

## Longitudinal analysis of dietary patterns in Chinese adults from 1991 to 2009

Carolina Batis<sup>1</sup>, Daniela Sotres-Alvarez<sup>2</sup>, Penny Gordon-Larsen<sup>1</sup>, Michelle A. Mendez<sup>1</sup>, Linda Adair<sup>1</sup> and Barry Popkin<sup>1\*</sup>

<sup>1</sup>Department of Nutrition, Carolina Population Center, University of North Carolina at Chapel Hill, 123 West Franklin Street, Chapel Hill, NC 27516, USA

<sup>2</sup>Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

(Submitted 2 April 2013 – Final revision received 17 October 2013 – Accepted 28 October 2013 – First published online 13 December 2013)

### Abstract

In the present study, we aimed to identify the changes or stability in the structure of dietary patterns and tracking, trends and factors related to the adherence to these dietary patterns in China from 1991 to 2009. We analysed dietary data collected during seven waves of the China Health and Nutrition Survey and included 9253 adults with complete dietary data for three or more waves. Dietary intake assessment was carried out over a 3 d period with 24 h recalls and a household food inventory. Using factor analysis in each wave, we found that the structure of the two dietary patterns identified remained stable over the studied period. The traditional southern dietary pattern was characterised by high intakes of rice, fresh leafy vegetables, low-fat red meat, pork, organ meats, poultry and fish/seafood and low intakes of wheat flour and maize/coarse grains and the modern high-wheat dietary pattern was characterised by high intakes of wheat buns/breads, cakes/cookies/pastries, deep-fried wheat, nuts/seeds, starchy root/tuber products, fruits, eggs/egg products, soya milk, animal-based milk and instant noodles/frozen dumplings. Temporal tracking (maintenance of a relative position over time) was higher for the traditional southern dietary pattern, whereas adherence to the modern high-wheat dietary pattern had an upward trend over time. Higher income, education and urbanicity levels were positively associated with both the dietary patterns, but the association became weaker in the later years. These results suggest that even in the context of rapid economic changes in China, the way people chose to combine their foods remained relatively stable. However, the increasing popularity of the modern high-wheat dietary pattern, a pattern associated with several energy-dense foods, is a cause of concern.

**Key words:** Dietary patterns: Factor analysis: China: Longitudinal dietary patterns

China is one of the countries with the most rapid increases in the prevalence of chronic diseases; the prevalence of overweight/obesity (BMI  $\geq 25$  kg/m<sup>2</sup>) among adults was 15% in 1992, 22% in 2002<sup>(1)</sup> and 33% in 2007–8<sup>(2)</sup>. The prevalence of diabetes, hypertension and dyslipidaemia reached 10, 27 and 65%, respectively, in 2007–8<sup>(2)</sup>. Due to the key role played by the diet in all these chronic diseases, it is important to better understand the eating behaviour of the Chinese population.

Over the last two decades, many important changes have been identified in the diet of the Chinese population. Eating behaviours such as snacking have emerged and continue to grow, and cooking methods have also changed from predominantly steaming and boiling to frying<sup>(3,4)</sup>. A marked increase in dietary diversity has also been observed, with more people consuming foods from more number of food groups<sup>(5)</sup>. The intake of vegetables, fruits, cakes, and milk and other animal products such as pork, poultry and eggs has increased, whereas that of cereals and tubers has decreased<sup>(6)</sup>.

In addition to the study on the number of meals consumed, cooking methods or food group consumption, data-driven dietary pattern analyses, such as factor or cluster analyses, are also useful for studying dietary patterns. Dietary patterns more closely resemble actual eating behaviours because multiple food groups instead of single food groups or nutrients are considered during analyses, and they give insights into how people eat by identifying the foods that are eaten in combination<sup>(7–9)</sup>. Previous studies conducted in a Chinese national survey in 2002 have identified several dietary patterns such as the ‘Yellow Earth’ or ‘traditional northern’ pattern, which is high in wheat, wheat products, maize, sorghum and tubers; the ‘green water’ or ‘traditional southern’ pattern, which is high in rice, vegetables, seafood, pork and poultry; and the ‘Western’ or ‘new affluence’ pattern, which is high in beef, lamb, milk, cheese, yogurt, cakes, juices and nuts<sup>(10–12)</sup>. However, to the best of our knowledge, no longitudinal analysis of dietary patterns over time has been conducted in the Chinese population.

**Abbreviations:** CHNS, China Health and Nutrition Survey; INFS, Chinese Institute of Nutrition and Food Safety; UNC-CH, University of North Carolina at Chapel Hill.

\* **Corresponding author:** B. Popkin, fax +1 919 966 9159, email popkin@unc.edu

To fill this gap, in the present study, we used measurements of dietary intake recorded during seven occasions over the course of 18 years in the China Health and Nutrition Survey (CHNS). When examining dietary patterns over time, there are two different dimensions one can look at: one is changes or stability in the general structure of the dietary patterns (are foods combined differently in each year?) and the other is the individuals' level of adherence to each dietary pattern over time. Therefore, first, we aimed to derive the dietary patterns independently at each point to determine whether the structure of dietary patterns has changed or remained stable. Second, we assessed tracking, trends over time and socio-demographic characteristics associated with the adherence to each dietary pattern.

## Methods

### *Study design and participants*

The CHNS is an ongoing study in which detailed information on income, employment, education, demographics, health and diet is collected. The survey was designed to examine across space and time how economic and social changes are associated with a range of health behaviours. A multi-stage, random-cluster process was used to draw samples in nine provinces. Survey protocols and instruments and the process followed for obtaining informed consent for the present study were approved by the institutional review committees of the University of North Carolina at Chapel Hill (UNC-CH) and the Chinese Institute of Nutrition and Food Safety (INFS), China Center for Disease Control and Prevention. The participants provided written informed consent. Additional details about the CHNS data have been provided elsewhere<sup>(13)</sup>.

Surveys were conducted in 1989, 1991, 1993, 1997, 2000, 2004, 2006 and 2009. We used data from 1991 to 2009 because only adults aged 20–45 years were included in 1989. In all the waves of the CHNS, identical clinical, dietary and anthropometric data were obtained from each household member. In each wave, we included all adults aged 18–65 years with complete dietary data for at least three waves ( $n$  9253); from these, 20% had complete dietary for all the seven waves, 50% for five or more waves, and 75% for four or more waves.

### *Dietary assessment and food grouping*

In the CHNS, dietary assessment is a combination of three consecutive 24 h recalls at the individual level and a food inventory at the household level carried out over the same 3 d period. The three consecutive days were randomly allocated to start from Monday to Sunday. For the food inventory, all the foods available in the household (purchased, stored or home produced) were measured on a daily basis with a Chinese balance (1991–7) or digital scales (2000–9). The changes in the household food inventory as well as wastage were used to estimate total household food consumption. For the 24 h recalls, trained interviewers recorded the types and amounts of foods consumed and the type of meals consumed and place of meal consumption. For dishes prepared at home, the amount of

each dish was estimated from the household food inventory, based on the proportion of each dish that the person reported to have consumed.

The food groups included in the present analysis were based on a food grouping system developed specifically for the CHNS by researchers from the UNC-CH and INFS<sup>(5)</sup>; this system separates foods into nutritionally and behaviourally meaningful food groups. We did not include alcoholic beverages, as mostly males consume them, with females consuming very low amounts. Further description of the food group classification that we used is given in Table S1 (available online).

### *Statistical analyses*

Because we included subjects who were 18–65 years old during each wave, it was possible for subjects to be excluded during certain waves (e.g. a subject aged 64 years in 2000 would be excluded in 2004 and afterwards or a subject aged 14 years in 1991 would be included only when he or she became 18 years in 1993). This exclusion criterion resulted in the inclusion of younger subjects in the later waves. In addition, the proportion of subjects in the sample residing in the North region had increased since 2000. This was because in 1997, a province in the North region was unable to participate and a substitute province was included in the study, and in 2000, both the original and substitute provinces were included again. Therefore, all the analyses were adjusted by geographical region and age in 1991 (equivalent to adjusting by birth year), so that the dietary trends found over time were not related to these sample distribution changes in age and region.

Most of the food groups had a high proportion of non-consumers, possibly because dietary intake was measured over a 3 d period. Therefore, we categorised consumption as a binary variable (non-consumers *v.* consumers) for food groups with <25% of consumers in all the waves or as a three-level variable (non-consumers, consumers with dietary intake below or above the median estimated using data collected during all waves).

We carried out an exploratory factor analysis for categorical variables using the robust weighted least-squares estimator in Mplus 6.1 (Muthén & Muthén) during each wave. Factor analysis on ordinal variables is carried out with a polychoric correlation matrix. Conceptually, ordinal variables have an underlying continuous normally distributed variable, so thresholds for the levels of the categorical variables are estimated. These thresholds are normal *z*-scores corresponding to the cumulative proportion of subjects in each category (i.e. for a binary variable (non-consumers *v.* consumers) with a 0.87 proportion of non-consumers, the threshold between the two categories would be equal to the corresponding normal *z*-score 1.13)<sup>(14,15)</sup>.

Food groups with  $\leq 5\%$  of consumers in all the waves were not included in the factor analysis. These food groups are uninformative because of a lack of variability in their intake in the population and also because they could produce bivariate tables with empty cells and affect the polychoric correlation.

Based on eigenvalues  $> 1$ , inspection of the scree plot and interpretability, we retained two dietary patterns in each wave.

Factors were rotated with the varimax procedure, because this method seeks to maximise the variability among the loadings in each factor and hence gives simpler and more easily interpreted factors<sup>(16)</sup>. To evaluate the similarity between factor loadings across all the waves in each dietary pattern, we estimated a coefficient of congruence between each pair of waves. The coefficient of congruence is an index of factor similarity and is estimated as:

$$\sum x_i y_i / \sqrt{\sum y_i^2 x_i^2}$$

where  $x_i$  and  $y_i$  are the factor loadings of food group  $i$  ( $i = 1, \dots, n$ ) on factor  $x$  and  $y$  (i.e.  $x$  = dietary pattern 1 in 1991 and  $y$  = dietary pattern 1 in 1993). A coefficient of congruence ranges from  $-1$  to  $1$ , a value between  $0.85$  and  $0.94$  indicates that the factors are fairly similar, and a value greater than  $0.95$  indicates that the two factors can be considered equal<sup>(17)</sup>.

Because we found similarity between the factor loadings across all the waves, we computed applied scores as many studies with longitudinal dietary patterns had done previously<sup>(18–21)</sup>. The factor loadings from year 2000 were used to calculate the factor scores in all the other waves; we used the loadings of all the food groups included in the dietary pattern analysis, regardless of the loadings being high or low. We chose year 2000 because it was the mid-point of the studied period and the loadings were close to the mean of the loadings of all the other waves and therefore most representative of the entire period. Because this applied score measures the same dietary pattern across time, it allows meaningful temporal comparisons. To estimate the scores with the procedure implemented by Mplus 6.1 (similar to the regression method, but with an iterative technique for categorical variables<sup>(22)</sup>), we carried out a confirmatory factor analysis in each wave specifying the factor loadings and thresholds estimated from year 2000. Because thresholds are  $z$ -scores of the intake variables and they are used in the estimation of the scores, it is important that they are not year specific; therefore, we also fixed them to the estimates of year 2000.

Pearson's correlation coefficients of the factor scores between each pair of waves were computed to assess the tracking (maintenance of a relative position or rank over time) of each dietary pattern. Additionally, we estimated a single tracking coefficient for the entire period. In a longitudinal model, the regression of the value of the score of one wave is performed simultaneously against the scores of all the other waves; the coefficient is standardised so that it can represent a correlation coefficient<sup>(23)</sup>. This coefficient can be interpreted as the average correlation between one wave and all the other waves. We first standardised the dietary pattern score independently in each wave and then used generalised estimating equations to account for the repeated measures (with an unstructured working correlation matrix and robust estimation of standard errors). Because the sample size was the largest in 2000, we performed the regression of the standardised dietary pattern scores of wave 2000 against the standardised dietary pattern scores of all the other waves. We tested for interactions of this tracking coefficient with sex, age, income and region. We present

stratified results and specify when the  $P$  value of the interaction term is less than  $0.10$ :

$$Y_{it} = \beta_0 + \beta_1 Y_{i2000} + \beta_2 t + \beta_3 X_{it} + \beta_4 Y_{i2000} \times X_{it} + \varepsilon_{it}$$

where  $Y_{it}$  is the standardised dietary pattern score for subject  $i$  at wave  $t$  (1991, 1993, 1997, 2004, 2006 and/or 2009),  $Y_{i2000}$  is the standardised dietary pattern score for subject  $i$  at wave 2000,  $\beta_1$  is the tracking coefficient and  $X$  is sex, age, income or region.

We computed factor score means over time, adjusted by age and region, to examine trends in each dietary pattern. Finally, multiple linear regressions with each factor score as the outcome and sociodemographic variables as the predictors were carried out independently in each wave; the clustering at the household level was accounted for in the estimation of the variance. Except for the factor analysis, all the other analyses were conducted in Stata 12.1 (StataCorp).

### Sensitivity analysis

To corroborate that the dietary patterns obtained during each wave were not affected by the sample changes in geographical region and age, we computed inverse probability weights and included them in the factor analysis. We fitted two logistic regressions: one to predict the probability of being in each wave and the other to predict the probability of being in each wave conditional on region and age in 1991. Then, stabilised weights were estimated as the ratio of these two probabilities. The application of these weights is an alternative to standardisation<sup>(24)</sup>. We found that the dietary patterns were very similar to those that we found in our original analysis; the difference between the factor loadings of the two analyses was below  $0.08$  in all the food groups. Similarly, to assess whether changes in energy intake over time affected the dietary patterns, we created inverse probability weights to standardise the patterns by total energy intake in addition to region and age. We found that this standardisation did not affect the composition of the dietary patterns meaningfully.

In addition, because subjects belonging to the same household were more likely to consume the same type of food groups, we repeated the analysis using only a single member per household (selected randomly;  $n$  4837). We found that the dietary patterns did not change meaningfully; all the loadings had a difference below  $0.08$ . Also, the correlation coefficients between the scores remained basically unchanged (all the differences in the coefficients were below  $0.03$ ).

### Results

There was a dramatic increase in the prevalence of overweight from 1991 to 2009 and in the proportion of subjects classified as medium- and high-income earners over the same period, which paralleled changes in urbanisation. The proportion of male smokers declined over time and that of female smokers remained low (Table 1).

The mean total energy intake increased by  $267$  kJ/d ( $64$  kcal/d) from 1991 to 2009 (Table 2). In general, it can be

**Table 1.** General characteristics of the study sample by wave\*  
(Mean values with their standard errors and percentages)

	1991	1993	1997	2000	2004	2006	2009	P for trend
<i>n</i>	5521	5732	6702	7410	6831	6488	5658	
Age (years) in 1991								
Mean	37.3	36.8	35.0	33.5	32.2	31.7	30.7	<0.001
SE	0.14	0.14	0.13	0.12	0.13	0.13	0.14	
Geographical region† (%)								
North	9.6	9.3	9.9	19.7	21.5	22.1	21.9	<0.001
Central	37.2	37.4	36.9	33.4	32.8	32.9	33.1	<0.001
South	53.2	53.3	53.1	46.9	45.7	45.0	45.0	<0.001
Male (%)	50.2	49.8	49.7	49.1	48.5	47.8	48.3	0.002
Overweight (BMI $\geq$ 23 kg/m <sup>2</sup> (36)) (%)	32.9	35.4	43.2	50.6	52.4	54.3	56.7	<0.001
Education (%)								
None	25.5	24.0	23.7	20.0	18.0	23.1	24.4	<0.001
Primary school	23.7	23.7	24.7	24.4	25.8	20.1	21.9	<0.001
Lower middle school or higher education	50.8	52.1	51.3	55.2	56.0	56.6	53.6	<0.001
Income‡ (%)								
Low	87.8	82.2	74.5	63.6	56.0	50.5	33.6	<0.001
Medium	7.5	13.0	19.8	27.9	28.9	30.2	33.2	<0.001
High	0.6	1.2	2.5	5.9	14.2	18.6	33.4	<0.001
Urbanicity§ (%)								
Low	72.8	67.9	57.6	49.3	45.9	40.7	32.7	<0.001
Medium	25.2	30.6	32.1	26.6	22.7	27.0	34.0	0.138
High	2.1	1.5	10.4	24.1	31.5	32.3	33.3	<0.001
Currently smoking (%)								
Females	3.1	3.5	3.8	3.3	3.2	2.9	3.3	0.524
Males	69.8	67.9	64.5	63.6	60.8	58.5	58.6	<0.001
Alcohol intake $\geq$ 3 times/week (%)								
Females	1.7	2.1	2.9	2.7	1.9	1.9	2.4	0.913
Males	25.6	29.9	30.2	32.1	31.9	31.7	31.7	<0.001

\* All variables (except age in 1991 and geographical region) were adjusted by age in 1991 and geographical region.

† North, Heilongjiang and Liaoning provinces; Central, Shandong, Henan and Jiangsu provinces; South, Hubei, Guizhou, Hunan and Guangxi provinces.

‡ Estimated from per capita household income inflated to 2009, categories based on the cut-off values of tertiles in 2009.

§ Estimated from urbanicity index, a multi-component scale that considers population density, economic activity, modern markets, transportation, etc.<sup>(37)</sup>, categories based on the cut-off values of tertiles in 2009.

seen that dietary diversity increased over time, the mean number of food groups consumed increased over time and, for most of the food groups, the percentage of consumers also increased. The only food groups with a decline in the percentage of consumers from 1991 to 2009 were wheat flour, dried legumes, pickled/salted/canned vegetables and low-fat red meat. In contrast, starchy root/tuber products, fresh leafy vegetables, dried vegetables and organ meats had a stable percentage of consumers over time (difference  $\leq$  1 point between 1991 and 2009). In addition, for all the food groups that had an increase in the percentage of consumers, there was also an increase in the amounts consumed (g/capita), except for rice, maize/coarse grains and starchy roots/tubers, for which even if the percentage of consumers increased, the g/capita actually decreased over time (data not shown).

The following food groups had  $\leq$  5% of consumers in all the waves (data not shown): deep-fried rice/legumes; dried fruits; preserved fruits with added sugar; seaweed; processed meats; dairy products; sweetened dairy products; Western-style fast foods; salty snacks; ready-to-eat cereals/porridge; sugar-sweetened beverages; low-energy beverages. Overall, these modern-type foods, over a 3 d period, were not yet widely consumed by 2009.

Using factor analysis for a total of twenty-nine food groups (excluding food groups with  $\leq$  5% of consumers), we identified two dietary patterns, which we called 'traditional southern' and 'modern high-wheat' dietary patterns. These dietary patterns, despite the increase in dietary diversity, remained fairly similar (Table 3). The coefficient of congruence between each pair of waves was generally higher than 0.84, and it was below this value only between 1991 and 2009, 1993 and 2009, and 1997 and 2009 for the traditional southern dietary pattern (0.74, 0.78 and 0.81, respectively) and between 1991 and 2006 and 1991 and 2009 for the modern high-wheat dietary pattern (0.80 and 0.80, respectively) (data not shown). The coefficients of congruence between each wave and wave 2000 are given in Table 3. Across time, the traditional southern dietary pattern was positively associated with the intake of rice, fresh leafy vegetables, low-fat red meat, low-fat and high-fat pork, organ meats, poultry and fish/seafood and inversely associated with that of wheat flour and maize/coarse grains. Interestingly, the factor loadings of rice, wheat flour and maize/coarse grains markedly declined over time, which was not related to the percentage of consumers (as described above, of these food groups, only wheat flour had a decrease in the percentage of consumers). These decreases in the factor loadings indicate that the influence or importance of these food groups in the

**Table 2.** Energy intake, number of food groups consumed and percentage of consumers for different food groups by wave\* (Mean values with their standard errors and percentages)

	1991	1993	1997	2000	2004	2006	2009	P for trend
Energy (kJ/d)								
Mean	6585	6546	6429	6605	6623	6746	6855	< 0.001
SE	28	27	25	24	25	26	28	
Number of food groups consumed								
Mean	7.2	7.5	7.9	8.2	8.7	9.1	9.8	< 0.001
SE	0.04	0.04	0.03	0.03	0.03	0.03	0.04	
Percentage of consumers								
Rice	84	84	87	88	87	87	90	< 0.001
Wheat noodles	23	32	38	42	47	54	55	< 0.001
Wheat flour	51	42	40	41	36	28	25	< 0.001
Wheat buns/breads	3	12	13	15	28	34	37	< 0.001
Cakes, cookies and pastries	3	4	3	3	6	8	8	< 0.001
Deep-fried wheat	6	8	7	8	12	14	15	< 0.001
Maize and coarse grains	21	23	22	22	24	23	26	< 0.001
Starchy roots and tubers	41	39	41	37	44	45	50	< 0.001
Fresh legumes	38	37	41	47	47	54	53	< 0.001
Dried legumes	14	14	13	12	13	12	12	< 0.001
Legume products	41	41	50	49	50	49	54	< 0.001
Nuts and seeds	8	6	8	9	9	8	11	< 0.001
Starchy root and tuber products	9	9	11	11	11	8	9	0.126
Fresh vegetables, non-leafy	75	77	76	83	84	87	90	< 0.001
Fresh vegetables, leafy	86	87	87	86	88	83	85	< 0.001
Pickled, salted or canned vegetables	30	32	26	25	22	22	23	< 0.001
Dried vegetables	9	11	9	9	7	8	10	0.005
Fruits	11	10	12	11	17	23	32	< 0.001
Low-fat red meat	6	9	7	7	4	3	4	< 0.001
High-fat red meat	4	5	8	8	12	13	10	< 0.001
Low-fat pork	9	11	11	10	12	12	13	< 0.001
High-fat pork	55	57	57	64	64	69	73	< 0.001
Organ meats	9	9	9	10	9	8	10	< 0.001
Poultry and game	11	12	16	18	18	18	24	< 0.001
Eggs and egg products	34	32	49	51	53	59	64	< 0.001
Fish and seafood	32	32	34	34	36	37	42	0.382
Soya milk	2	2	2	4	5	10	12	< 0.001
Animal-based milk	1	1	1	3	7	6	5	< 0.001
Instant noodles and frozen dumplings	0	0	2	2	4	12	15	< 0.001

\* All variables were adjusted by age in 1991 and region; data for food groups with  $\leq 5\%$  of consumers in all the waves are not reported.

traditional southern dietary pattern was decreasing over time. The modern high-wheat dietary pattern was positively associated, across time, with the intake of wheat buns/breads, cakes/cookies/pastries, deep-fried wheat, nuts/seeds, starchy root/tuber products, fruits, eggs/egg products, soya milk and animal-based milk. In the earlier years, this pattern was associated with the intake of high-fat red meat, high-fat pork, organ meats and poultry/game. In 1991 and 1993, there were no consumers of instant noodles/frozen dumplings, so this food group was not included in the factor analysis during these years, but when these foods began to be consumed in 1997, they became prominent in the modern high-wheat dietary pattern.

Because overall we found a similarity between dietary pattern structures over time, we computed applied scores, and hence the rest of the results refer to these applied scores. The tracking coefficient of factor scores between wave 2000 and all the other waves was considerably higher for the traditional southern dietary pattern than for the modern high-wheat dietary pattern (0.71 *v.* 0.55, respectively; Table 4). This means that the tracking of the traditional southern dietary pattern was higher or, in other words, that individuals strongly maintained their relative position over time. However, the tracking of the traditional

southern dietary pattern decreased over time; between 1991 and 1993, Pearson's correlation coefficient was 0.79, whereas in more recent years, with a similar 2-year or 3-year interval in between, the correlations were lower (0.71 in 2004–6 and 0.67 in 2006–9). Interestingly, we found that the tracking varied importantly by sociodemographic variables. The tracking was much higher in the Central region, particularly for the traditional southern dietary pattern. The tracking coefficient was also slightly higher among the younger subjects (traditional southern dietary pattern) and those with lower income (both the dietary patterns), whereas it did not vary by sex.

The mean dietary pattern score increased over time for both the dietary patterns; however, the slope was considerably higher for the modern high-wheat dietary pattern (Fig. 1(a)). To understand how much of the increase in the scores of both the dietary patterns was related to the increase in dietary diversity, we adjusted the scores by the number of food groups consumed (Fig. 1(b)). After adjustment, the traditional southern dietary pattern had a flat trend over time and the modern high-wheat dietary pattern had only a slight increase in trend. This means that for the modern high-wheat dietary pattern, to increase the score while holding the number of food groups constant, individuals were either substituting

**Table 3.** Factor loadings for dietary patterns derived with exploratory factor analysis during each wave\*

	Traditional southern dietary pattern							Modern high-wheat dietary pattern						
	1991	1993	1997	2000	2004	2006	2009	1991	1993	1997	2000	2004	2006	2009
Rice	0.87	0.87	0.73	0.56	0.76	0.64	0.50	–	–0.25	–0.25	–0.41	–0.20	–0.28	–0.30
Wheat noodles	0.30	0.27	0.28	–	–	–	–	–	–	–	–	–	–	–0.22
Wheat flour	–0.80	–0.82	–0.86	–0.66	–0.59	–0.67	–0.50	–	–	–	0.26	–	–	–
Wheat buns, breads	–	–	–0.30	–0.27	–0.53	–0.44	–0.49	0.33	0.40	0.39	0.40	0.37	0.48	0.45
Cakes, cookies and pastries	–	–	–	–	–	–	–	0.31	0.48	0.53	0.51	0.57	0.53	0.41
Deep-fried wheat	–	–0.28	–	–	–0.37	–	–	0.46	0.70	0.57	0.71	0.63	0.76	0.74
Maize and coarse grains	–0.84	–0.78	–0.76	–0.65	–0.64	–0.53	–0.35	–	–	–	0.28	–	–	0.28
Starchy roots and tubers	–	–	–	–	–	–	–0.27	–	–	–	–	–	0.20	–
Fresh legumes	–	–	–	–	–	–	–0.20	–	–	–	–	–	0.20	–
Dried legumes	–	–	–	–	–	–	0.23	–	–	–	–	–	–	–
Legume products	–	–	–	–	–	–	–	0.24	–	–	–	–	–	–
Nuts and seeds	–	–	–	–	–	–	0.21	0.42	0.33	0.24	0.26	0.28	0.34	0.30
Starchy roots and tubers products	–0.30	–	–	–	–	–	–	0.30	0.32	0.26	0.26	0.24	0.28	0.26
Fresh vegetables, non-leafy	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Fresh vegetables, leafy	0.25	0.21	0.22	0.22	0.21	0.32	0.46	–	–	0.23	–	–	–	–
Pickled, salted or canned vegetables	0.28	–	0.21	0.25	–	–	–	–	–	–	–	–	–	–
Dried vegetables	–	–	0.21	–	0.31	0.24	0.33	–	–	0.25	–	0.25	–	–
Fruits	–	–	–	–	–	–	–	0.43	0.57	0.46	0.35	0.56	0.55	0.44
Low-fat red meat	0.31	0.50	0.40	0.48	0.33	0.47	0.38	0.40	0.25	0.28	0.20	–	–	–
High-fat red meat	–	–	–	0.24	–	0.28	0.27	0.37	0.37	0.46	0.24	0.36	0.22	0.29
Low-fat pork	0.32	0.40	0.26	0.32	0.49	0.30	0.35	0.25	0.30	–	–	–	–	–
High-fat pork	0.32	0.36	0.44	0.48	0.27	0.34	0.43	0.49	0.34	0.44	0.25	0.32	–	–
Organ meats	0.40	0.49	0.39	0.53	0.32	0.43	0.48	0.45	0.32	0.46	0.24	0.37	–	–
Poultry and game	0.31	0.40	0.32	0.53	0.34	0.49	0.48	0.44	0.44	0.47	0.25	0.40	–	–
Eggs and eggs products	–	–	–	–	–	–	–	0.48	0.45	0.42	0.38	0.39	0.37	0.30
Fish and seafood	0.34	0.39	0.37	0.48	0.45	0.46	0.43	0.44	0.33	0.41	–	0.34	0.26	0.24
Soya milk	–	–	–	–	–0.27	–	–	0.55	0.79	0.58	0.72	0.69	0.76	0.78
Animal-based milk	–	–	–	–	–	–	–	0.64	0.68	0.56	0.58	0.67	0.53	0.49
Instant noodles and frozen dumplings	NA	NA	–	–	–	–	–	NA	NA	0.47	0.40	0.27	0.41	0.44
Percentage of variation in the intake of food groups explained	0.12	0.12	0.11	0.10	0.10	0.10	0.09	0.11	0.13	0.12	0.10	0.11	0.11	0.10
Coefficient of congruence with wave 2000	0.89	0.93	0.93	1.00	0.90	0.97	0.89	0.86	0.95	0.92	1.00	0.95	0.94	0.96

NA; not applicable.

\*Loadings < |0.20| are not shown; for a table with all loadings, see Table S2 (available online).

C. Batis *et al.*



**Table 4.** Tracking of dietary patterns

	Traditional southern dietary pattern						Modern high-wheat dietary pattern					
	1991	1993	1997	2000	2004	2006	1991	1993	1997	2000	2004	2006
Pearson's correlation coefficients of factor scores between the waves*												
1993	0.80						0.63					
1997	0.79	0.81					0.55	0.61				
2000	0.73	0.76	0.76				0.52	0.56	0.57			
2004	0.74	0.77	0.74	0.72			0.52	0.57	0.56	0.56		
2006	0.73	0.76	0.72	0.70	0.72		0.49	0.52	0.52	0.53	0.62	
2009	0.69	0.72	0.71	0.68	0.67	0.67	0.46	0.48	0.48	0.48	0.56	0.59
Tracking coefficients of factor scores between wave 2000 and all the other waves†												
All			0.71						0.55			
South region			0.44						0.38			
Central region			0.77‡						0.47‡			
North region			0.40‡						0.43‡			
Males			0.72						0.54			
Females			0.71						0.56			
Age <35 years			0.73						0.56			
Age 35–49 years			0.70‡						0.55			
Age ≥50 years			0.71‡						0.53			
Low income			0.72						0.55			
Medium income			0.67‡						0.51‡			
High income			0.68‡						0.48‡			

\* Pearson's correlation coefficients were estimated for all the available subjects for each pair of waves; therefore, sample size for each coefficient is different. Sample size ranges from *n* 2888 (1991–9) to *n* 5884 (2000–4).

† Sample size: *n* 7410.

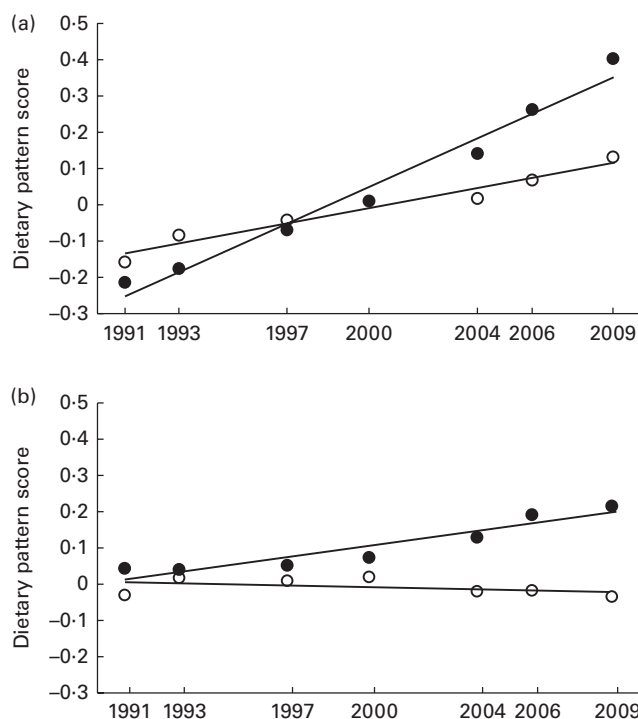
‡ *P* for interaction term < 0.10.

other foods with food groups related to this pattern or consuming higher amounts of these foods.

To examine how the trends in scores were translated into changes in dietary intake, we estimated the percentage of consumers among those in the fourth highest quartile of the scores in each dietary pattern in each wave (Table 5). The trends in consumption observed in the entire sample (i.e. increases in the consumption of most of the food groups) were equally found in subjects with high scores for either pattern. However, the increase in the percentage of consumers of the food groups positively related to the modern high-wheat dietary pattern, among the followers of this pattern, was dramatic. For example, from 1991 to 2009, the intake of wheat buns/breads increased from 7 to 67%; cakes/cookies/pastries from 8 to 24%; deep-fried wheat from 23 to 51%; fruits from 25 to 55%; soya milk from 6 to 44%; and instant noodles/frozen dumplings from 0 to 36%. This large increase in the consumption of key food groups associated with the modern high-wheat dietary pattern can explain the upward trend observed in the scores of this pattern. Although we only present data for food groups relevant for the dietary patterns in Table 5, we also examined food groups excluded from the dietary pattern analysis due to a low proportion of consumers (not shown); we found that none of these food groups had a higher percentage of consumers in one of the dietary patterns, with the exception of deep-fried rice/legumes (5–6% of consumers in the fourth highest quartile in the traditional southern dietary pattern *v.* 1–2% of consumers in the modern high-wheat dietary pattern).

Finally, we ran multiple linear regressions in each wave and examined how sociodemographic factors were related to the traditional southern and modern high-wheat dietary patterns (Table 6). Geographical region was strongly associated with

the traditional southern dietary pattern, with lower scores observed in the Central and North regions than in the South region. Interestingly, the difference in scores between the regions decreased over time, mainly due to a slight increase



**Fig. 1.** Mean factor scores over time. (a) Adjusted by geographical region and age in 1991. (b) Additionally adjusted by the number of food groups consumed. Difference between slopes was statistically significant (*P* < 0.01) in both (a) and (b). ○, Traditional southern; ●, Modern high-wheat.

**Table 5.** Energy intake, number of food groups consumed and percentage of consumers for relevant food groups, among those in the fourth highest quartile of the factor scores in each dietary pattern and year\* (Mean values with their standard errors and percentages)

	Traditional southern dietary pattern			Modern high-wheat dietary pattern		
	1991	2000	2009	1991	2000	2009
Energy (kJ/d)						
Mean	6947	7272	7430	7158	7140	7667
SE	54	46	53	57	48	56
Number of food groups consumed						
Mean	8.4	9.4	10.7	9.1	10.9	12.9
SE	0.07	0.06	0.07	0.07	0.06	0.07
Percentage of consumers						
Rice	100†	100†	100†	67	82	84
Wheat flour	9†	7†	4†	82	61	35
Wheat buns/breads	3	5	15	7†	33†	67†
Cakes, cookies and pastries	3	3	9	8†	12†	24†
Deep-fried wheat	5	4	9	23†	31†	51†
Maize and coarse grains	1†	2†	4†	38	37	42
Nuts and seeds	10	12	16	17†	17†	22†
Starchy root and tuber products	6	11	7	21†	19†	16†
Fresh vegetables, leafy	97†	95†	96†	87	88	88
Fruits	13	19	40	25†	26†	55†
Low-fat red meat	16†	19†	9†	9	12	6
High-fat red meat	5	15	18	9†	16†	18†
Low-fat pork	18†	20†	24†	8	11	15
High-fat pork	87†	91†	93†	75	82	83
Organ meats	27†	29†	25†	13	19	14
Poultry and game	30†	44†	50†	18	31	33
Eggs and egg products	37	55	64	57†	77†	83†
Fish and seafood	60†	65†	68†	42	48	56
Soya milk	1	2	8	6†	15†	44†
Animal-based milk	2	3	7	4†	11†	18†
Instant noodles and frozen dumplings	0	2	14	0†	8†	36†

\* All variables were adjusted by age in 1991 and region.

† Percentage of consumers for relevant food groups in each dietary pattern (loadings  $\geq$  |0.20| in all the waves); percentage of consumers for food groups not relevant for any dietary pattern are not reported.

**Table 6.** Factor score differences by sample characteristics in each wave

	Traditional southern dietary pattern			Modern high-wheat dietary pattern		
	1991	2000	2009	1991	2000	2009
Age in 1991, 10 years' increment	0.01	-0.02*	-0.04*	0.02*	0.01	-0.02
Region						
Central v. South	-1.07*	-0.83*	-0.77*	0.61*	0.72*	0.67*
North v. South	-0.85*	-0.78*	-0.65*	0.60*	0.38*	0.60*
Males v. females	0.03	0.07*	0.05*	-0.07*	-0.06*	-0.08*
Education						
Primary school v. none	0.13*	0.11*	0.06*	0.11*	0.08*	0.06*
Lower middle school or higher education v. none	0.15*	0.09*	0.04	0.23*	0.20*	0.19*
Income						
Medium v. low	0.27*	0.24*	0.11*	0.06*	0.11*	0.05
High v. low	0.17	0.35*	0.22*	0.22*	0.18*	0.18*
Urbanicity						
Medium v. low	0.26*	0.24*	0.20*	0.49*	0.38*	0.38*
High v. low	0.08	0.45*	0.43*	0.78*	0.61*	0.72*
Smoking v. non-smoking	-0.05	-0.03	0.02	0.02	-0.01	0.01
Alcohol intake						
$\geq$ 3 times/week v. $<$ 3 times/week	0.16*	0.08*	0.06*	0.08*	0.10*	0.06*
n	4785	6172	5553	4785	6172	5553
Score mean	-0.16	0.01	0.13	-0.16	0.05	0.47
Score range (minimum-maximum)	-2.3-2.4	-2.4-2.3	-2.5-2.2	-1.5-2.3	-1.6-2.6	-1.6-3.0
Score SE	0.009	0.008	0.009	0.009	0.008	0.009

\*  $P < 0.05$ ; score differences based on a multiple linear regression that included all the variables reported in the table.



over time in the scores in the North and Central regions and a slight decrease in the scores in the South region. Urbanicity was strongly associated with the modern high-wheat dietary pattern, with higher scores being obtained for those living in more urbanised areas. Males had a slightly higher score for the traditional southern dietary pattern and slightly lower score for the modern high-wheat dietary pattern than females. Higher education, income and urbanicity levels, and alcohol intake were associated with higher scores in both the dietary patterns; however, the strength of the association decreased over time. Adjusting by the number of food groups consumed did not affect the estimates, although differences for education, income and urbanicity levels were largely attenuated (data not shown).

## Discussion

In the present longitudinal study that analysed dietary data collected in seven waves over the course of 18 years (from 1991 to 2009), we derived two dietary patterns using factor analysis: a traditional southern dietary pattern characterised by rice, vegetables, meat, poultry and fish and a modern high-wheat dietary pattern characterised by wheat products, nuts, fruits, eggs, milk and instant noodles/frozen dumplings. Despite the rapid increase in dietary diversity, the structure of these two dietary patterns remained stable over time, meaning that the type of foods that the population chose to eat in combination had not changed. However, compared with those of the traditional southern dietary pattern, the scores of the modern high-wheat dietary pattern exhibited lower tracking and had a steeper upward trend over time, indicating that this pattern is more dynamic and is becoming more popular over time. Higher education, income and urbanicity levels were related to higher scores in both the dietary patterns, but the difference in dietary pattern scores between these socio-economic groups became smaller over time, reflecting that foods belonging to this dietary pattern are being consumed more widely.

Although this is the first study to assess longitudinal trends in dietary patterns over time in China, previous studies have analysed dietary patterns at one time point. Although our findings differ from those of studies carried out in specific urban areas, such as Shanghai<sup>(25–28)</sup>, they are most similar to findings of surveys of nationally representative samples, such as the China National Nutrition and Health Survey. Previously identified patterns are the ‘Yellow Earth’ or ‘traditional northern’ pattern high in wheat products, maize or sorghum and tubers; the ‘green water’ or ‘traditional southern’ pattern high in rice, vegetables, seafood, pork and poultry; and the ‘Western’ pattern high in beef, fruits, eggs, poultry, seafood, tofu, milk, cakes, fruit juice, beverages and nuts<sup>(11,12)</sup>. Our ‘traditional southern’ dietary pattern is comparable to the one previously described. However, our ‘modern high-wheat’ dietary pattern could be considered as a combination of the ‘traditional northern’ and ‘Western’ dietary patterns that these previous studies have found. This is reasonable because even in these studies, a considerably high intake of wheat, fried wheat and other cereals was observed in the ‘Western’ dietary pattern. In addition, in the

Jiangsu Province, a pattern similar to our ‘modern high-wheat’ dietary pattern was found, in which the Western-type pattern was also related to the intake of fried wheat and cakes<sup>(29)</sup>. Taken together, it is clear that the regional tradition and availability of wheat in the North region and of rice in the South region are still the main drivers of food selection in the Chinese population. It is interesting that wheat patterns are more likely characterised by the intake of more varied and Western-type foods. Possibly, compared with rice, there is a wider range of foods that can be prepared with wheat, and a wider range of foods that can accompany a wheat-based food, therefore facilitating the incorporation of new foods in the dietary patterns of subjects used to consuming wheat.

Even if the general structure of both the dietary patterns remained stable over time, the tracking or, in other words, the stability of the relative position of each subject’s factor score between time points was higher for the traditional southern dietary pattern than for the modern high-wheat dietary pattern, particularly in the South and Central regions. This finding is comparable to that of a study carried out in Japan<sup>(30)</sup>, where among the prudent, traditional and Westernised patterns, the traditional dietary pattern had the highest tracking among men. The degree of tracking that we found for the traditional southern dietary pattern in the entire sample was remarkably high; the correlation coefficient of scores obtained 18 years apart (1991–2009) was 0.68, which is similar to what other studies carried out in the USA<sup>(31)</sup>, Sweden<sup>(32)</sup> and Japan<sup>(30)</sup> have found but in only 1 year apart (0.56–0.77). In studies carried out 4 or 10 years apart in England<sup>(20)</sup> and Sweden<sup>(33)</sup>, correlation coefficients have been found to be about 0.30–0.52. It is possible that the tracking was lower for the modern high-wheat dietary pattern because of the introduction of new food groups (e.g. instant noodles/frozen dumplings) and/or rapid increase in the intake of the food groups associated with this dietary pattern. It is worth noting that the high tracking of the traditional southern dietary pattern was actually driven by the Central region, whereas the South and North regions had tracking coefficients comparable to those of the modern high-wheat dietary pattern. The Central region had the lowest level of adherence to the traditional southern dietary pattern and the highest level of adherence to the modern high-wheat dietary pattern, and still it was the region with the highest degree of tracking for both the dietary patterns. Further research could explore why the stability of the dietary patterns is higher in this region.

In the present study, we confirmed that the increase in dietary diversity previously observed in 1989–97<sup>(5)</sup> is a phenomenon still going on through 2009. Factor scores would be affected by dietary diversity, because they are similar to a weighted sum of several food groups, so with the exception of food groups with negative loadings, it is expected that the higher the number of food groups consumed, to a lesser or greater extent, the higher the factor scores. Therefore, in the context of increasing dietary diversity, the absolute factor score trend over time represents the degree of adherence to a certain dietary pattern as well as the overall number of food groups consumed. An alternative is to adjust by the number of food groups consumed, because the only way to increase the

factor score while holding the number of food groups consumed constant is to substitute with food groups related to the dietary pattern and/or to increase the amount consumed of these food groups. The present results indicate that in comparison with those of the traditional southern dietary pattern, the factor scores of the modern high-wheat dietary pattern increased significantly over time and, that even after adjusting by the number of food groups consumed, these scores exhibited an upward trend. A clear indication of this trend is the remarkable changes in the diets of those following this pattern (subjects with scores in the highest quartile of their year); although the increase in the intake of food groups such as fruits and soya milk is beneficial, the sharp increase in that of items such as cakes/cookies/pastries, deep-fried wheat or instant noodles is worrisome. Indeed, the adherence to this dietary pattern might be associated with overweight and obesity; we found that for the entire follow-up period, one unit increase in the modern high-wheat dietary pattern was associated with a BMI  $\geq 23$  kg/m<sup>2</sup> (OR 1.27, 95% CI 1.24, 1.30), whereas that in the traditional southern dietary pattern was inversely associated (OR 0.94, 95% CI 0.92, 0.97) (data not shown). In addition, it is of public health importance that with time, food groups of the modern high-wheat dietary pattern seem to be becoming more available to all urban/rural, income and education groups.

The key strength of the present study is the several repeated measures of diet in the same subjects, which allowed us to assess for the first time the stability and tracking of dietary patterns in the Chinese population. Although there were changes in the study sample because of there being subjects without complete dietary data in all the seven waves, we believe that our comparisons between years were not affected by this, not only because all the subjects were included in at least three waves but also because we adjusted all the results by age and geographical region.

A limitation of the present study is that dietary assessment captured dietary intake for only three consecutive days. Although the advantages of this method are important because dietary assessment is very detailed and precise, and the same method can remain constant over time, a disadvantage is that the intake of important food groups that are still only episodically consumed in this country could not be captured. Items such as dairy products, candies/other high-sugar foods, Western-style fast foods, salty snacks or sugar-sweetened beverages had a very low proportion of consumers in a 3 d period and hence we were not able to include these in the factor analysis. Therefore, we do not exclude the possibility that a new, separate Western pattern emerged within the studied period and we were not able to identify it.

In line with this, another important limitation common to all the countries is that the number of foods available in the food supply exceeds by far the number of those available in food composition tables. In the USA, for example, there are over 85 000 uniquely formulated products in the food system, whereas national food composition tables only have about 7600 unique foods<sup>(34)</sup>. The fastest growth in supermarkets occurs in China than in other countries<sup>(35)</sup>, and the present study, because of the inability to keep up with the rapidly changing food supply landscape, was unable to capture all

the changes in dietary intake, particularly of processed packaged foods. In addition, in the present study, dietary assessment focused on the measurement of foods at the ingredient level and not at the dish level, and this is very advantageous because recipes vary considerably between households<sup>(5)</sup> and few mixed dishes are available in the Chinese food composition table; however, we missed this important behavioural aspect of diet, and important changes in dish selection might be going on.

In summary, to the best of our knowledge, this is the first study to assess the stability and tracking over time of dietary patterns in the Chinese population. We found that the way foods are eaten in combination has not changed much between 1991 and 2009. The degree of adherence to a traditional southern dietary pattern has remained unchanged, and there is remarkable inter-individual consistency in the Central region in the intake of foods related to this pattern. On the other hand, the modern high-wheat dietary pattern has become more prominent in the population, and there is more within-subject variation in the adherence level. The trend observed for this modern high-wheat dietary pattern may as well reflect the global influence and economic rise in the last few decades in China. Close attention should be paid to this pattern because it is associated with many energy-dense foods that may affect diet quality. So far, the northern part of the country and subjects with better socio-economic position or in urban areas are more prone to follow this pattern, but it seems that the reach of this pattern is extending towards the general population. A willingness to diversify dietary intake and increase the intake of equally healthful foods such as fruits was evident among the followers of this modern high-wheat dietary pattern; therefore, a great opportunity also lies in the promotion of healthy foods and in the efforts to increase their intake in this population.

### Supplementary material

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S0007114513003917>

### Acknowledgements

The authors thank the National Institute of Nutrition and Food Safety, Chinese Center for Disease Control and Prevention, the Carolina Population Center, the UNC-CH and all the people involved in the CHNS. They also thank Bo Qin for her helpful comments on the manuscript, Ms Frances L. Dancy for her administrative assistance and Mr Daniel Blanchette for his exceptional programming assistance.

The present study received funding from the NIH (R01-HD30880, DK056350, R24 HD050924, R01-HD38700 and R01-HL108427) and the Fogarty International Center. The NIH provided financial support for the collection of CHNS data from 1989 to 2011. C. B. was supported by a scholarship from the Mexican council Consejo Nacional para la Ciencia y Tecnologia (CONACYT). None of the funders, NIH, Fogarty International Center and CONACYT, had any role in the design and analysis of the study or in the writing of this article.

The authors' contributions are as follows: C. B. designed and conducted the analyses and wrote the manuscript; D. S.-A., P. G.-L., M. A. M., L. A. and B. P. contributed to the interpretation of the data and reviewed the manuscript.

None of the authors has any conflicts of interest to declare.

## References

- Wang Y, Mi J, Shan X, *et al.* (2006) Is China facing an obesity epidemic and the consequences? The trends in obesity and chronic disease in China. *Int J Obes (Lond)* **31**, 177–188.
- Yang ZJ, Liu J, Ge JP, *et al.* (2012) Prevalence of cardiovascular disease risk factor in the Chinese population: the 2007–2008 China National Diabetes and Metabolic Disorders Study. *Eur Heart J* **33**, 213–220.
- Wang Z, Zhai F, Du S, *et al.* (2008) Dynamic shifts in Chinese eating behaviors. *Asia Pac J Clin Nutr* **17**, 123–130.
- Wang Z, Zhai F, Zhang B, *et al.* (2012) Trends in Chinese snacking behaviors and patterns and the social-demographic role between 1991 and 2009. *Asia Pac J Clin Nutr* **21**, 253–262.
- Popkin BM, Lu B & Zhai F (2002) Understanding the nutrition transition: measuring rapid dietary changes in transitional countries. *Public Health Nutr* **5**, 947–953.
- Zhai F, Wang H, Du S, *et al.* (2009) Prospective study on nutrition transition in China. *Nutr Rev* **67**, S56–S61.
- Moeller SM, Reedy J, Millen AE, *et al.* (2007) Dietary patterns: challenges and opportunities in dietary patterns research. An Experimental Biology workshop, April 1, 2006. *J Am Diet Assoc* **107**, 1233–1239.
- Hoffmann K, Schulze M, Boeing H, *et al.* (2002) Dietary patterns: report of an international workshop. *Public Health Nutr* **5**, 89–90.
- Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* **13**, 3–9.
- Wang D, He Y, Li Y, *et al.* (2011) Dietary patterns and hypertension among Chinese adults: a nationally representative cross-sectional study. *BMC Public Health* **11**, 925.
- Li Y, He Y, Lai J, *et al.* (2011) Dietary patterns are associated with stroke in Chinese adults. *J Nutr* **141**, 1834–1839.
- Zhang X, Dagevos H, He Y, *et al.* (2008) Consumption and corpulence in China: a consumer segmentation study based on the food perspective. *Food Policy* **33**, 37–47.
- Popkin BM, Du S, Zhai F, *et al.* (2010) Cohort Profile: The China Health and Nutrition Survey – monitoring and understanding socio-economic and health change in China, 1989–2011. *Int J Epidemiol* **39**, 1435–1440.
- Bollen KA (1989) *Structural Equations with Latent Variables*. New York: Wiley.
- Maydeu-Olivares A, García-Forero C, Gallardo-Pujol D, *et al.* (2009) Testing categorized bivariate normality with two-stage polychoric correlation estimates. *Methodology* **5**, 131–136.
- Kim JO & Mueller CW (1978) *Factor Analysis: Statistical Methods and Practical Issues*. Oaks, CA: Sage.
- Lorenzo-Seva U & Ten Berge JM (2006) Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology* **2**, 57–64.
- Mishra GD, McNaughton SA, Bramwell GD, *et al.* (2006) Longitudinal changes in dietary patterns during adult life. *Br J Nutr* **96**, 735–744.
- Crozier SR, Robinson SM, Godfrey KM, *et al.* (2009) Women's dietary patterns change little from before to during pregnancy. *J Nutr* **139**, 1956–1963.
- Northstone K & Emmett PM (2008) A comparison of methods to assess changes in dietary patterns from pregnancy to 4 years post-partum obtained using principal components analysis. *Br J Nutr* **99**, 1099–1106.
- Elstgeest LEM, Mishra GD & Dobson AJ (2012) Transitions in living arrangements are associated with changes in dietary patterns in young women. *J Nutr* **142**, 1561–1567.
- Muthén BO (1998–2004) *Mplus Technical Appendices*. Los Angeles, CA: Muthén & Muthén.
- Twisk JW (2003) Tracking. In *Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide*, pp. 225–240 [JW Twisk, editor]. Cambridge, UK: Cambridge University Press.
- Hernán MA & Robins JM (2006) Estimating causal effects from epidemiological data. *J Epidemiol Community Health* **60**, 578–586.
- Zhang Y, Tan H, Dai X, *et al.* (2012) Dietary patterns are associated with weight gain in newlyweds: findings from a cross-sectional study in Shanghai, China. *Public Health Nutr* **15**, 876–884.
- Villegas R, Yang G, Gao YT, *et al.* (2010) Dietary patterns are associated with lower incidence of type 2 diabetes in middle-aged women: the Shanghai Women's Health Study. *Int J Epidemiol* **39**, 889–899.
- Cui X, Dai Q, Tseng M, *et al.* (2007) Dietary patterns and breast cancer risk in the Shanghai Breast Cancer Study. *Cancer Epidemiol Biomarkers Prev* **16**, 1443–1448.
- Cai H, Zheng W, Xiang YB, *et al.* (2007) Dietary patterns and their correlates among middle-aged and elderly Chinese men: a report from the Shanghai Men's Health Study. *Br J Nutr* **98**, 1006–1013.
- Zuo H, Shi Z, Yuan B, *et al.* (2013) Dietary patterns are associated with insulin resistance in Chinese adults without known diabetes. *Br J Nutr* **109**, 1662–1669.
- Nanri A, Shimazu T, Ishihara J, *et al.* (2012) Reproducibility and validity of dietary patterns assessed by a food frequency questionnaire used in the 5-year follow-up survey of the Japan Public Health Center-Based Prospective Study. *J Epidemiol* **22**, 205–215.
- Hu FB, Rimm E, Smith-Warner SA, *et al.* (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* **69**, 243–249.
- Khani BR, Ye W, Terry P, *et al.* (2004) Reproducibility and validity of major dietary patterns among Swedish women assessed with a food-frequency questionnaire. *J Nutr* **134**, 1541–1545.
- Newby PK, Weismayer C, Akesson A, *et al.* (2006) Long-term stability of food patterns identified by use of factor analysis among Swedish women. *J Nutr* **136**, 626–633.
- Ng SW & Popkin BM (2012) Monitoring foods and nutrients sold and consumed in the United States: dynamics and challenges. *J Acad Nutr Diet* **112**, 41–45.
- Hu D, Reardon T, Rozelle S, *et al.* (2004) The emergence of supermarkets with Chinese characteristics: challenges and opportunities for China's agricultural development. *Dev Policy Rev* **22**, 557–586.
- WHO Expert Consultation (2004) Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* **363**, 157–163.
- Jones-Smith JC & Popkin BM (2010) Understanding community context and adult health changes in China: development of an urbanicity scale. *Soc Sci Med* **71**, 1436–1446.