Dietary patterns generated by the Treelet Transform and risk of stroke: a Danish cohort study

Signe B Frederiksen[1],†, Heidi H Themsen[1],†, Kim Overvad[1,2] and Christina C Dahm[1],*

[1]Department of Public Health, Aarhus University, Bartholins Allé 2, DK-8000 Aarhus C, Denmark: [2]Department of Cardiology, Aalborg University Hospital, Hobrovej 18, DK-9000 Aalborg, Denmark

Abstract

Objective: To relate empirically derived dietary patterns identified using the Treelet Transform (TT) to risk of stroke.

Design: A prospective cohort study using the Danish Diet, Cancer and Health cohort. Dietary information was obtained in 1993–1997 using a validated semi-quantitative FFQ. Incident stroke diagnoses, obtained from the Danish National Patient Register, were verified by record review. Dietary patterns were generated using TT, and participants were categorised into quintiles based on their adherence to each pattern. Sex-specific Cox proportional hazard models estimated associations between dietary patterns and stroke.

Setting: Denmark.

Participants: 55 061 men and women aged 50–64 years at the time of enrolment.

Results: Three dietary patterns explaining 15·4 % of the total variance were identified: a *Prudent* pattern, a *Western* pattern and a *Wine & Snacks* pattern. During a follow-up time of 10 years, 1513 cases occurred. Comparing the highest to lowest quintiles of intake, adherence to a Prudent pattern was inversely associated with stroke (HR$_{men}$ 0·74, 95 % CI 0·60, 0·91; HR$_{women}$ 0·82, 95 % CI 0·62, 1·08), while adherence to a Western pattern was associated with greater risk (HR$_{men}$ 1·61, 95 % CI 1·23, 2·10; HR$_{women}$ 2·01, 95 % CI 1·48, 2·72). No association was found for a Wine & Snacks pattern for women, but a weak inverse association was found for men (HR 0·81, 95 % CI 0·67, 0·99).

Conclusions: The results of this study are broadly in line with current recommendations for a healthy diet to prevent stroke.

CVD, including stroke, is the leading cause of mortality among adults. According to the WHO, 6·7 million deaths were due to stroke in 2015, and the European Heart Network estimated that stroke causes close to one million deaths in Europe each year[1–3]. The incidence of stroke is increasing globally, and as more people survive a stroke to live with life-long consequences[4], the burden of disease is expected to increase in the years to come[4,5]. Prevention is, therefore, imperative to reduce the risk of stroke and thereby the burden that stroke imposes on healthcare systems and on the individual.

One of the prominent modifiable risk factors for stroke is diet, which is also important in developing other conditions known as risk factors for stroke, such as hypertension, hypercholesterolaemia and diabetes mellitus[1,6–8]. So far,

no clear association has been found between any single food or nutrient and risk of stroke[9]. A review by Sherzai showed that the associations between single foods and risk of stroke differed depending on the study population. This could be due to the complexity of dietary preferences and complicated interactions between foods that are not accounted for by examining single foods, including cultural differences in overall dietary habits and food preparation[9–15]. To address this, dietary patterns can be derived using exploratory methods. These methods are data-driven and are used to identify patterns based on information on real dietary habits[14,15]. Frequently used exploratory methods include cluster and factor analysis. Cluster analyses derive groups of people with similar dietary habits and categorise individuals into mutually exclusive clusters[10,14,15]. Principal component analysis (PCA) is a factor analysis often used to create meaningful

---

†These authors contributed equally to this work.

Corresponding author: Email ccd@ph.au.dk

© The Authors 2020

factors, or patterns, based on correlations between individual food groups. PCA uses all available food groups to describe the factors, resulting in patterns that are complex to interpret[10]. The Treelet Transform (TT) is a recently developed exploratory method that combines the strengths of PCA and cluster analysis. Similar to PCA, TT uses the correlation matrix to derive factors based on most correlated foods; however, the factors produced by TT contain a smaller number of food groups and are thereby simpler to interpret[13,16]. To date, very little is known about how TT-derived patterns relate to the risk of disease.

In a meta-analysis of studies that used cluster analysis, factor analysis or PCA to generate dietary patterns and to assess their associations with CVD[17], 'Prudent/Healthy' patterns were associated with an inverse risk of coronary heart disease, but not statistically significantly with the risk of stroke. The 'Western/unhealthy' patterns were not consistently associated across global regions, such as North America/Europe and Asia. Hence, there could be different prevalent dietary patterns associated with the risk of CVD, and dietary patterns that share similarities across global regions are not consistently associated with disease development. This may be due to differing methodologies between studies, or true cultural differences in food habits. A dimension reduction method to derive dietary patterns that requires fewer investigator-driven decisions may aid in disentangling this.

Thus, the aims of this study were to use TT to identify dietary patterns that explained variations in dietary habits in this Danish population, and to examine the associations between these dietary patterns and the risk of stroke. Based on previous studies, we expected to identify two to three dietary patterns[17–20], and that high adherence to any pattern, including foods recommended by the American Stroke Association, would be associated with a lower incidence of stroke[8].

Materials and methods

Study design and population

We used the Danish Diet, Cancer and Health cohort study. Men and women aged 50–64 years born in Denmark, living in the greater areas of Aarhus and Copenhagen and without any previous cancer diagnosis registered in the Danish Cancer Registry, were invited to participate in the study between December 1993 and May 1997. Participants were excluded if information on diet, covariates or enrolment date was missing, if the participant had previously had cancer unregistered at the time of invitation, or a stroke prior to enrolment.

At enrolment into Diet, Cancer and Health, participants filled in an FFQ and visited one of the study clinics. In the study clinic, a questionnaire on lifestyle factors and health was filled in and participants underwent a physical examination where anthropometrical measurements, such as

weight, height and blood pressure, were taken and biological samples were collected[21].

Diet assessment

Prior to the study clinic visit, participants completed a self-administered, 192-item semi-quantitative FFQ, specifically developed for the Diet, Cancer and Health cohort. The FFQ contained questions relating to different beverages, foods and dishes, and participants reported how often within the past year they had consumed each food item within twelve categories, ranging from 'never' to 'eight times a day'[21]. The average daily intake of specific foods and nutrients was calculated for each participant using the software programme FoodCalc, using specifically developed standardised recipes and portion sizes[21,22]. The development and validation of the FFQ has been described elsewhere[22,23].

Each food item was grouped into larger food groups based on the general guidelines from the European Prospective Investigation into Cancer and Nutrition, of which the Diet, Cancer and Health study is a part[24]. Further groupings were made based on cultural considerations so that similar foods were aggregated into a total of thirty-one food groups.

Outcome: stroke

Possible stroke cases, ischaemic or haemorrhagic, which had occurred before 30 November 2009 were identified by linkage to the Danish National Patient Register based on the International Classification of Disease codes (ICD-8: 430, 431, 433, 434, 436.01 or 436.90; ICD-10: I60, I61, I63 or I64). Diagnoses were then verified by a physician with neurological experience, who reviewed all medical and discharge records[25].

Covariates

Potential confounders were selected *a priori* as established risk factors for stroke.

Information on sex, education, smoking, alcohol intake, physical activity and hypertension, hypercholesterolaemia and diabetes was obtained at baseline from the self-administered lifestyle questionnaire. In the questionnaire, participants stated their educational level after elementary school (none, short, medium and long), current smoking habits (never, former or current), how often they consumed alcohol, including beer, wine, spirits and fortified wine (never, <one occasion per month, one–three occasions per month, one occasion per week, two–four occasions per week, five–six occasions per week, daily) and their habitual leisure-time physical activities (hours per week engaged in walking, biking, housework, home maintenance, gardening and sports). Participants stated if they had hypertension, hypercholesterolaemia and diabetes mellitus, and if they were taking medications for these conditions, and the hypertension, hypercholesterolaemia and

diabetes mellitus variables were constructed combining the responses to these questions.

Height and weight were measured by a lab technician during physical examination. Height was measured, to the nearest 0·5 cm, with the participant standing without shoes. Weight was measured, to the nearest 0·1 kg, using a digital scale, with the participant wearing light clothing or underwear. BMI was calculated as weight/height^2 (kg/m^2).

Statistical analysis

Dietary intake of each food group was adjusted for total energy intake by the residual method to prevent participants with a high energy intake from dominating the dietary patterns[26].

The Treelet Transform

Dietary patterns were exploratively generated using TT as described by Gorst-Rasmussen[13,16]. TT is a dimension reduction method, based on the statistical principles of cluster analysis and PCA. Few investigator-led decisions are required to derive meaningful factors in a given dataset. The TT algorithm introduces sparsity in factor loadings, making many factor loadings exactly zero. This results in fewer variables (food groups) being included in each factor, compared to a PCA, and thereby simplifies the investigators' interpretation of which variables are important contributors to the pattern. The TT algorithm locates the food groups with highest correlations and performs a local PCA. The two food groups are merged at the bottom of the cluster tree and replaced by a sum factor. Then a local PCA is performed on the sum factor and the next most correlated food group. The sum factor and the next most correlated food group are then replaced by a sum factor at the next level of the tree. This is repeated for all food groups, so that a local PCA is made for all the remaining correlations between food groups until all food groups are represented in the cluster tree. A sum factor is estimated at each level of the cluster tree representing a connection between included food groups. Selecting a cut level for the TT cluster tree amounts to choosing the level of detail desired in the dimension reduction of data. More variation can be explained at the cost of factor sparsity with a higher cut level. Here, a ten-fold cross-validation was used to define the optimal cut level for the cluster tree, that is, the point at which increasing the cut level does not substantially increase the variation of retained patterns.

The number of meaningful dietary patterns was decided at this cut level based on the increase in explained variation that another pattern could contribute with, compared to a single food group. To only identify meaningful patterns, the eigenvalue cut-off for the patterns was set to >1·25[27].

Within the generated patterns, a factor loading for each underlying food group was estimated based on the sum factor for the pattern itself.

The participants' adherence to each dietary pattern (TT scores) was estimated as the average of their energy-adjusted intake of food groups contributing to the factor weighted by food group factor loadings. Positive scores indicated a larger-than-average intake of contributing food groups and vice versa. Participants were categorised into quintiles (Q) based on their adherence to each pattern. Independence between patterns was verified using Spearman's correlation.

Cox proportional hazard model

Cox proportional hazard models were used to estimate hazard ratios (HR) with 95 % CI for the association between adherence to dietary patterns and stroke. Age was used as the underlying timescale. Time at risk was defined as time from enrolment until an event – stroke, death by other cause, emigration or 30 November 2009, whichever came first. All analyses were stratified by quintiles of date of enrolment, since conditional independence between participants' risk of stroke and date of enrolment was not confirmed. Model 1 was adjusted for age and total energy intake, and model 2 was further adjusted for potential confounders: educational level (none, short, middle long, and long), smoking (never, former and current), physical activity (>/<30 min/d), BMI (underweight – <18·5 kg/m^2; normal weight – 18·5–24·9 kg/m^2; overweight – 25–29·9 kg/m^2) and obesity (≥30 kg/m^2), together with self-reported hypertension, hypercholesterolaemia and diabetes mellitus (yes, no or unknown). Model 2 was additionally adjusted for alcohol intake (<one occasion per month, one–four occasions per month, two–four occasions per week, >four occasions per week), for dietary patterns not including alcohol.

HR were estimated for total stroke in men and women separately, to evaluate whether associations between dietary patterns and risk of stroke differed by sex. As this was an *a priori* decision, a test for effect modification on multiplicative scale was not performed. The proportional hazards assumption was assessed using log-log plots and global tests of proportional rates for dietary patterns during follow-up. A sensitivity analysis was performed after adjusting for smoking in five categories (never, former, current <15 g tobacco per day, 15–25 g/d, >25 g/d) instead of three, alcohol consumption as g/d, and physical activity expressed as hours per week instead of categorical variables.

A complete case analysis was conducted. All statistical analyses were conducted in Stata/IC 15.1. Two-sided *P* values ≤0·05 were considered statistically significant.

Results

A total of 160 725 Danish men and women were invited to participate in the Diet, Cancer and Health study, of which 57 053 (35·5 %) chose to participate. 55 061 were included
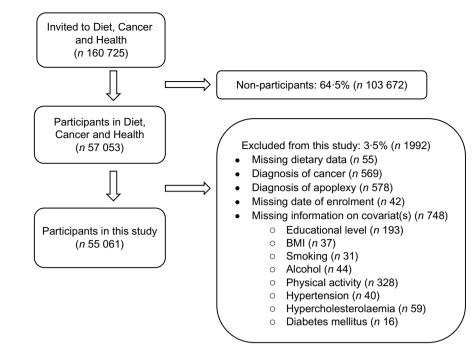
Fig. 1 Flowchart of Diet, Cancer and Health cohort and exclusion of participants from this study

in the present study. In total, 1992 participants were excluded (Fig. 1). During a median follow-up time of 13·5 years, 1357 verified incident stroke cases occurred among men and 901 cases among women. Compared to the cohort, stroke cases were more often male, had no education after elementary school, were overweight or obese, were smokers, drank alcohol daily, were less physically active and reported hypertension, hypercholesterolaemia and diabetes mellitus (Table 1).

When conducting the TT analysis, the cross-validation showed an optimal cut level at fourteen for two to four patterns, at which level three patterns had eigenvalues >1·25. The correlation matrix for the thirty-one food groups with these three dietary patterns is illustrated in the cluster tree in Fig. 2. The cumulative variance in diet described by the three patterns was 15·4 %. The eigenvalue and variance of each pattern is shown in Table 2. The first pattern derived from TT, named *Prudent,* was characterised by a high intake of, for example, fish, fruit and vegetables. The second pattern was named *Western* and had a high intake of animal products, and the third pattern was named *Wine & Snacks* with a high intake of alcohol and salty snacks. Food groups represented in the dietary patterns as well as factor loadings for each food group are listed in Table 2.

The distribution of potential confounders across quintiles of the three dietary patterns is shown in Table 3. Participants with high adherence to the Prudent pattern were more often women, had longer education, were less likely to be smokers, drank alcohol more frequently, were more physically active and had hypertension,

hypercholesterolaemia and diabetes mellitus compared to participants with low adherence. Within the Western pattern, participants with high adherence were more likely to be male, had shorter education, less healthy habits and hypercholesterolaemia and diabetes mellitus compared to participants with low adherence. Participants with high adherence to Wine & Snacks were more often male, had longer education, few were obese, more were previous smokers, more drank alcohol frequently and few had diabetes mellitus compared to participants with low adherence.

The associations between adherence to the Prudent, Western and Wine & Snacks patterns and stroke are shown in Table 4 (men) and Table 5 (women). The proportional hazards assumption could not be accepted, as the association with baseline dietary patterns changed over the complete follow-up time. To obtain proportional hazards, the follow-up time was, therefore, truncated at 10 years, during which 913 cases occurred among men and 600 cases among women.

In men, adherence to the Prudent pattern was associated with a lower risk of stroke after adjusting for confounders (Q5 *v.* Q1 HR 0·74, 95 % CI 0·60, 0·91). Point estimates in women were similar to those for men, but not statistically significant. High adherence to the Western pattern was associated with a greater risk of stroke in both men and women after adjusting for confounders (men: Q5 *v.* Q1 – HR 1·61, 95 % CI 1·23, 2·10; women: Q5 *v.* Q1 – HR 2·01, 95 % CI 1·48, 2·72).

Adherence to the Wine & Snacks pattern was weakly inversely associated with stroke in men (Q5 *v.* Q1 – HR 0·81, 95 % CI 0·67, 0·99), but not in women.

Table 1 Baseline characteristics of study cohort and participants who developed stroke during follow-up

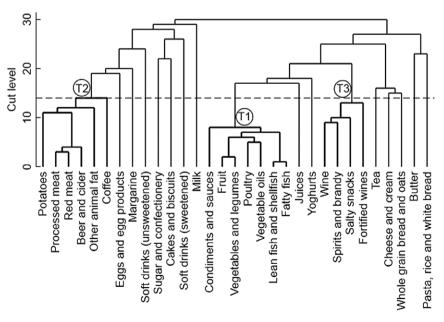|	All *n*, 55 061	p10–p90	Cases *n*, 2258	p10–p90
Age, median	56	51–63	59	52–64
Sex, women (%)	52·4		39·9	
Educational level after elementary school (%)				
None	14·8		17·7	
Short	23·0		21·3	
Middle long	40·1		38·4	
Long	22·2		22·7	
BMI (%)				
<18·5	0·8		0·8	
18·5–24·9	43·1		37·1	
25–29·9	41·7		43·4	
≥30	14·5		18·7	
Smoking (%)				
Never	35·3		23·9	
Former	28·8		26·1	
Current	35·9		50·0	
Alcohol (%)				
<One occasion per month	8·1		10·7	
One–four occasions per month	25·0		21·7	
Two–four occasions per week	33·3		28·5	
>Four occasions per week	33·6		39·1	
Physical activity (%)				
≥30 min/d	39·5		34·2	
Hypertension (%)				
Yes	16·0		27·7	
No	71·0		58·5	
Unknown	13·0		13·8	
Hypercholesterolaemia (%)				
Yes	7·4		10·9	
No	50·4		48·5	
Unknown	42·3		40·5	
Diabetes mellitus (%)				
Yes	2·0		4·3	
No	93·4		90·0	
Unknown	4·5		5·8	

Fig. 2 Cluster tree of food groups produced by Treelet Transform with three factors at cut level 14: T1 (Prudent), T2 (Western) and T3 (Wine & Snacks)

Table 2 Factor loadings for food groups within the derived dietary patterns from Treelet Transform analyses

Food group	Prudent	Western	Wine & Snacks
Potatoes		0·3020	
Processed meat		0·5254	
Red meat		0·5254	
Beer and cider		0·4619	
Other animal fat		0·2901	
Coffee		0·2433	
Eggs and egg products			
Margarine			
Unsweetened or artificially sweetened carbonated/soft/isotonic drinks, diluted syrups			
Sugar and confectionery			
Cakes and biscuits			
Sweetened carbonated/soft/isotonic drinks, diluted syrups			
Milk			
Condiments and sauces	0·2738		
Fruit	0·3956		
Vegetables and legumes	0·3956		
Poultry	0·3550		
Vegetable oils	0·3550		
Lean fish and shellfish	0·4243		
Fatty fish	0·4243		
Fruit and vegetable juices			
Yoghurts			
Wine			0·5462
Spirits, brandy			0·5462
Salty snacks			0·4763
Fortified wines			0·4202
Tea			
Cheese and cream			
Wholegrain bread and oats			
Butter			
Pasta, rice and white bread			
Explained variation (%)	5·9	5·2	4·3
Eigenvalue	1·8	1·6	1·3

Sensitivity analyses with more fine-grained control for confounding showed no changes in associations (data not shown).

Discussion

In this large cohort study, three dietary patterns were derived based on information about 55 061 participants' dietary intake. Adherence to a Prudent dietary pattern was associated with a lower risk of stroke among men, but not significantly so among women. High adherence to a Western pattern was associated with a greater risk for both men and women. Adherence to a Wine & Snacks pattern was associated with a slightly lower risk of stroke among men, but no association was found in women.

Our study has strengths and weaknesses. Risk of bias related to the selection of participants was unlikely in our study, thanks to the very limited loss to follow-up (0·6 % due to emigration).

Information about dietary intake was obtained through a semi-quantitative FFQ, which is a common choice of dietary instrument in large cohort studies[26]. The FFQ used in this study has been validated in previous studies[22,23,28,29]. Some misclassification related to exposure is inevitable as dietary information was self-reported; however, recall-bias was likely to have been non-differential with respect to the outcome, resulting in an underestimation of associations. Based on previous studies, measurement error due to systematic underreporting as a function of total energy intake, underreporting of alcohol, and sex-dependent overreporting of healthy foods was expected[26,30]. To reduce the influence of measurement errors related to systematic underreporting of all food groups, participants' intake of each food group was adjusted for total energy intake, making their intake relative rather than absolute. Nevertheless, over- or underreporting of specific food groups could have affected the generation of our dietary patterns, both in terms of which food groups were represented in the patterns and in the calculation of factor loadings for each food group. Thus, non-differential misclassification of adherence to the dietary patterns among participants with a high alcohol intake and sex-specific overreporting of healthy foods could have occurred. Overall, the dietary patterns may have included other foods if true intake had been available, but neither of these potential measurement errors or misclassifications are expected to be related to participants' risk of stroke, as they occurred prior to incident disease. Therefore, these errors are most likely non-differential and would tend to underestimate the true associations.

Table 3 Baseline characteristics of participants according to selected quintiles of derived dietary patterns

n 55 061	Prudent						Western						Wine & Snacks					
	Q1	p10–p90	Q3	p10–p90	Q5	p10–p90	Q1	p10–p90	Q3	p10–p90	Q5	p10–p90	Q1	p10–p90	Q3	p10–p90	Q5	p10–p90
Age, median	56	51–63	56	51–63	56	51–63	56	51–63	56	51–63	56	51–63	57	51–64	56	51–63	56	51–63
Sex, women (%)	35·1		56·4		59·4		74·8		60·1		14·6		57·8		53·9		42·5	
Educational level after elementary school (%)																		
None	21·0		13·8		10·4		10·8		15·5		17·2		24·3		13·5		8·0	
Short	20·7		24·4		23·8		26·0		24·7		17·3		24·8		23·1		20·3	
Middle long	39·2		40·3		39·5		39·6		39·1		41·3		36·0		42·6		39·8	
Long	19·1		21·6		26·3		23·6		20·7		24·1		14·9		20·8		31·9	
Smoking (%)																		
Never	28·6		36·9		37·7		47·4		37·4		18·6		36·6		38·6		28·2	
Former	24·7		28·9		33·5		30·1		29·4		26·7		23·6		29·2		32·7	
Current	46·8		34·2		28·8		22·5		33·2		54·7		39·8		32·2		39·1	
Alcohol (%)																		
<One occasion per month	11·7		7·0		7·6		12·1		8·3		4·8		30·0		1·9		0·9	
One–four occasions per month	27·3		24·9		23·1		32·0		26·6		13·8		46·0		22·4		3·2	
Two–four occasions per week	29·3		34·3		34·5		33·5		35·7		26·3		14·5		50·2		21·2	
>Four occasions per week	31·8		33·8		34·8		22·4		29·5		55·2		9·5		25·6		74·8	
Physical activity (%)																		
≥30 min/d	32·9		38·8		48·2		50·7		38·0		32·1		38·8		38·9		39·0	
BMI (%)																		
<18·5	1·0		0·9		0·7		1·3		0·7		0·5		1·3		0·7		0·6	
18·5–24·9	42·1		44·0		41·7		54·1		42·7		33·0		41·6		43·1		44·4	
25–29·9	42·3		41·0		42·4		34·3		42·2		48·1		39·0		42·6		43·5	
≥30	14·6		14·1		15·3		10·3		14·4		18·4		18·2		13·6		11·6	
Hypertension (%)																		
Yes	13·0		15·7		18·3		14·6		16·8		15·2		16·8		15·5		16·4	
No	70·2		71·8		71·1		75·3		71·2		67·0		70·2		71·4		70·9	
Unknown	16·8		12·5		10·6		10·1		12·0		17·9		13·0		13·1		12·7	
Hypercholesterolaemia (%)																		
Yes	5·4		7·0		10·5		6·3		7·7		8·1		7·7		7·0		8·1	
No	49·2		50·6		51·5		53·6		50·1		47·9		50·2		49·5		51·6	
Unknown	45·4		42·3		38·0		40·1		42·2		44·0		42·1		43·5		40·3	
Diabetes mellitus (%)																		
Yes	1·4		1·9		3·1		1·3		2·1		2·7		3·5		1·4		1·7	
No	92·5		93·8		93·1		94·9		93·6		91·0		91·0		93·9		94·1	
Unknown	6·1		4·4		3·8		3·8		4·3		6·4		5·4		4·7		4·2	

Table 4 HR (95 % CI) for associations between adherence to dietary patterns and risk of total stroke in men, truncated at 10 years of follow-up

n 26 219	Q1	Q2 HR	Q2 95 % CI	Q3 HR	Q3 95 % CI	Q4 HR	Q4 95 % CI	Q5 HR	Q5 95 % CI
Prudent									
Model 1*	1	0.88	0.73, 1.05	0.67	0.55, 0.83	0.83	0.68, 1.01	0.73	0.60, 0.89
Model 2†	1	0.91	0.76, 1.09	0.69	0.56, 0.85	0.87	0.71, 1.06	0.74	0.60, 0.91
Western									
Model 1*	1	1.19	0.87, 1.63	1.36	1.02, 1.82	1.29	0.97, 1.70	1.99	1.53, 2.57
Model 2†	1	1.14	0.83, 1.56	1.25	0.93, 1.67	1.11	0.84, 1.47	1.61	1.23, 2.10
Wine & Snacks									
Model 1*	1	0.78	0.63, 0.96	0.74	0.60, 0.90	0.66	0.54, 081	0.75	0.62, 0.91
Model 2†	1	0.84	0.68, 1.04	0.80	0.65, 0.98	0.72	0.59, 0.88	0.81	0.67, 0.99

HR, hazard ratios.
*Model 1 is adjusted for age and total energy intake.
†Model 2 is further adjusted for education, smoking, alcohol (Prudent), physical activity, BMI, hypertension, hypercholesterolaemia and diabetes mellitus.

Misclassification of outcomes in this study could only be non-differential with respect to the exposure, as the classification and verification of cases were obtained independently from knowledge about dietary intake. Verification through medical journals furthermore excluded false-positive cases[31]. Thus, bias related to misclassification of outcome was minimised.

Adjusting for potential confounders, identified *a priori*, showed that these confounded the associations between dietary patterns and risk of stroke, but a further fine-grained adjustment did not alter the results. Nevertheless, residual confounding in this observational study cannot be ruled out.

The aim of this study was to describe dietary patterns explaining variations in actual dietary habits. For this purpose, an exploratory method such as TT is well suited. A strength of exploratory methods is their ability to account for the correlated combinations of foods, which are part of the cumulative effect of diet. Dietary patterns generated using PCA included all food groups, making the patterns complex and difficult to interpret. Using TT, we generated patterns that were easier to interpret due to their simplicity, compared to patterns generated with PCA. The dietary patterns generated in this study explained 15·4 % of variance in diets. This may seem low compared to the dietary patterns generated with PCA, which in other studies explained between 16 and 27 % of variance with two to three patterns; yet, Gorst-Rasmussen *et al*. found that TT described almost as much variation as PCA for the same number of patterns[13]. However, as the aim of this study was to explain the variance in diet, and not to generate patterns most correlated with disease, an outcome-independent method, rather than an outcome-dependent method, was chosen. Therefore, these patterns cannot be interpreted as patterns most associated with the risk of stroke[10,14].

Using exploratory methods necessitates subjective choices made by the investigator[10,13–15]. In this study, standardisation and cultural adaption of food groups, adjustment for total energy, the expected number of patterns generated, at what level the cluster tree was to be cut, and the naming of patterns were given a careful consideration. No separate patterns were derived for subgroups, as the study population was considered homogenous in relation to diets, and other studies have not found sex differences in dietary patterns[11]. The reproducibility and validity of exploratory-derived dietary patterns are often discussed, as knowledge on this subject is still insufficient[10,12,15]. A disadvantage of these methods is that identical patterns can never be derived across different study populations, making the comparison of results difficult[14,15]. However, the patterns determined here had high loadings of food groups similar to those found in other studies[9,12,17]. This indicates some reproducibility of exploratory-generated dietary patterns across study populations.

The proportional hazards assumption was not met in this study, perhaps because other age-related causes of

Table 5 HR (95 % CI) for associations between adherence to dietary patterns and risk of total stroke in women, truncated at 10 years of follow-up

n 28 842	Q1	Q2 HR	Q2 95 % CI	Q3 HR	Q3 95 % CI	Q4 HR	Q4 95 % CI	Q5 HR	Q5 95 % CI
Prudent									
Model 1*	1	0·91	0·70, 1·19	0·92	0·71, 1·19	0·75	0·57, 0·97	0·74	0·57, 0·97
Model 2†	1	0·95	0·73, 1·24	1·00	0·77, 1·30	0·83	0·63, 1·09	0·82	0·62, 1·08
Western									
Model 1*	1	1·23	0·97, 1·57	1·51	1·19, 1·91	1·50	1·16, 1·94	3·00	2·24, 4·01
Model 2†	1	1·12	0·88, 1·42	1·27	1·00, 1·61	1·12	0·86, 1·47	2·01	1·48, 2·72
Wine & Snacks									
Model 1*	1	0·84	0·66, 1·06	0·94	0·74, 1·19	0·87	0·68, 1·12	0·92	0·71, 1·19
Model 2†	1	0·89	0·70, 1·13	1·04	0·82, 1·33	0·93	0·72, 1·20	0·92	0·71, 1·19

HR, hazard ratios.
*Model 1 is adjusted for age and total energy intake.
†Model 2 is further adjusted for education, smoking, alcohol (Prudent), physical activity, BMI, hypertension, hypercholesterolaemia and diabetes mellitus.

stroke increased in importance over time[7], leading to changes in baseline hazards, or because participants' dietary habits changed during follow-up. Thus, participants' adherence to dietary patterns at enrolment probably does not reflect their true exposure across the entire follow-up period, and an HR for the entire follow-up time would have been deceptive. However, our literature review did not observe a cut in follow-up time in any other studies of long follow-up using baseline dietary measurements. This phenomenon may, therefore, depend on the specific dietary exposure under study. Truncating follow-up at 10 years ensured that model assumptions were met but decreased the number of incident cases for analyses to 1513 over 10 years.

In this study, we found inverse associations between adherence to a Prudent pattern and risk of stroke, which has also been found by others using exploratory-derived patterns. For example, Fung *et al.* and Stricker *et al.* found similar patterns consisting of high intakes of fruit, vegetables, whole grains, fish and poultry that were associated with a lower risk of stroke (HR 0·78, 95 % CI 0·61, 1·01; and 0·69, 95 % CI 0·53, 0·88, respectively)[19,35].

Comparing the dietary patterns derived by other studies, the recurrent food groups in a Western pattern are red and processed meat[17,35,36], but apart from meats, the important food groups within Western-type patterns differed between studies. Unsurprisingly, the associations between Western/unhealthy patterns and stroke also vary. In a meta-analysis by Rodriguez-Monteforte, no associations between adherence to a Western/unhealthy pattern and stroke were found[17]. This could be explained by the differing food groups within these patterns, which might have influenced the overall associations with stroke.

Contrary to the results of this study, a high adherence to the Wine & Snacks pattern might be expected to be associated with a greater risk of stroke due to the high intake of alcohol in the upper quintiles of the pattern. Our results could be due to the non-uniform reference group (Q1), as this quintile included both participants who never

consume alcohol – possibly for health-related reasons, and thus might have a high underlying risk of stroke[37] – and those with frequent alcohol intake. A quarter of participants in Q1 consumed alcohol more than two times a week (Table 3). Thus, adherence to the Wine & Snacks pattern is not a representation of total alcohol intake, but rather of specific alcoholic beverages, and the associations between the patterns and risk of stroke can, therefore, not be interpreted as representing a greater consumption of alcohol as a whole.

Our results are likely to be transferrable to other populations of the same age and within similar social and cultural settings as the study participants. While the participants in the Diet, Cancer and Health cohort were better off in relation to socioeconomic factors compared to non-participants, there were no major differences between the participants and Danish population in relation to dietary habits at entry into the study[21]. In addition, the dietary patterns derived by TT were somewhat reproducible[13], although it is unlikely that studies in other populations will derive entirely identical patterns due to differences in food availability and cultural preferences.

Conclusion

We used TT to generate dietary patterns, of which three explained 15·4 % of variation in energy-adjusted dietary intake. Adherence to a Prudent pattern was inversely associated with stroke, whereas adherence to a Western pattern was associated with a higher risk of stroke. We found no association between a Wine & Snacks pattern and stroke in women, but a weak inverse association in men.

The findings of this study support the general recommendation for a healthy diet to prevent stroke as described by the ASA. In future studies, it would be relevant to investigate the role of older age and changing dietary habits.

Also, it would be relevant to compare dietary patterns generated with an outcome-dependent method with

current dietary recommendation to clarify the dietary components that explain the most variation in health outcomes.

References

1. World Health Organization (2017) Cardiovascular Diseases (CVDs). http://www.who.int/mediacentre/factsheets/fs317/en/ (accessed October 2018) .
2. World Health Organisation (2003) *The World Health Report 2003. Shaping the Future.* Geneva: WHO.
3. Wilkins E & Wilson L (2017) *European Cardiovascular Disease Statistics 2017 Edition.* Brussels: European Heart Network.
4. Feigin VL, Krishnamurthi RV, Parmar P *et al.* (2015) Update on the global burden of ischemic and hemorrhagic stroke in 1990–2013: the GBD 2013 study. *Neuroepidemiology* **45**, 161–176.
5. Langhorne P, Bernhardt J & Kwakkel G (2011) Stroke rehabilitation. *Lancet* **377**, 1693–1702.
6. Mackay J & Mensah G (2004) *Atlas of Heart Disease and Stroke.* Geneva: World Health Organization.
7. Boehme AK, Esenwa C & Elkind MS (2017) Stroke risk factors, genetics, and prevention. *Circ Res* **120**, 472–495.
8. American Stroke Association (2017) Food for Thought: Heart-Healthy Diet Is also Good for Your Brain. http://www.strokeassociation.org/STROKEORG/LifeAfterStroke/Healthy LivingAfterStroke/Nutrition/Food-for-Thought-Heart-healthy-Diet-is-Also-Good-For-Your-Brain_UCM_449603_Article.jsp# (accessed October 2018).
9. Sherzai A, Heim LT, Boothby C *et al.* (2012) Stroke, food groups, and dietary patterns: a systematic review. *Nutr Rev* **70**, 423–435.
10. Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* **13**, 3–9.
11. Schulze MB, Hoffmann K, Kroke A *et al.* (2003) An approach to construct simplified measures of dietary patterns from exploratory factor analysis. *Br J Nutr* **89**, 409–418.
12. Newby PK & Tucker KL (2004) Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev* **62**, 177–203.
13. Gorst-Rasmussen A, Dahm CC, Dethlefsen C *et al.* (2011) Exploring dietary patterns by using the treelet transform. *Am J Epidemiol* **173**, 1097–1104.
14. Krebs-Smith SM, Subar AF & Reedy J (2015) Examining dietary patterns in relation to chronic disease: matching measures and methods to questions of interest. *Circulation* **132**, 790–793.
15. 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.
16. Gorst-Rasmussen A & Aalborg Universitet Institut for Matematiske Fag (2011) *Tt: Treelet Transform With Stata.* Aalborg: Department of Mathematical Sciences, Aalborg University.
17. Rodriguez-Monforte M, Flores-Mateo G & Sanchez E (2015) Dietary patterns and CVD: a systematic review and meta-analysis of observational studies. *Br J Nutr* **114**, 1341.
18. Saadatnia M, Shakeri F, Keshteli AH *et al.* (2015) Dietary patterns in relation to stroke among Iranians: a case-control study. *J Am Coll Nutr* **34**, 32–41.
19. Stricker MD, Onland-Moret NC, Boer JM *et al.* (2013) Dietary patterns derived from principal component- and k-means cluster analysis: long-term association with coronary heart disease and stroke. *Nutr Metab Cardiovasc Dis* **23**, 250–256.
20. Chan R, Leung J & Woo J (2013) A prospective cohort study examining the associations of dietary calcium intake with all-cause and cardiovascular mortality in older Chinese community-dwelling people. *PLoS One* **8**, e80895.
21. Tjonneland A, Olsen A, Boll K *et al.* (2007) Study design, exposure variables, and socioeconomic determinants of participation in diet, cancer and health: a population-based prospective cohort study of 57,053 men and women in Denmark. *Scand J Public Health* **35**, 432–441.
22. Tjonneland A, Overvad K, Haraldsdottir J *et al.* (1991) Validation of a semiquantitative food frequency questionnaire developed in Denmark. *Int J Epidemiol* **20**, 906–912.
23. Overvad K, Tjonneland A, Haraldsdottir J *et al.* (1991) Development of a semiquantitative food frequency questionnaire to assess food, energy and nutrient intake in Denmark. *Int J Epidemiol* **20**, 900–905.
24. Riboli E, Hunt K, Slimani N *et al.* (2002) European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* **5**, 1113–1124.
25. Laursen ASD, Dahm CC, Johnsen SP *et al.* (2017) Substitutions of dairy product intake and risk of stroke: a Danish cohort study. *Eur J Epidemiol* **33**, 201.
26. Willett W (2012) *Nutritional Epidemiology.* Oxford: Oxford University Press Inc.
27. Slattery ML, Boucher KM, Caan BJ *et al.* (1998) Eating patterns and risk of colon cancer. *Am J Epidemiol* **148**, 4–16.
28. Tjonneland A, Haraldsdottir J, Overvad K *et al.* (1992) Influence of individually estimated portion size data on the validity of a semiquantitative food frequency questionnaire. *Int J Epidemiol* **21**, 770–777.
29. Haraldsdottir J, Tjonneland A & Overvad K (1994) Validity of individual portion size estimates in a food frequency questionnaire. *Int J Epidemiol* **23**, 786–796.
30. Juul S (2012) *Epidemiologi Og Evidens*, 2nd ed. Kbh.: Munksgaard.
31. Lühdorf P, Overvad K, Schmidt EB *et al.* (2017) Predictive value of stroke discharge diagnoses in the Danish National Patient register. *Scand J Public Health* **45**, 630–636.
32. Mohammadifard N, Talaei M, Sadeghi M *et al.* (2017) Dietary patterns and mortality from cardiovascular disease: Isfahan Cohort Study. *Eur J Clin Nutr* **71**, 252–258.
33. Mertens E, Markey O, Geleijnse JM *et al.* (2017) Dietary patterns in relation to cardiovascular disease incidence and risk markers in a middle-aged British male population: data from the Caerphilly Prospective Study. *Nutrients* **9**, 75.

Public Health Nutrition

34. Shimazu T, Kuriyama S, Hozawa A *et al.* (2007) Dietary patterns and cardiovascular disease mortality in Japan: a prospective cohort study. *Int J Epidemiol* **36**, 600–609.

35. Fung TT, Stampfer MJ, Manson JE *et al.* (2004) Prospective study of major dietary patterns and stroke risk in women. *Stroke* **35**, 2014–2019.

36. Larsson SC (2017) Dietary approaches for stroke prevention. *Stroke* **48**, 2905–2911.

37. Wood AM, Kaptoge S, Butterworth AS *et al.* (2018) Risk thresholds for alcohol consumption: combined analysis of individual-participant data for 599 912 current drinkers in 83 prospective studies. *The Lancet* **391**, 1513–1523.

Public Health Nutrition