10.1002/14651858.CD009744.pub2 [published online Jun 14].
[19] Whayne Jr TF. Nicotinic acid: current status in lipid management and cardiovascular disease prevention. Angiology 2014;65:557–9.
[20] Liu M, Li L, Chu J, Zhu B, Zhang Q, Yin X, et al. Serum N(1)-methylnicotinamide is associated with obesity and diabetes in Chinese. J Clin Endocrinol Metab 2015;100:3112–7.
[21] Zhou SS, Li D, Zhou YM, Sun WP, Liu QG. B-vitamin consumption and the prevalence of diabetes and obesity among the US adults: population based

ecological study. BMC Public Health 2010. https://doi.org/10.1186/1471-2458-10-746 [published online Dec 2].
[22] Canoy D, Wareham N, Welch A, Bingham S, Luben R, Day N, et al. Plasma ascorbic acid concentrations and fat distribution in 19 068 British men and women in the European Prospective Investigation into Cancer and Nutrition Norfolk cohort study. Am J Clin Nutr 2005;82:1203–9.
[23] Ntaios G, Savopoulos C, Grekas D, Hatzitolios A. The controversial role of B-vitamins in cardiovascular risk: an update. Arch Cardiovasc Dis 2009;102:847–54.
[24] da Silva RP, Kelly KB, Al Rajabi A, Jacobs RL. Novel insights on interactions between folate and lipid metabolism. Biofactors 2014;40:277–83.
[25] Kelly KB, Kennelly JP, Ordonez M, Nelson R, Leonard K, Stabler S, et al. Excess folic acid increases lipid storage, weight gain, and adipose tissue inflammation in high fat diet-fed rats. Nutrients 2016. https://doi.org/10.3390/nu8100594 [published online Sep. 23].
[26] Zhou S, Li D, Chen N, Zhou Y. Vitamin paradox in obesity: deficiency or excess? World J Diabetes 2015;6:1158–67.