FOOD INTAKE PATTERNS AND DEVELOPMENT OF OBESITY

PhD Thesis by

Per Togo, MD

Research Unit for Dietary Studies
and
Danish Epidemiology Science Centre
at
Institute of Preventive Medicine,
Copenhagen University Hospital

Copenhagen County Research Centre for Prevention and Health,
Glostrup University Hospital

University of Copenhagen

Supervisors
Merete Osler, MD, D.Sc.
Berit L. Heitmann, dentist, Ph.D.
Thorkild I.A. Sørensen, MD, D.Sc.

Table of Contents

1. PREFACE AND ACKNOWLEDGEMENTS

'At vide hvad man ikke ved er dog en slags alvidenhed' ('after all, knowing what you don't know is a kind of omniscience')

Piet Hein (1905-1996).

This truth should be a comfort to nutritional epidemiologists, who struggle to deliver a consistent message concerning a healthy diet!
Indeed, working with this project has shown me that the association between diet and obesity involves a Pandora's box of known and unknown interacting factors. Nonetheless, it has been an educational process, and hopefully this project contributes to the understanding of the association, and shows new ways to analyse it.

I wish to thank the participants in the MONICA studies who gave their time and efforts to provide the data. The staff working at the Copenhagen County Centre for Preventive Medicine, at Glostrup University Hos-pital, and the Research Unit for Dietary Studies (steering comittee: Berit L. Heitmann, Lillian Mørch Jørgensen, Merete Osler, Agnes N. Pedersen and Marianne Schroll) are thanked for the work of collecting and registering the data, and for making the data available for this project. I owe many thanks for all the suggestions and admo-nitions 'to focus' during the project from Merete Osler[a], Berit L. Heitmann* and Thorkild I.A. Sørensen*, who formed an excellent team of competent advisors. For suggestions regarding methodology and statistics in discus-sions early in the project, I thank Anne H. Andreasen, Jakob B. Bjørner, Klaus Larsen, Svend Kreiner and Niels Keiding. The reviewers Erik Lykke Mortensen, Lauren Lissner and Ulrik Gerdes are thanked for their thorough review of the first edition of the thesis, which has lead to many improvements of the thesis.

I have had a wonderful time at the scientifically and socially privileged Institute of Preventive Medicine in Copenhagen and Centre for Preventive Medicine in Glostrup, and I wish to thank everyone there for making my time there enjoyable.

Finally, I am very grateful for the financial support. It was provided by the University of Copenhagen, DK (grant: j.nr. 301-116-5/99), Else and Mogens Wedell-Wedellsborgs Foundation (donation 664) and the FREJA (Female Researchers in Joint Action) programme from the Danish Medical Research Council, which also financed the establishment of Research Unit for Dietary Studies.

This thesis is dedicated to my dear grandfather Oscar Foght-Sørensen who passed away this winter in the age of 93 years. Last but not least, I owe Inge many thanks for being patient and bearing with me in the numerous hours, days, months behind the computer I have spent during the thesis preparation.

Per Togo, September 2003

The Ph.D.-thesis will be defended October 21st, 2003 at 2 pm in Bartholin auditoriet at Kommunehospitalet, Copenhagen.

[a] Merete Osler is affiliated with *Institute of Public Health, University of Copenhagen*.
Berit L. Heitmann is affiliated with *Research Unit for Dietary Studies and Danish Epidemiology Science Centre at the Institute of Preventive Medicine, Copenhagen University Hospital*.
Thorkild I.A. Sørensen is affiliated with *Danish Epidemiology Science Centre at the Institute of Preventive Medicine Copenhagen University Hospital*.

2. SUMMARY

According to a recent estimate the prevalence of obesity has been increasing in Denmark, and around 300 000 adult Danes are obese today. If the increase follows the trends seen in many other Western countries, the problem of obesity and comorbidities such as diabetes, hypertension, osteoarthritis and some cancers will be an overwhelming challenge for the healthcare system. The habitual diet is one of many factors, which have been proposed to lead to the accumulation of body fat. However, previous studies have failed to give consistent evidence for an (energy independent) effect of macronutrients in the diet on weight gain. Interactions between macro- and micronutrients and the appetite regulating effect of some foods may well be of importance for weight balance. Studying the foods as they are eaten rather than macronutrients in isolation may better capture this interaction. Therefore, we examined whether the combinations of foods in the diet - food intake patterns -have an influence on weight gain.

All previously published papers studying the association between food intake patterns and body mass index (BMI) or obesity development was reviewed. This showed that all 30 studies were cross-sectional and the measurement of body mass was mostly secondary to the different prime targets of the analyses. The results of the studies were inconsistent and the comparison was hampered by the difference in methodology used to define the food intake patterns.

Consequently, we studied this relation in a longitudinal design using data from the Danish MONICA studies, collected by the Research Centre for Prevention and Health at the Copenhagen County Hospital in Glostrup between 1982 and 1994. Food intake patterns were identified by factor analysis on data from a food frequency questionnaire in the participants who completed the food frequency questionnaire as well as a 7-day diet record. The food factors were labelled after their food composition, for men, ‘Green’, ‘Sweet’ and ‘Traditional’ and for women ‘Green’ and ‘Sweet-Traditional’ and they were similar to factors found in other studies. The factors could be reproduced using categorised data from the diet record, and in data from the participants who only filled in the food frequency questionnaire.

Each of the factor scores were tested for association with current BMI, subsequent change in BMI and obesity over a period of 11 years, while adjusting for age, education level, physical activity, smoking, parity, (baseline BMI and previous change in BMI) and the other food factor scores. The cross-sectional analyses showed an inverse association between scores on the ‘Sweet’ or ‘Sweet-Traditional’ factor (men and women, respectively) and BMI. Prospective analyses showed a significant inverse association between baseline score on the ‘Traditional’ factor and subsequent 11-year BMI change for men, whereas for women, an inverse association between baseline score on the ‘Sweet-Traditional’ factor and subsequent five-year BMI change was observed. Fully adjusted analyses of the associations between factor score changes and subsequent BMI changes or the odds of being obese at 11-year follow-up gave no significant results. In conclusion, the associations between food intake patterns and obesity observed in this study were only modest and gender specific. However, the concept of food intake patterns can be further developed for statistical analysis, and should be applied on other data materials that include frequent longitudinal observations of changes in body weight, to establish a more solid evidence of the possible relation.

The PhD.-thesis is based on three published or submitted papers. They are included as the following enclosures I, II and III.

I: Food Intake Patterns and Body Mass Index in Observational Studies.
Togo P, Osler M, Sørensen TIA, & Heitmann BL (2001) *International Journal of Obesity* 25;12:1741-1751

II: Consistency of Food Intake Factors by Different Dietary Assessment Method and Population Groups
Togo P, Heitmann BL, Sørensen TIA, & Osler M (2003) *British Journal of Nutrition* 90: 667–678

III: A Longitudinal Study of Food Intake Patterns and Obesity in Adult Danish Men and Women.
Togo P, Osler M, Sørensen TIA, & Heitmann BL (2001) *Submitted, under review (International Journal of Obesity* May 2003)

See 'Publications, list of tables and figures' for location of tables and figures referred to in the thesis.

3. Aim and Research Questions

The main aim of this project was to:
'Investigate the development of overweight and obesity associated with patterns of food intake'.

The following research questions were addressed:

1. How can food intake patterns be identified and applied in epidemiological studies? (Paper I and II)
2. What have previous investigations shown, regarding the association between food intake patterns and obesity? (Paper I)
3. Which food intake patterns can be identified in a Danish population? (Paper II and III)
4. Can food intake patterns identified using food frequency questionnaire data, be reproduced using data from diet records? (Paper II)
5. Are the food intake patterns identified in one subgroup of the Danish population reproducible in another subgroup of the same population? (Paper II)
6. Are changes in relative weight (body mass index) associated with particular food intake patterns or a change in these patterns in the Danish population? (Paper III)
7. Is obesity development associated with particular food intake patterns or a change in these patterns in the Danish population? (Paper III)

4. BACKGROUND

4.1 THE IMPACT OF OBESITY

Overweight and in particular obesity is an important risk factor for a range of diseases (Bray, 1987; Kopelman, 2000; Visscher & Seidell, 2001). In particular, body mass index (BMI) is associated with diabetes type II and hypertension as demonstrated in several follow-up studies (Colditz et al., 1990b; Carey et al., 1997; Kroke et al., 1998; Brown et al., 2000; Hu et al., 2001). Obesity and in particular abdominal fatness is an important part of the metabolic syndrome, which include a combination of risk factors of cardiovascular disease (including abdominal fatness, insulin resistance, impaired glucose tolerance, dyslipidaemia, microalbuminuria and hypertension). The existence of the syndrome was suggested by Camus, 1966 and described by Reaven, 1988. A definition was proposed by the WHO in 1998 (Alberti & Zimmet, 1998) and recently updated in a review by Groop & Orho-Melander, 2001 as the *dys*metabolic syndrome. In a recent study where changes over time in different components of the syndrome were analysed, BMI, as a measure of obesity, was found to be the central feature of the syndrome in both men and women (Maison et al., 2001).

Other diseases associated with obesity include ischaemic heart disease, osteoarthritis, endometrial-, renal- and breast cancer (Goodman et al., 1997; Oliveria et al., 1999; Shapiro et al., 1999; Friedenreich, 2001; Field et al., 2001). Obese women have higher prevalence of infertility and higher incidence of maternal and fetal complications (e.g. neural tube defects) during pregnancy, delivery and in the postpartum period (Cogswell et al., 2001; Shaw et al., 2002).

Although Denmark has a relatively low prevalence of obesity, compared to the US and most European countries, the secular trend is: an increasing prevalence of obese Danes (Heitmann et al., 2001; Statens Institut for Folkesundhed, 2002), as shown in Table 4-1 (See also Table 6-7 and Figure 6-1 based on follow-up data).

Table 4-1 Prevalence of BMI >= 30 kg/m2 in Denmark 1982-2000

Year:	1982-84	1986-87	1991-92	1994	2000
Copenhagen County Research Centre for Prevention and Health: Men and women 30-60 years of age [a]	9.3%	9.1%	12.0%		
National Institute of Public Health: Men and women 16 years and above [b]		5.5%		7.6%	9.5%

[a] Measured weight and height at study facility
[b] Self-reported weight and height

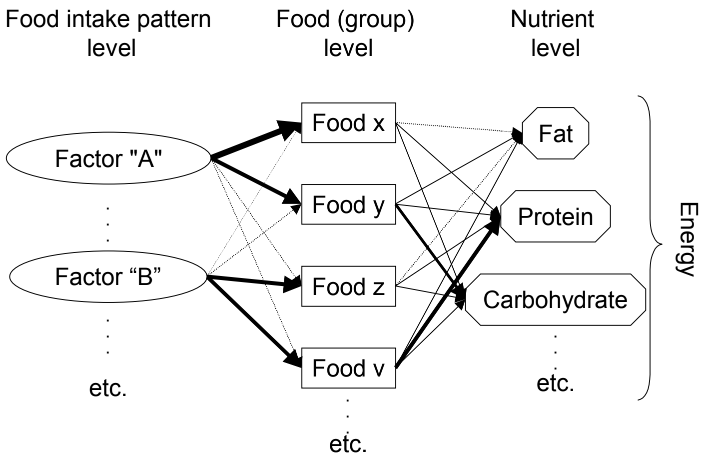
The adverse health effects of obesity are considerable. Thus it has been estimated that 280 000 annual deaths in the US (1991) and 30 000 in the UK (2001) could be avoided, if no Americans or British were obese (Allison et al., 1999; Vass, 2002; The Committee of Public Accounts, 2002). In Denmark, an estimated 14% of premature (adult) deaths were attributed obesity by 1991 (Heitmann et al., 2001). Given the *increase* in the prevalence of obesity and diabetes in Denmark (Beck-Nielsen et al., 2000; Statens Institut for Folkesundhed, 2002), as in the US (Must et al., 1999; Mokdad et al., 2001), this number is likely to increase in the coming years. Thus, obesity is becoming a major factor in the accumulated health and social sector costs (Narbro et al., 1996; Thompson et al., 2001). It has been estimated that in the US up to 5.7% (Wolf & Colditz, 1998) and in Denmark 4-8% of the national health expenditure (Svendsen et al., 2001) could be linked to obesity.

Consequently, there is an increasing call for an understanding of the determinants behind the epidemic and preventive strategies of overweight and obesity. Knowledge of the etiologic factors, which lead to the development of obesity, is a prerequisite for prevention. However, this knowledge is incomplete.

4.2 DIET AND OBESITY

Studying the diet and its possible effect on health requires information on the food intake, most commonly obtained by a food frequency questionnaire, a diet history interview or a diet record completed by the individuals participating in the study. This information can be studied at different levels depending on the comprehensiveness of the dietary assessment instrument. Figure 4-1 illustrates that: the patterns of food intake determine which foods are eaten more often; the foods are composed of nutrients in different proportions and the nutrients add up to an amount of energy intake. The present thesis focuses on the food intake pattern level.

Figure 4-1 Levels of dietary assessment



The diet has traditionally been attributed a major role in the development of obesity, as the source of energy in the energy balance (energy intake – energy expenditure = change in the body's energy stores). Whereas macronutrient composition, in particular the proportion of energy intake from fat, has been (inversely) associated with weight loss in intervention studies (Hill et al., 2000; Astrup et al., 2000), the results from epidemiology do not support an association between fat intake and subsequent weight gain (Lissner & Heitmann, 1995; Jørgensen et al., 1995; Willett, 1998; Willett, 2002). Instead, studying the association between obesity and foods or food groups, as they are eaten, may provide more useful information than studies of macronutrients and obesity.

The term food intake pattern is defined in this thesis as *the distribution by frequency and/or amount of foods in the habitual diet.* The advantages of studying food intake patterns as opposed to studying macronutrients was also described by Hu, 2002, who listed five arguments (here with minor additions/modifications). 1. People do not eat isolated nutrients. Except for very low diversity diets (of poor farmers or in remote areas) the most common way of eating is to *combine* different food groups over the day and at each meal e.g. 'toast, spread and jam', 'meat, bread and different vegetables' etc. 2. Nutrients are inter-correlated. This hampers the study of separate effects of single nutrients. 3. The effect of a single nutrient may be too small to detect. However, *combinations* of foods and their preparation entail the intake of a number of nutrients and possible interactions between the nutrients, which are only captured by studying them as a whole, as food intake patterns. 4. The risk of chance findings is smaller when the many foods and nutrients are reduced to a few pattern scores and entered in the analyses as such, instead of individually. 5. Because nutrient intakes are commonly associated with certain dietary patterns, 'single nutrient' analysis may potentially be confounded by the effect of dietary patterns. Using both the relevant pattern (with which the nutrient is associated) and the nutrient in the same analyses may solve this problem.

4.2.1 Food intake patterns and energy balance

The hypothesis that the food intake pattern is associated with *weight gain* is based on a number of elements related to energy metabolism:

- Differences in satiating effect, energy density, taste, palatability, (neuro-) stimulating ability and 'bioavailability' of the components of the food intake pattern may lead to *different energy intakes*;
- 'Diet induced thermogenesis', energetic cost of preparing/eating/digesting the meal and breaking down its components for storage or 'fuel', and the effect on later physical activity level (foods for sports) may lead to *different energy expenditure*;
- Interactions of nutrients or unknown elements in the food may affect the propensity of the adipocytes to *store energy* as fat, and not adequately signalling to the brain that the necessary amount of food has been eaten.

To answer the second research question: what previous investigations have shown, we reviewed and compared studies of food intake patterns that included body mass index or other measures of relative body weight and/or obesity (Togo et al., 2001, Paper I).

4.3 Food Intake Patterns and BMI in Observational Studies. Review

There are a number of possible ways to describe the food intake patterns of individuals or a group of individuals. Diet index, cluster analysis and factor analysis are the three methods most often used for this purpose (Kant, 1996). Even though there are considerable differences in the methodology, the concept is the same: classifying or ranking subjects according to properties of their habitual diet. A typical diet index is the Healthy Eating Index based on a 10-component system of five food groups, four nutrients, and a measure of variety in food intake (Kennedy et al., 1995). An example of the application of cluster analysis could be the study of Wirfalt & Jeffery, 1997, who found six clusters of subjects based on the percentage of energy from 38 food groups. Whereas cluster analysis is used to segregate individuals into clusters based on resemblance between food values, factor analysis is used to describe the interrelationships between food variables. The use of factor analysis to study food associations was reported in the mid 1970'ies (Howell, 1974; Guthrie & Guthrie, 1976) but from the late 1980'ies (with the help of computers) it has become a significant alternative and supplement to the analysis of macronutrients in nutritional epidemiology (Jacobson & Stanton, 1986; Hu, 2002). The relative merits of each of the three methods are listed in the 'Material and Methods' *choice of method* section (5.3.1); factor analysis is described in general in Appendix 3, and its application was addressed in Paper II and III.

Table 4-2 Factor scores and BMI/obesity in the reviewed studies using factor analysis

Reference	Factors (labels)	**Men**	**Women**
Gex-Fabry et al., 1988	*Satiating*	↓[b]	↓
	Healthful	(↓)	(↓)
	Low culinary complexity	↓[b]	(↓)
Barker et al., 1990	*Traditional*	(↑)	(↓)
	Cosmopolitan	(↓)	(↓)
	Convenience	(↑)	↓[b]
	'Meat & two vegetables'	(↓)	(↓)
Randall et al., 1991	*Salad*	↑	NS
	Fruit	NS	↑
	Healthful	NS	NS
	Traditional (vegetables)	↓	-
	Snacks	NS	-
	High fat	↑[a]	NS
	Whole grain	NS	NS
	Low cost	-	NS
	Light	-	NS
Wolff & Wolff, 1995[c]	*Nutrient dense*		NS
	Traditional		NS
	Transitional		NS
	Nutrient dilute		NS
	Protein rich		NS
	High-fat dairy		NS
	Mixed dishes		NS
Beaudry et al., 1998 [e]	*High-energy density*	NS	
	Traditional	NS	
	Health-conscious	NS	
Gittelsohn et al., 1998 [e]	*Vegetables*	NS	
	Junk foods	NS	
	Bush foods	↑	
	Breakfast foods	NS	
	Hot meal foods	(↑)	
	Tea foods	(↑)	
	Bread & Butter'	NS	
Slattery et al., 1998	*Western*	↑[a]	↑[a]
	Prudent	(↓)	(↑)
	High fat/sugar-dairy	↓[b]	(↓)
	Drinker	↑	↓
	'Substituter'	(↓)	(↓)
	Fruit Juice	↑	-
	'Coffee and Roll'	-	(↓)
Hu et al., 2000 [d]	*Prudent*	↓	
	Western	↑[a]	
Fung et al., 2001 [d]	*Prudent*	(↑)	
	Western	(↑)	
Terry et al., 2001 [c] **New**	*Healthy*		(↓)
	Western		(↓)
	Drinker		↓
Fung et al., 2003 [c] **New**	*Prudent*		(↑)
	Western		(↑)
Sanchez-Villegas et al., 2003 **New**	*Western*	BMI: ↑[a]	Obesity [f]: ↓
	Spanish-Mediterranean	BMI: (↑)	Obesity: ↑

[a b] Factors resembling a Western or high fat diet that were significantly positively (a) or inversely (b) assoc. with BMI
[c] Women only [d] Men only [e] Men and women united; **New:** Additional study, not included in Paper I; [f] 'History of obesity'
↑: Factor score was significantly ($p<0.05$) or non-significantly (↑) positively associated with BMI or obesity
↓: Factor score was significantly ($p<0.05$) or non-significantly (↓) inversely associated with BMI or obesity
NS: Factor scores were not associated with BMI or obesity and the association was not shown or indeterminable

Out of 108 papers identified for the review, 30 were selected. Twelve studies used a diet index, nine cluster analyses and nine used factor analysis to characterise populations by food intake pattern. Details of the material and methods of each study are given in Paper I, Tables 1 through 3. The associations between food intake and BMI/obesity in the 30 studies are summarised in Paper I, Tables 4 to 6 (which include more detailed descriptions of the indexes, clusters or factors). Only the results of the studies using *factor* analysis are included in Table 4-2, in which three new studies (published after the review analysis) using factor analysis in different populations are added (Terry et al., 2001, Fung et al., 2003 and Sanchez-Villegas et al., 2003).

In brief, the review showed that ten (three using factor analysis) of the 30 studies were in agreement with the hypothesis that a food intake pattern reflecting a Western diet (here defined as: more fatty food, more meat, less fruit and vegetables) was positively associated with BMI (or a BMI >=30). Four studies (two using factor analysis) showed inverse associations and 11 studies (three using factor analysis) non-significant associations with BMI. Part of the inconsistency can be explained in the heterogeneous food intake patterns derived from different dietary assessment methods, variation in number of foods and factors included, difference in cooking traditions and properties of foods (e.g. fat content) in different populations/countries etc. In addition, the limitations of the methods discussed later may apply differently for different populations (e.g. the accuracy of the diet reporting). However, the labels ‘prudent’ or ‘healthful’ and ‘Western’ or ‘Traditional’, respectively, seem to cover ‘universal’ patterns of eating and they have often been used for factors.

In all the studies reviewed, the food intake factors and factor scores (the individuals’ relative intake of the foods included in the factor) were not evaluated for reproducibility, using different dietary assessment methods and population groups. Further, all studies were cross-sectional in design, with regard to food intake and weight assessments, and hence could not take into account the temporal relation between the food intake and weight (and BMI) development.

The analyses described in Paper II and III were designed to meet these shortcomings and to investigate the association between food intake patterns and obesity development in a Danish population.

5. MATERIAL AND METHODS

5.1 STUDY POPULATION

The population under study represent part of several populations investigated by the Copenhagen County Research Centre for Prevention and Health (Københavns Amts Center for Forebyggelse og Sundhed, formerly known under the names Copenhagen County Centre of Preventive Medicine and the Glostrup Population Studies) and harmonised at the Research Unit for Dietary Studies (Enhed for Epidemiologisk Kostforskning). (DAN)-MONICA is the Danish contribution to the WHO initiated international MONICA survey (Monitoring of trends and determinants in cardiovascular diseases, MONICA data centre, 2000).

5.1.1 The baseline survey (M-82)

The population invited for the baseline examination (MONICA 1, M-82 in the text) (1982-84) included 4807 citizens, aged approximately 30, 40, 50 or 60 years. The group was a random sample of the Danish population, selected from the Central Person Register among citizens who all lived in the Western part of the Copenhagen County. Of the invited, 79 percent (n=3785) attended M-82, the '*study-population'* in Figure 5-1. The remaining 21 % who did not attend the examination have been found not to differ from the attendees with respect to gender, age, educational level, occupation and housing, except for a slight under-representation of unemployed and self-employed, as well as people employed in agriculture, horticulture and fishery. Characteristics of the attendees and the non-attendees were given in a previous paper (Jørgensen, 1987).

5.1.2 Follow-up of the cohort (M-87, M-93)

The subjects invited for M-82 were re-examined twice: in 1987-88 ('*GenMONICA'* = M-87) and in 1993-94 ('*MONICA 10'* = M-93). Of the 3785 baseline attendees, 550 (15%) participated in M-87 only and 117 (3 %) in M-93 only. In total, 2436 participants (51% of the invited for- and 64% of the people examined at- baseline) attended the full examination programme - M-82, M-87 and M-93.

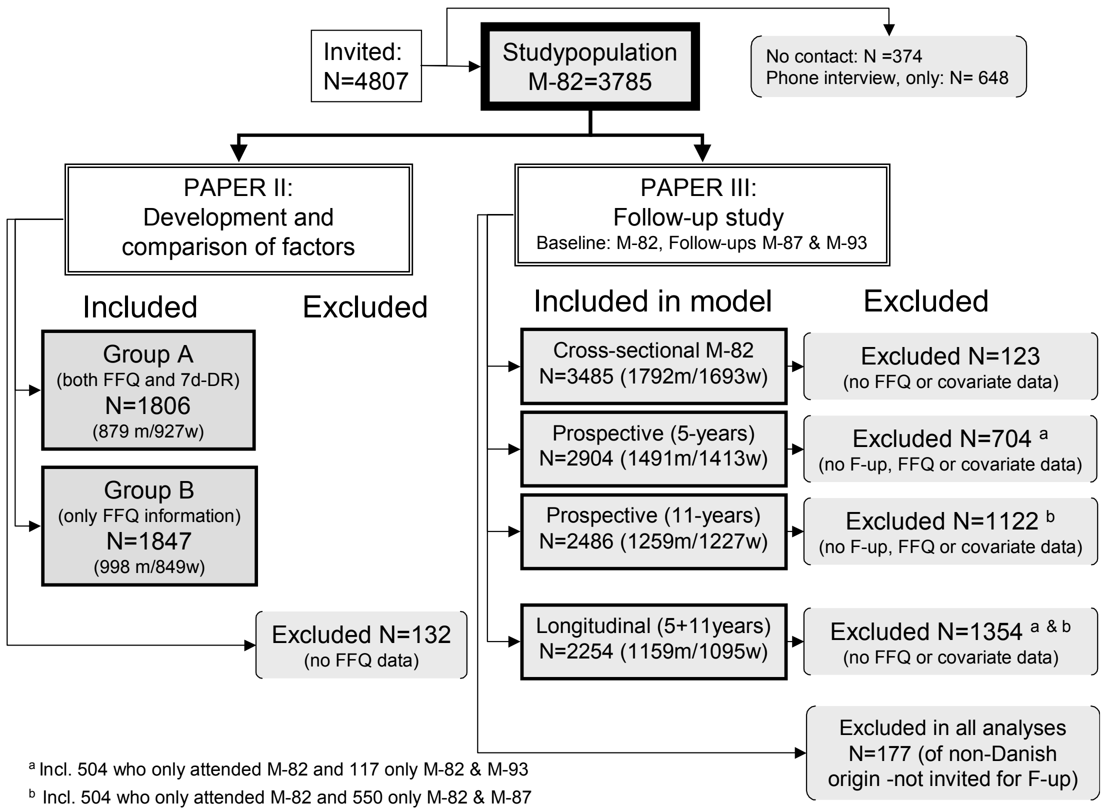
5.1.3 Number of subjects included in the analyses reported in Paper II - III

At M-82, 3629 (95.9%) and at M-87, 2740 (91.7%) of the attendees completed the food frequency questionnaire. In a few subjects, the factor scores could not be estimated by the program and they were excluded from additional analysis (and excluded as 'no FFQ data' along with those who did not complete the FFQ). Approximately half of the population (n = 1851) completed a seven-day diet record.

In Paper II (Figure 5-1, left), the analyses of the question, whether food intake patterns could be reproduced in data based on different dietary assessment methods, were conducted using data from the 869 men and 927 women who filled in both food frequency questionnaires (FFQ) and 7-day diet records (group A). In addition, data from the 998 men and 849 women who only completed the FFQ (group B) were used for analyses of the reproducibility of the food intake pattern in different population groups.

In Paper III (Figure 5-1, right), the number of subjects included were maximised for each analysis and therefore depended on the available dietary, anthropometrical and covariate data. The cross-sectional and the prospective analyses (predicting BMI at 5 and 11-year follow-up) were based on the subjects who had dietary information at M-82 (FFQ) and the relevant follow-up data. The longitudinal analysis required most information, as the dietary information at M-82 *and* M-87 as well as BMI and information on a number of covariates at different time-points were used (see Figure 5-1 for details on the number of subjects in each analysis).

Figure 5-1 Number of participants (N) in the analyses for Paper II and III



5.2 Variables

5.2.1 Assessment of habitual food intake

Only a smaller part of the population filled in a diet record or gave a diet history interview at follow-ups, and therefore, the identification of the food factors was based on the food frequency questionnaire (FFQ), which was distributed to all subjects at baseline and follow-ups. However, diet-record data was used to reproduce the food factors identified, and analyse the consistency of the correlation between factor scores, based on the two dietary assessment methods, in subgroups defined by age and other covariates as described in Paper II.

Food frequency questionnaire

All study participants were asked to complete questions on how frequent foods from 26 food groups usually were eaten.

Intake frequency categories were:

1: never	5: two to three times a week
2: once a month or less	6: once a day
3: twice a month	7: two to three times a day
4: once a week	8: four or more times a day

The participants were asked to state which category of intake frequency best fitted their usual intake. The food groups included are listed in the first column of Table 1 in the Appendix along with the Danish labels. In less than 4.1% and 8.3% of the attendees, at M-82 and M-87 respectively, the FFQ information needed for the analyses was incomplete (as reported under 5.1.3). Those cases were removed by list-wise deletion, which was preferred to imputation techniques that would be more complicated for the skewed, ordered, categorical FFQ data. The missing information was at random, not dependent on age, gender, smoking-status, physical activity, education or BMI group when analysed by logistic regression (data not shown). It was assumed that the ranking of subjects was not affected by the missing options '3 times a month' or '4-6 times a week'.

Diet record

The diet record was completed on a pre-typed form containing 113 foods and drinks and empty lines for additional foods (See Paper II and Appendix 2. Translations' for a list of the foods included in the diet record and how they were grouped and categorised in octiles of gram intake/week to match the FFQ variables). The record should reflect a 'normal' week within three weeks from the examination and completion of the FFQ. Based on the diet record data, macronutrient and energy composition of the diet was computed (by a dietician using Microcamp® by Mørup, 1986) and used to characterise the factor scores in Paper II.

5.2.2 Anthropometry

Height was measured to the nearest 0.5 cm with subjects standing without shoes, heels together and head in horizontal Frankfurter plane. Body weight was measured to the nearest 0.1 kg using a SECA scale, subjects wearing only light indoor clothes. As a measure of relative weight and as a simple measure of body fat, Body Mass Index = bodyweight/height^2 (kg/m^2) was computed for all available examinations. For the prospective and longitudinal analyses, height at baseline only, was used to disregard BMI differences due to changes in height alone (more information in Paper III). Subjects who had a BMI $\geq$ 30 kg/m^2 were categorised as obese.

5.2.3 Other covariates

Physical activity level in leisure time was assessed by a four-category question. *Smoking habits* were included in the questionnaires distributed at all surveys. Since the same questions were not asked at follow-up, a new universal variable with three categories (non-smoker, ex-smoker and smoker) was defined for use in all analyses. *Parity* was included as a five-category variable and women who gave birth during the first follow-up period were excluded. *Educational level (school or vocational)* serves as an indicator of social position in the analyses. This variable was highly correlated with the other social position indicators such as schooling, own-, partner's- or family's income and 'social class' that were not included in the analyses. For more details on covariates, see the Papers II and III.

5.3 ANALYSIS STRATEGY AND DESIGN OF THE STUDY

This section was included to highlight the methodological considerations we had, regarding how to answer the third research question: which food intake patterns can be identified in a Danish population.

5.3.1 Choice of method for the identification of food intake patterns

The most important advantages and drawbacks of each of the three methods most commonly used to identify food intake patterns can be listed as in Table 5-1.

Table 5-1 Some advantages and drawbacks of three common methods to identify food intake patterns in epidemiological studies

	Advantages	*Drawbacks*
Diet Index	✓ Simple to use ✓ No advanced computation needed ✓ Data independent ✓ Continuous outcome possible	A priori knowledge or hypothesis needed Arbitrary definition of cut-points or scoring method No biological foundation
Cluster Analysis	✓ Simple to understand ✓ Relatively easy to compute with few variables ✓ Good at classifying subjects	Highly data dependent No biological foundation Difficult to measure changes Highly data reductive Categorical No gold standard for choosing number of clusters Difficult to use with many variables with small variance Sensitive to variables with skewed distribution Within cluster variation can be large
Factor Analysis	✓ Flexible if used 'confirmatory' ✓ Testing of different factor solutions possible (if 'confirmatory') ✓ Advanced modelling with inclusion of covariates possible (if 'confirmatory') ✓ Suitable for longitudinal analysis (if used 'confirmatory') ✓ Continuous output and less data reduction ✓ Each subject can be ranked by several factors ✓ Positive use of colinearity	Difficult to understand Data dependent especially if used 'exploratory' Gold standard for choosing number of factors not useful for low variance input as dietary intake Not necessarily biological foundation (but highlighting correlations among several variables)

Factor analysis was found to be the most appropriate for identification of food intake patterns in this investigation for a number of reasons. 1) The current knowledge of the relation between food intake and obesity is limited and most research has focused on associations between macronutrients and obesity, not on food groups and obesity. Furthermore, associations with macronutrients are inconsistent. This makes it difficult to compose a diet index *a priori*. 2) The longitudinal design with two to three repeated dietary assessments allows for an analysis of *change* in food intake pattern, for which the cluster analysis is less suitable. 3) Factor analysis has been used in many studies, and can be applied in an exploratory as well as confirmatory fashion, which can be used to identify and confirm the food intake patterns, respectively.

The *confirmatory* factor analysis method was preferred to compute the factor scores for the longitudinal study (Paper III) because it could reflect a change in adherence to a well-defined food intake pattern over time as a change in score, and potentially be used in another population. Using the principal components analysis only (without rotation of factors) or the 'exploratory' factor analysis without the selection of relevant foods would produce factors that were more complicated, which would be difficult to use with repeated measurements because of data dependency. However, changing the factor extraction methodology or rotation method did not change the factor structure much (data not shown).

5.3.2 Identification of food intake patterns by factor analysis

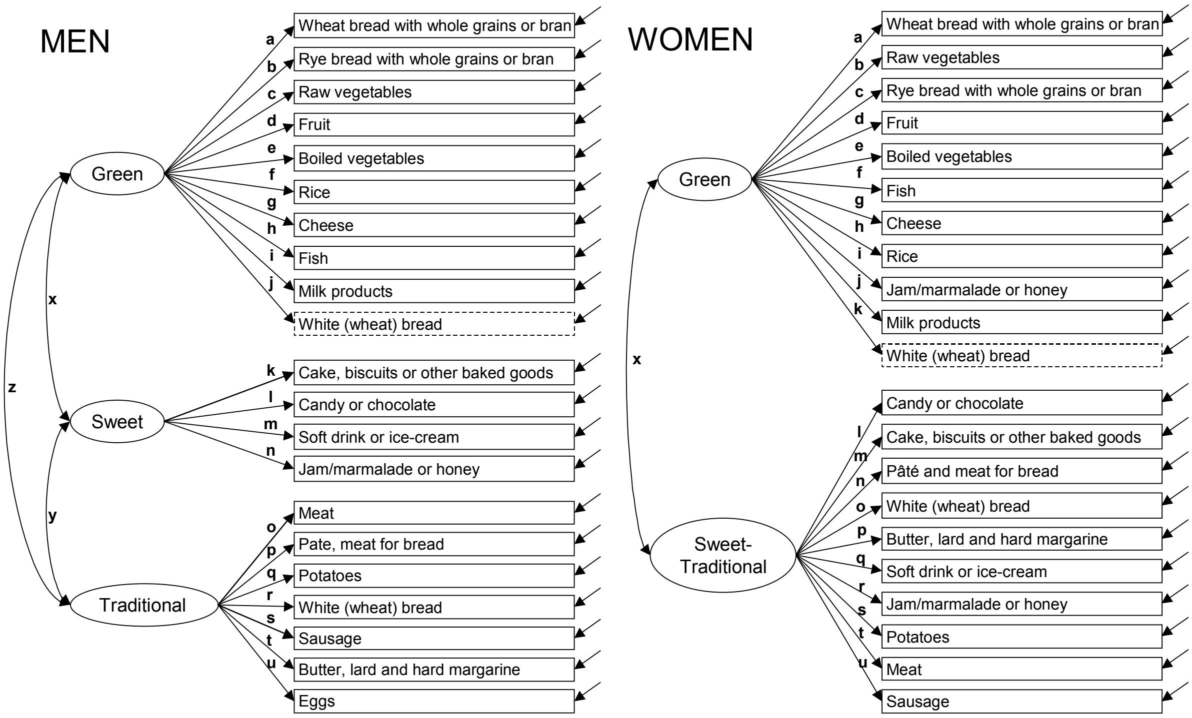
Using the factor analytic framework, the overall food intake pattern is determined by a number of *latent variables.* The latent variables could be e.g.: a higher compliance with dietary recommendations to adapt a Mediterranean diet, a craving for sweet and tasty foods or a continuation of established dietary habits. Each latent variable can be identified by a number of *indicator foods,* the intakes of which hypothetically are particularly 'influenced' by the latent variables. Factors that approximate the true latent variables can be constructed/computed by factor analysis. The correlation matrix of all the individuals' intake frequencies of the indicator foods is used to determine these food intake factors. A score can be computed for each individual on each factor, reflecting the individuals inclination to the intake of the given indicator foods, and -hypothetically - to the overall latent pattern of intake. See also Appendix 3 regarding factor analysis in general and score computation in Mplus.

The identification of the food intake patterns in this study was based on an exploratory factor analysis on the group of participants who had both FFQ and diet record information (Paper II, scree plot Figure 1 and loadings in Table 2).

For men, the final factors were labelled *'Green'* (highest loadings on wheat bread and rye bread with whole grains/bran; raw- and boiled vegetables and fruit), *'Sweet'* (cake; biscuits and baked goods; candy or chocolate; soft drink or ice-cream and jam, marmalade or honey) and '*Traditional'* (meat; pâté and meat for bread; potatoes; and white bread). For women, the factors were labelled '*Green*' (high loadings on the same foods as for men) and '*Sweet-traditional*' (candy or chocolate; cake, biscuits and baked goods; pâté and meat for bread; white bread; butter, lard and hard margarine; and soft drink or ice-cream).

Absolute loadings >0.3 in the exploratory factor analysis on the FFQ data determined which food groups were identifying each factor and those food groups were included in a confirmatory analysis to compute factor scores (Model in Figure 5-2, Loadings in Paper II, Table 3, score computation described in Appendix 3). The analyses are described in depth in Paper II-III.

Figure 5-2 Food intake factors as modelled on the baseline data to compute factor scores. X, Y and Z denote factor-factor correlations and letters A through U denote factor loadings of the food groups in boxes (negative loading if dashed box and different loadings for men and women)



5.3.3 Analysis of non-participation in follow-up (dropout)

Dropout may bias the association between food intake patterns and obesity development. Consequently, comparisons were made between the 2437 subjects who attended full follow-up (completers) and the remaining 1348 who did not (i.e. those who only participated in the baseline examination, or in the baseline and only *one* of the two follow-up examinations = dropouts). This was done in two ways: by linear regression analyses of the cross-sectional associations between factor scores and BMI at baseline stratified by participation in subsequent follow-up, and by logistic regression analyses of the associations between the covariates and dropout status adjusted for age. (Results, Table 6-1 and Table 6-2, respectively; number of participants were reduced to those who had complete dietary and covariate information).

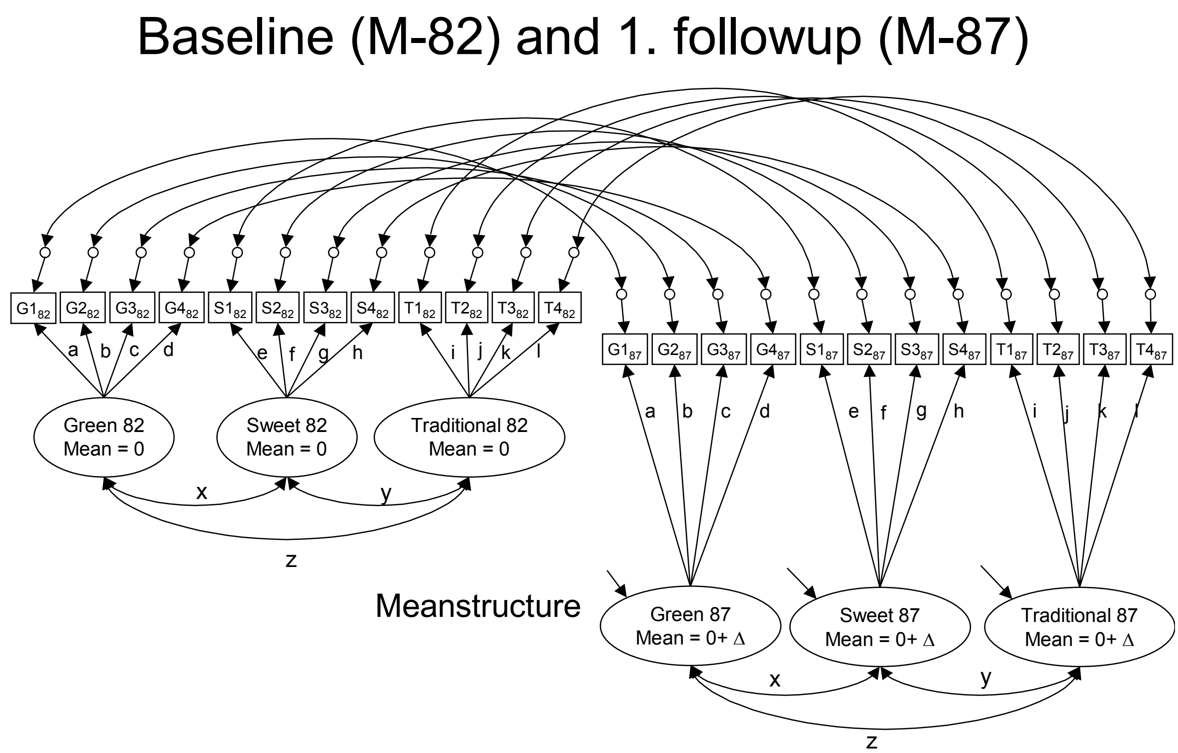
5.3.4 Consistency of Food Intake Factors (Paper II)

The dietary data from the FFQ and the diet record were harmonized as described in Paper II. Exploratory factor analyses were carried out in both types of dietary assessment data to determine if similar factors were identified. Confirmatory factor analyses were performed on both FFQ and diet record data and loadings of the food groups compared. The resulting scores for each sex using the two factor analysis methods and the two dietary assessment methods were compared by correlation analysis. Differences in correlations in subgroups defined by the covariates were examined. The correlations between factor scores and macronutrient and energy composition were described. In addition, exploratory and confirmatory factor analysis, using the '*a priori'* model (Figure 5-2) developed using data from the participants who provided both diet record and FFQ information (group A), were carried out using data from participants who only had FFQ (group B).

5.3.5 Food Intake Pattern and Obesity (Paper III)

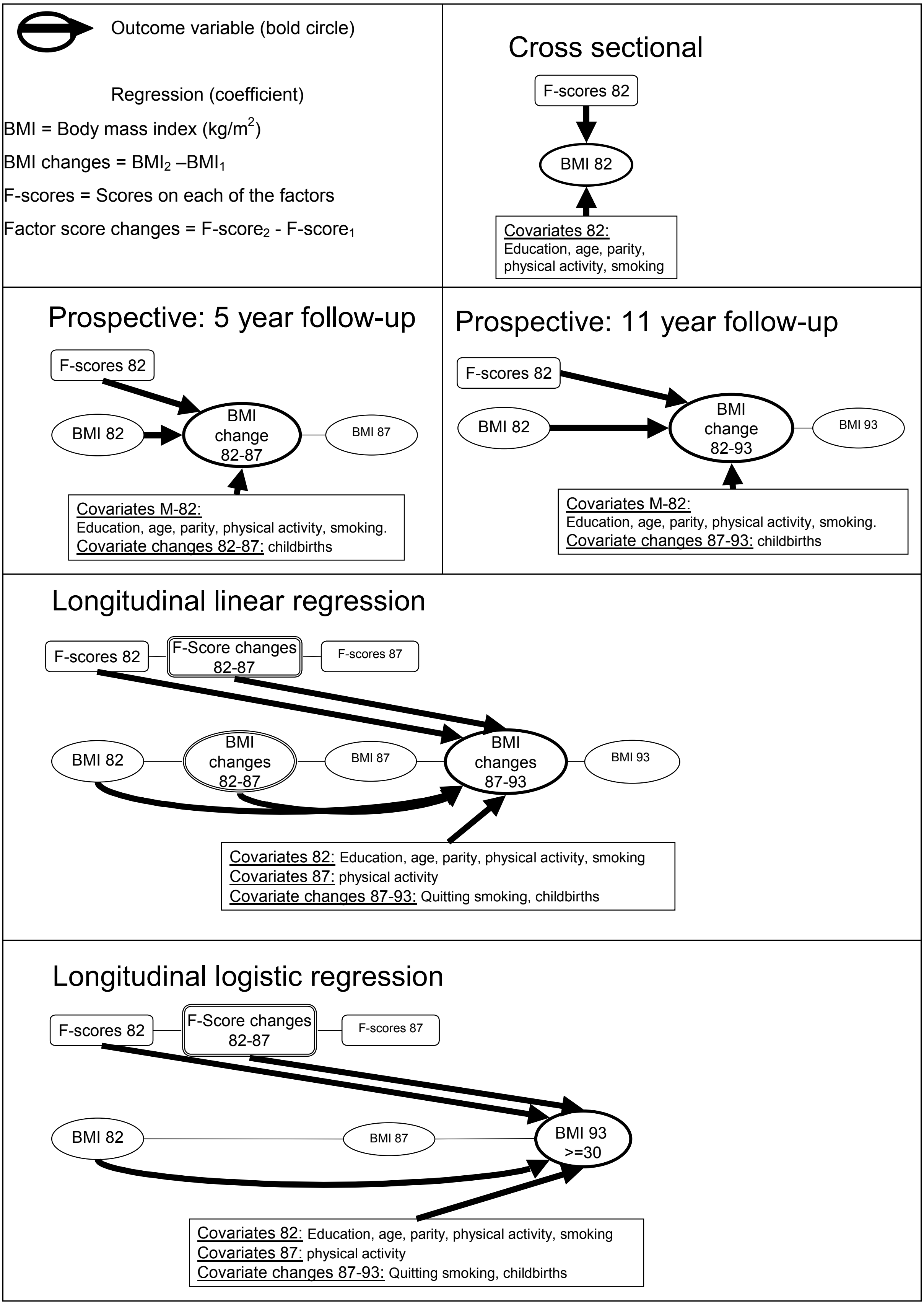
Analyses of the possible associations between food intake patterns, BMI, BMI change and obesity were conducted using different regression model designs (Figure 5-4).
To investigate the associations between *changes* in food intake pattern over time, food frequency data from baseline (M-82) and first follow-up (M-87) was incorporated in the confirmatory factor analysis model shown in Figure 5-3.
The FFQ contained the same foods at the M-82 and M-87 examinations and the factor model was constructed in a way, which 'forced' the factors to be identified by the same indicator foods at each examination (e.g. Green 82 by the foods $G1_{82}$, $G2_{82}$, $G3_{82}$ and $G4_{82}$ and Green 87 by $G1_{87}$, $G2_{87}$, $G3_{87}$ and $G4_{87}$ in Figure 5-3). In addition, the loadings of each indicator food on the factors (a, b, c, d etc) and the factor-factor correlations (x, y, z) were set to be the same over time. Further, a 'mean-structure' was applied to allow the factor score means to change over time from M-82 to M-87 - instead of being standardized to 0 - to reflect changes in diet (See Appendix 4. 'Model specification for the computation of factor scores in Mplus with comments').

Figure 5-3 Model for the factor score computation in the longitudinal analyses, names of the four or more foods loading on each factor were replaced by G1, G2, G3, G4; S1, S2, S3 S4; T1, T2, T3, T4 for simplification. Loadings (a to l in the figure), and factor correlations (x, y, z in the figure), were fixed to the same values over time. See paper III and Appendix 3 and 4 for additional information. This model is for MEN (for women the same model with two factors was used)



The computed factor scores and difference between scores in M-82 and M-87 (difference = Δ factor = score M-87 minus score M-82) were entered as covariates in linear or logistic regression models with BMI-change or obesity as dependent variables, while controlling for age, physical activity, smoking, education, baseline BMI and parity (see Paper III and Figure 5-4 for details)

Figure 5-4 Outline of the designs of the adjusted regression analyses of the associations between food factor scores and BMI, BMI-change or obesity



6. RESULTS

This chapter provides analyses of non-participation, summaries of the observations in Paper II and III, distributions of factor scores and BMI in more detail and an extraction of the major gender differences observed in the study.

6.1 NON-PARTICIPATION IN FOLLOW-UP (DROPOUT)

Our main concern regarding dropout was that the relation between factor scores and BMI was different in the two groups of baseline participants (completers and those who later dropped out), which could indicate that the groups were different with regard to the association between diet and BMI development.

Table 6-1 Cross-sectional linear regression analysis of the crude and adjusted association between factor scores and BMI at baseline, stratified by sex and participation in subsequent follow-up (completers vs. dropouts)

	Crude[a] associations with BMI				**Adjusted[b] associations with BMI**			
MEN	*Completers* (N=1215)		*Dropouts* (N=585)		*Completers*		*Dropouts*	
	beta	95 % CI	Beta	95 % CI	Beta	95 % CI	Beta	95 % CI
Factor scores								
Green	-0.33	(-0.70; 0.05)	0.37	(-0.20; 0.94)	-0.04	(-0.41; 0.32)	0.58*	(0.00; 1.16)
Sweet	-0.86**	(-1.26;-0.47)	-0.49	(-1.07; 0.09)	-0.72**	(-1.10;-0.34)	-0.53	(-1.11; 0.04)
Traditional	-0.44	(-1.12; 0.23)	0.27	(-0.86; 1.39)	-0.29	(-0.94; 0.35)	0.27	(-0.84; 1.37)
Covariates								
Age[c]					0.70**	(0.53; 0.87)	0.51**	(0.24; 0.78)
P.A.[d]					-0.43**	(-0.70;-0.17)	-0.18	(-0.63; 0.27)
Education					-0.26**	(-0.38;-0.13)	-0.31**	(-0.52;-0.09)
Smoking					-0.65**	(-0.89;-0.42)	-0.71**	(-1.19;-0.24)
WOMEN	*Completers* (N=1175)		*Dropouts* (N=539)		*Completers*		*Dropouts*	
	beta	95 % CI	Beta	95 % CI	Beta	95 % CI	Beta	95 % CI
Factor scores								
Green	-0.12	(-0.61; 0.37)	-0.42	(-1.19; 0.35)	0.21	(-0.28; 0.70)	0.00	(-0.76; 0.77)
Sweet-Traditional	-0.94**	(-1.41;-0.48)	-0.77*	(-1.50;-0.04)	-0.84**	(-1.28;-0.39)	-0.80*	(-1.48;-0.11)
Covariates								
Age[c]					0.85**	(0.65; 1.06)	-0.80*	(-1.48 -0.11)
P.A.[d]					-0.54**	(-0.87;-0.21)	0.96**	(0.63; 1.29)
Education					-0.18*	(-0.31;-0.04)	-0.01	(-0.60; 0.59)
Smoking					-0.37**	(-0.60;-0.13)	-0.34**	(-0.57;-0.10)
Parity					0.29**	(0.08; 0.50)	-0.94**	(-1.37;-0.51)

* p< 0.05; [a] Model included the factor scores and constant only; [b] Model included the variables listed + constant;
[c] 30 years=3, 40=4, 50=5, 60=6; [d] P.A. = physical activity in leisure time

Although the associations in the dropouts were weak (partly due to the smaller sample size), and there were no *significant* results suggesting that they should be different from the completers in this respect, the factor score-BMI associations in those who dropped out were in some cases inverse of the findings in the completers (Table 6-1).

The logistic regression (Table 6-2) showed that the groups *were* different in the way that the odds of dropout tended to be higher in the older, the less active, the less educated, the smokers, the obese and the women who bore many children. In addition, the odds of dropout were lower by increasing score on the 'Green', 'Sweet' and 'Sweet-Traditional' factor.

Table 6-2 Age adjusted logistic regression analysis with odds ratios of dropout (lowest = reference group) [a] and χ^2- tests for trend (linear by linear association)

		MEN			WOMEN		
		OR	95% CI	P	OR	95% CI	P
Age (years)	30	1			1		
	40	0.79	(0.59 - 1.05)	0.10	1.14	(0.86 - 1.52)	0.37
	50	1.04	(0.78 - 1.38)	0.79	1.29	(0.97 - 1.72)	0.08
	60	2.02	(1.54 - 2.65)	0.00	1.99	(1.49 - 2.65)	0.00
	χ^2-test for 'trend'			<0.001			<0.001
Factor scores	Green	0.76	(0.63 - 0.93)	0.01	0.61	(0.49 - 0.77)	0.00
(by one unit increase)	Sweet	0.56	(0.45 - 0.68)	0.00			
	Traditional	1.40	(0.97 - 2.02)	0.07			
	Sweet-Traditional				0.71	(0.57 - 0.88)	0.00
Physical activity	Sedentary	1			1		
(leisure time)	Walking etc	0.66	(0.52 - 0.84)	0.00	0.68	(0.55-0.84)	0.00
	Sports/heavy gardening	0.47	(0.35 - 0.62)	0.00	0.42	(0.29-0.61)	0.00
	χ^2-test for 'trend'			<0.001			<0.001
Education	None	1			1		
	Semiskilled worker	1.11	(0.73 - 1.68)	0.63	1.20	(0.87 - 1.64)	0.27
	One-year education (theoretical)	0.98	(0.53 - 1.80)	0.94	0.93	(0.64 - 1.36)	0.72
	Basic vocational courses/apprentice	0.74	(0.53 - 1.03)	0.07	0.77	(0.58 - 1.02)	0.07
	Shorter theoretical edu.	0.39	(0.23 - 0.66)	0.00	0.55	(0.36 - 0.85)	0.01
	Medium length theoret. edu. (3-4y)	0.63	(0.41 - 0.97)	0.04	0.42	(0.26 - 0.67)	0.00
	Longer edu. (5+ y) or student	0.36	(0.19 - 0.66)	0.00	1.53	(0.72 - 3.27)	0.27
	χ^2-test for 'trend'			<0.001			<0.001
Smoking	Never-smoker	1			1		
	Ex-smoker	1.14	(0.79 - 1.63)	0.49	1.13	(0.81 - 1.58)	0.47
	Smoker	2.14	(1.58 - 2.89)	0.00	2.14	(1.69 - 2.71)	0.00
	χ^2-test for 'trend'			<0.001			<0.001
BMI-group	Normal weight	1			1		
	Overweight (>=25)	1.18	(0.95 - 1.46)	0.13	1.24	(0.97 - 1.59)	0.09
	Obese (>=30)	1.39	(0.99 - 1.94)	0.06	1.41	(0.97 - 2.03)	0.07
	Underweight (<18.5)	2.11	(0.80 - 5.57)	0.13	1.53	(0.89 - 2.62)	0.12
	χ^2-test for 'trend'			0.01			0.02
Parity	No children				1		
	1 child				1.11	(0.77 - 1.60)	0.57
	2 children				0.92	(0.66 - 1.28)	0.60
	3 children				1.25	(0.85 - 1.82)	0.26
	4 or more children				1.70	(1.08 - 2.68)	0.02
	χ^2-test for 'trend'						0.01

[a] One analysis for each covariate. Model = constant + age + covariate (except for analysis with age as only predictor, and factor scores which were entered all at once with age in the model)

6.2 Consistency of Food Intake Factors by Different Dietary Assessment Method and Population Groups

6.2.1 Food intake factors by different dietary assessment methods

Spearman correlations between the food groups in the FFQ and the corresponding food groups in the diet record ranged between meat = 0.10 to 'rye bread with whole grains/bran' or 'milk or yoghurt' = 0.68 (Paper II, Table 1). Very similar factors were found by exploratory factor analyses on data based on the FFQ (Paper II, Table 2) and the diet records, and scores on the corresponding factors based on the different dietary assessment methods had a correlation between 0.34 and 0.61.

Table 6-3 Pearson correlation coefficients[a] between factor scores on similar factors based on food frequency questionnaire and diet record data in exploratory and confirmatory factor analyses, respectively (from Paper II)

	MEN			WOMEN	
Correlations between factors based on	'Green'	'Sweet'	'Traditional'	'Green'	'Sweet-Traditional'
Different dietary assessment methods					
Exploratory FA[b] on FFQ[c] data vs. diet record data	0.61	0.55	0.34	0.61	0.57
Confirmatory FA on FFQ data vs. diet record data	0.57	0.53	0.37	0.64	0.56
Different factor analysis methods					
Exploratory FA vs. confirmatory FA on FFQ data	0.95	0.94	0.91	0.95	0.96
Exploratory FA vs. confirmatory FA on diet record data	0.82	0.90	0.90	0.94	0.94

[a] All Correlations, $p < 0.01$
[b] FA: Factor Analysis
[c] FFQ: Food Frequency Questionnaires

Stratification by age, BMI, energy intake, education, physical activity and smoking did not alter the correlations markedly with a few exceptions described in Paper II. Scores based on the confirmatory factor analyses (on the foods loading more than 0.3 in the exploratory factor analyses, Paper II, Table 3) had about the same correlation between dietary assessment methods, and the factor scores were highly correlated with the exploratory factor score counterparts (see Table 6-3).

6.2.2 Food intake factors by different population groups

Very similar factors were identified by exploratory factor analyses on the two population groups defined by completion of a FFQ *and* a diet record (group A) vs. completion of FFQ, only (group B). This is illustrated by the loadings in Table 6-4 and suggests that the factors are reproducible in different population segments.

Table 6-4 Comparison of loadings [a] in exploratory factor analyses of FFQ-data using participants completing both the FFQ and a diet record (A) and participants completing the FFQ, only (B). Used for scatterplots in Paper II, Fig. II-2,3

	MEN						**WOMEN**			
	'Green'		'Sweet'		'Traditional'		'Green'		'Sweet-Traditional'	
Food groups	A	B	A	B	A	B	A	B	A	B
Wheat bread w. whole grains/bran	**0.60**	**0.37**	0.07	0.24	-0.23	**-0.36**	**0.62**	**0.58**	-0.07	-0.02
Raw vegetables	**0.60**	**0.53**	-0.09	0.01	0.01	-0.13	**0.58**	**0.58**	-0.13	-0.22
Boiled vegetables	**0.57**	**0.68**	-0.10	-0.12	0.20	0.17	**0.41**	**0.43**	-0.09	-0.10
Rye bread with whole grains/bran	**0.55**	**0.43**	0.05	0.15	-0.17	-0.23	**0.57**	**0.51**	-0.12	-0.11
Fruit	**0.50**	**0.50**	0.20	0.13	0.02	0.00	**0.56**	**0.56**	0.10	0.07
Fish	**0.41**	**0.47**	-0.06	-0.10	0.24	0.20	**0.41**	**0.35**	-0.06	-0.06
Rice	**0.36**	**0.31**	0.29	0.27	0.00	-0.13	**0.38**	**0.33**	0.25	0.03
Cheese	**0.31**	**0.45**	0.01	-0.04	0.12	0.15	**0.31**	0.29	0.03	-0.03
Milk or yoghurt	**0.30**	0.15	0.13	0.24	0.17	0.09	**0.34**	**0.35**	0.10	0.24
Cake, biscuits and baked goods	0.07	0.02	**0.75**	**0.73**	0.00	0.10	0.23	**0.36**	**0.61**	**0.54**
Candy or chocolate	-0.05	-0.08	**0.73**	**0.79**	0.01	0.10	0.17	**0.32**	**0.59**	**0.54**
Soft drink or ice-cream	-0.12	-0.18	**0.56**	**0.61**	0.04	0.11	-0.07	0.15	**0.43**	**0.43**
Jam, marmalade or honey	0.18	0.15	**0.55**	**0.44**	0.01	0.11	**0.32**	**0.32**	**0.41**	**0.39**
Meat	0.07	0.21	-0.08	0.03	**0.63**	**0.63**	-0.10	-0.21	**0.33**	**0.45**
Pâté and meat for bread	0.13	-0.03	0.01	0.19	**0.61**	**0.58**	-0.09	-0.17	**0.55**	**0.59**
Potatoes	-0.02	0.24	-0.15	-0.14	**0.54**	**0.61**	-0.07	-0.12	**0.37**	**0.45**
Sausages	0.07	-0.01	0.13	0.27	**0.49**	**0.45**	-0.03	-0.10	**0.46**	**0.46**
White (wheat) bread	**-0.38**	-0.24	0.24	0.04	**0.41**	**0.56**	**-0.41**	**-0.35**	**0.47**	**0.50**
Butter, lard and hard margarine	-0.02	0.02	0.16	0.25	**0.37**	**0.38**	-0.05	-0.07	**0.45**	**0.50**
Eggs	0.29	**0.38**	-0.16	-0.07	**0.35**	**0.32**	0.30	0.10	0.12	0.05
Vegetable fat margarine/spread	0.10	0.05	0.14	0.09	0.20	0.20	0.12	0.07	0.26	0.08

[a] Loadings >0.30 or < -0.30 are **boldfaced**

6.2.3 Factor scores, covariates and nutrient intake

Factor scores based on confirmatory factor analysis using FFQ data, were analysed for linear trends across strata of covariates and correlations with energy and nutrient energy contribution (E%). Directions of the associations are shown in Table 6-5.

Table 6-5 List of associations between covariates, nutrient intake and factor scores

	MEN			**WOMEN**	
Factor scores	Green	Sweet	Traditional	Green	Sweet-Traditional
Linear trend (ANOVA) by					
Age	↓	↓	0	↓	0
Leisure activity	↑	↑	0	↑	0
Education	↑	↑	↓	↑	0
Smoking	↓	↓	↑	↓	↓
Parity				0	0
Correlation with					
Energy (kJ)	0	↑	↑	0	↑↑
Energy density (kJ/g)	0	↑	0	↓	↑↑
Partial correlation with					
Protein (E%)	↑	↓	↓	↑↑	↓↓
Carbohydrate (E%)	↑↑	↑↑	↓	↑	0
Fat (E%)	↓	↑	↑	↓↓	↑↑

Trend: ↑ positive; ↓ inverse; 0 non-significant
Correlations: ↑ < 0.2; ↑↑ > 0.2; ↓ >-0.2; ↓↓ < -0.2; 0 non-significant

The ‘Green’ factor score showed very similar trends across strata of covariates for men and women: younger, more active, higher educated and non-smokers had a higher score on the ‘Green’ factor. For men, the same applied to the ‘Sweet’ factor, while for women; the ‘Sweet-Traditional’ factor was lower in smokers, only. For men, participants with lower education and smokers had a higher traditional score. Regarding nutrient intake: Scores on the ‘Green’ factor were associated with the recommended nutrient distribution at the time of the examination (more protein and carbohydrate, less fat) while the ‘Sweet’, ‘Traditional’ and ‘Sweet-Traditional’ scores were associated with a less prudent combination of lower protein and higher fat intake.

6.3 A LONGITUDINAL STUDY OF FOOD INTAKE PATTERN AND OBESITY IN ADULT DANISH MEN AND WOMEN

6.3.1 Distribution of food intake factor scores and BMI

The distribution of food intake patterns at M-82 and changes to M-87 (described by factor scores and factor score differences) across sex and age groups is given in Table 6-6. The largest variation in mean – about half a standard deviation - between the 30 and 60 year-old men at baseline was observed for the ‘Sweet’ factor, while the ‘Traditional’ factor score mean was nearly constant across age-groups. For women, mean scores on both factors decreased with age. Independently of age, the mean changes in score between M-82 and M-87 indicated an *increase* in the intake of foods belonging to the ‘Green’ factor, a *decrease* in the intake of foods belonging to the ‘Traditional’/ ’Sweet-Traditional’ factor (men/women, respectively) and a *small increase* in the intake of ‘Sweet’ foods.

Table 6-6 Factor scores at M-82 and changes in scores between M-82 and M-87

	Factor scores M-82			**Change in factor scores M-82 to M-87**		
MEN	**Green**	**Sweet**	**Traditional**	**Green**	**Sweet**	**Traditional**
Age years	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
30	0.11 (0.55)	0.19 (0.56)	-0.02 (0.34)	+0.28 (0.17)	-0.01 (0.24)	-0.26 (0.16)
40	0.02 (0.57)	-0.04 (0.64)	0.02 (0.33)	+0.29 (0.20)	+0.02 (0.25)	-0.27 (0.19)
50	-0.03 (0.62)	-0.10 (0.63)	-0.01 (0.38)	+0.31 (0.20)	+0.07 (0.24)	-0.26 (0.18)
60	-0.10 (0.65)	-0.10 (0.66)	-0.01 (0.39)	+0.29 (0.21)	+0.06 (0.28)	-0.25 (0.18)
Total	0.00 (0.60)	-0.01 (0.63)	-0.01 (0.36)	+0.29 (0.20)	+0.04 (0.26)	-0.26 (0.18)
WOMEN	**Green**	**Sweet-Traditional**		**Green**	**Sweet-Traditional**	
Age years	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	
30	0.05 (0.47)	0.10 (0.46)		+0.23 (0.17)	-0.19 (0.21)	
40	0.01 (0.46)	-0.04 (0.49)		+0.21 (0.17)	-0.19 (0.21)	
50	0.02 (0.53)	-0.01 (0.50)		+0.24 (0.17)	-0.19 (0.23)	
60	-0.09 (0.54)	-0.06 (0.47)		+0.26 (0.17)	-0.16 (0.25)	
Total	0.00 (0.50)	0.00 (0.49)		+0.24 (0.17)	-0.18 (0.22)	

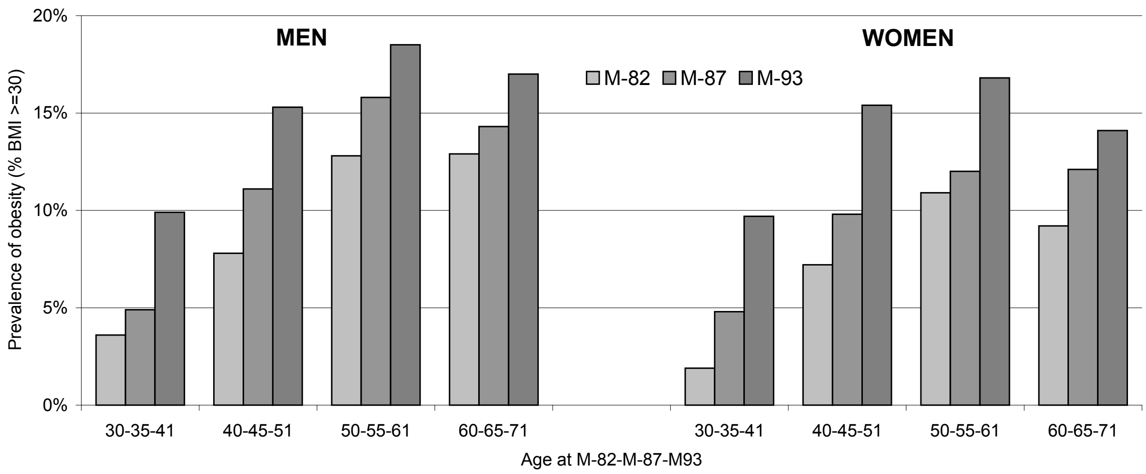
BMI in M-82 and mean changes in BMI between follow-up examinations are given by gender and age in Table 6-6. The table illustrates that height-standardised BMI increased with age and the BMI changes over time were larger for the younger participants. The prevalence of obesity showed a similar distribution (Figure 6-1) although the prevalence was not the highest among the oldest.

Table 6-7 Body Mass Index (BMI) at M-82 and differences in BMI between the three examinations

Age at M-82 Years (N)	BMI [a] kg/m^2(range) M-82	SD_{BMI}	BMI [a] difference kg/m^2 (range) M-82 to 87	M-87 to 93
MEN				
30 (304)	23.7(17; 38)	2.9	+0.9(-3; +5)	+1.0(-5; +6)
40 (333)	25.3(17; 38)	3.2	+0.5(-9; +5)	+0.7(-6; +7)
50 (298)	26.0(19; 40)	3.6	+0.3(-4; +6)	+0.4(-4; +7)
60 (224)	26.1(18; 37)	3.3	+0.0(-13; +4)	+0.2(-5; +6)
Total (1159)	25.2(17; 40)	3.4	+0.5(-13; +6)	+0.6(-6; +7)
WOMEN				
30 (310)	22.0(16; 36)	2.9	+1.0(-4; +8)	+1.1(-4; +8)
40 (305)	23.6(16; 39)	3.8	+0.9(-17; +7)	+1.2(-7; +11)
50 (274)	24.8(15; 41)	3.9	+0.5(-20; +6)	+0.9(-9; +8)
60 (206)	24.9(16; 40)	3.9	+0.3(-4; +7)	+0.1(-5; +4)
Total (1095)	23.7(15; 41)	3.8	+0.7(-20; +8)	+0.9(-9; +11)

[a] BMI's computed using height at M-82 and only weight at the respective examinations (see Paper III).

Figure 6-1 Prevalence of obesity at the three examinations by gender and age



6.3.2 Associations between factor scores and BMI

Overall, the ability of the food factor scores to predict BMI, BMI-changes or future obesity was low, both before and after adjustment for covariates. The results listed in Paper III, tables III-2 to III-5 are summarised here in Table 6-8.
For men, a significant inverse cross sectional association between the 'Sweet' factor score and BMI was observed. Inverse associations were also observed between the 'Sweet' factor score at baseline and subsequent five or 11 year BMI change, while the association with subsequent BMI change and obesity at M-93 was *positive*– when both baseline score and change in score was included (although none of the associations were significant). The inverse association between score on the 'Traditional' factor and 11 year subsequent BMI change was significant and in general, score and change in score on the 'Traditional' factor tended to be inversely associated with BMI, changes in BMI and obesity at follow-up. For both men, and women, the 'Green' factor score was not significantly associated with BMI, changes in BMI or obesity, but apart from at the cross-sectional level, the direction of the associations suggested an inverse relationship.

Table 6-8 Summary of the fully adjusted associations between food factor scores and BMI or obesity

		Men			Women	
Tables[a]	Associations [b] between:	Green	Sweet	Traditional	Green	Sweet-Traditional
2	Diet [c] and BMI, both at baseline (82)	(↑)	**↓**	(↓)	(↑)	**↓**
3	Diet (82) and changes in BMI from 82 to 87 (5-year follow-up)	(↓)	(↓)	(↓)	(↓)	**↓**
3	Diet (82) and changes in BMI from 82 to 93 (11-year follow-up)	(↓)	(↓)	**↓**	(↓)	(↓)
	In the same longitudinal model:					
4	Diet (82) and changes in BMI from 87 to 93 and	(↓)	(↑)	(↓)	(↓)	(↑)
	Changes in diet (82 to 87) and changes in BMI (87 to 93)	*(↓)*	*(↑)*	*(↓)*	*(↓)*	*(↓)*
	In the same longitudinal model:					
5	Diet (82) and BMI (93) ≥ 30 (obesity) and	(↓)	(↑)	(↓)	(↓)	(↓)
	Changes in diet (82-87) and BMI (93) ≥ 30 (obesity)	*(↓)*	*(↑)*	*(↓)*	*(↓)*	*(↑)*

[a] Table number refers to the tables of Paper III
[b] Controlled for the covariates listed in the relevant tables
[c] Diet quantified as food intake factor scores ('Green', 'Sweet', 'Traditional' and 'Green', 'Sweet-Traditional') one column for each
↑ Positive or **↓** inverse associations, (↑) or *(↑)* non-significant associations

For women, score on the 'Sweet-Traditional' factor was inversely associated with baseline BMI as well as five-year (and 11 year) subsequent BMI change, whereas a more inconsistent and insignificant pattern of association was observed when change in factor score was included in the model.
See Paper III for detailed information on the associations and the model elements.

6.4 GENDER DIFFERENCES

All analyses in the study were done separately for men and women because a gender difference in food intake pattern as well as in the development of obesity was expected. To sum up the differences observed:

- A few of the reviewed studies identified slightly different factors but only one factor (the 'drinker' factor seemed to be oppositely associated for men (positively) and women (inversely) with BMI.
- Three factors for men and two for women were identified (based on the scree plots) but the primary factor 'Green' was the same, and the 'Sweet'-Traditional' factor in women was closely related to the two namesake factors in men.
- There was a larger difference between male 'completers' and the men who dropped out, when the cross-sectional associations between factor scores and BMI were compared in the two groups, whereas the association between covariates/ factor scores and tendency to dropout was similar for the two genders.
- Correlations between factor scores based on the two different dietary assessment methods were similar for the 'Green' factor and the 'Sweet'/'Sweet-Traditional' factors but lower for the 'Traditional' factor in men.
- Factor scores were distributed equally across strata of covariates and showed similar changes over time. Women had a lower mean BMI but gained a little more (also when standardized height was used).
- In the adjusted analyses, the 'Green' factor score (or change in score) could not predict BMI or changes in BMI regardless of gender, but the direction of the association was the same in all analyses.
- The 'Sweet', 'Traditional' and 'Sweet-Traditional' factor scores were all inversely associated with BMI or change in BMI, when changes in scores were not included. In the longitudinal analyses the associations with subsequent BMI-change or obesity were positive for the 'Sweet' factor score but remained inverse for the 'Traditional' factor in men. For women, the positive association between the 'Sweet-Traditional' score and BMI-change was combined with an inverse association between change in score on the factor and BMI-change and this set of associations was reversed when obesity was used as endpoint. However, most of the adjusted associations were not significant.

7. Discussion

Is there an association between pattern of food intake and the development of overweight and obesity? This question has been the focus of the present project; but although there is a strong 'common belief' on the importance of diet and food intake patterns in the aetiology of obesity, the relation was not evident from our longitudinal analyses. Further, the few previous studies, in which the association between food intake patterns and obesity have been evaluated, were cross-sectional and their results were inconsistent.

7.1 Previous studies

The studies by Gex-Fabry et al., 1988 and Pryer et al., 2001, found inverse associations between BMI and a 'low-culinary-complexity' factor (dominated by confectionery, butter and cookies) or a 'mixed/sweet' diet cluster, respectively, and thus were in concurrence with our finding of an inverse cross-sectional association between the 'Sweet' or 'Sweet-Traditional' factor score and BMI (at baseline). On the other hand, Wirfalt & Jeffery, 1997 and Slattery et al., 1998 observed that the men in the 'soft-drink' cluster or the men loading high on the 'fruit-juice' (and sweet drinks-) factor, respectively, had a higher BMI. Likewise, McCrory et al., 1999 found that the 'variety (index) of sweets, snacks, condiments, entrées and carbohydrates' was positively associated with body fat percentage.

The inverse cross-sectional association between the 'Sweet-Traditional' factor score and BMI for women (and insignificantly between the 'Traditional' factor and BMI for men) at baseline was also confirmed by a similar association with another factor in Gex-Fabry et al., 1988 (the satiating factor dominated by macaroni, sausage and white bread), but the meat and potato dominated clusters found in Pryer et al., 2001 and Tucker et al., 1992 and the Western factor in the study by Slattery et al., 1998, Sanchez-Villegas et al., 2003 (and for men in Hu et al., 2000) were *positively* associated with BMI.

Finally, in the cross-sectional analyses the 'Green' factor tended to be positively associated with BMI for men and women in our study and this was also observed by Randall et al., 1992 and Fraser et al., 2000, who found that a 'fruit' factor and a 'fruit and vegetables' cluster, respectively, was positively associated with BMI. Contrary to that, McCrory et al., 1999 and Greenwood et al., 2000 found that a 'high vegetable variety' or being in a 'vegetarian' cluster, respectively, was *inversely* associated with BMI for women, and Hu et al., 2000 found the 'prudent' factor to be inversely associated with BMI for men. The inverse association was in accordance with the adjusted prospective and longitudinal analyses of this study, in which the 'Green' score and change in score tended to be associated inversely with subsequent BMI change. However, the associations were barely different from 0 and not significant.
Hence, overall, results from this and other literature have been mixed.

As summarized in the results chapter, gender played a role in the association between food intake pattern and obesity. A major difference was that for men, the initial analyses of possible factor solutions suggested three factors, whereas, for women, the second and third factor – equivalent to the 'Traditional' and 'Sweet' factors in men - shared high loadings for a number of foods (also after rotation) and therefore were reduced to one factor. The appropriateness of the number of factors was confirmed by the reproducibility analyses in Paper II (e.g. the scree plot). The different factors and the weak estimates made it difficult to identify clear gender differences in the association with obesity.

7.2 LIMITATIONS OF THE DESIGN

Given the inconsistency of the results and the weak associations observed, it is possible that bias and/or confounding have skewed or diluted the expected relations.
The number of participants may have been on the low side considering the number of variables studied. An 'inverse power calculation' was conducted (using nomogram reading, p. 456 in Altman, 1991) based on two equally sized groups of 600 subjects for each gender (e.g. low vs. high score on one of the diet factors), a chosen power of 80 percent and a five percent significance level. It showed that, with a standard deviation of 1.7 (computed for the entire population), a difference in mean BMI change from M-87 to M-93 larger than 0.3 kg/m^2 could be detected (equivalent to approximately one kg of weight change for a person of 70 kg/180 cm). However, this very simplified approach only provides a crude estimate of the 'power' of the study, which is complicated by the use of multivariate regression analysis.

In the longitudinal design of the analyses, using three BMI measurements and two dietary assessments, we stressed the importance of analysing the correct temporal association between diet and change in diet (and concurrent BMI change) and the subsequent change in BMI. The need for triple measurements of BMI was illustrated by the studies of Colditz et al., 1990a and French et al., 1994; Korkeila et al., 1999; Kroke et al., 2002 who all found that weight-cycling or weight loss attempts were major predictors of subsequent weight gain. However, the lag-time between the measurements in our study (up to 11 years) may have been too long. Furthermore, the exact weight history preceding the first assessment of the diet was not known. Additionally, frequent dieting makes it difficult to identify a 'habitual diet'. This may explain, in part, the differences between men and women that were observed in the longitudinal analyses. E.g. the inverse cross-sectional association between score on the 'Sweet' factor and BMI that, in women, was reversed by adjustment for previous weight change (and other covariates) could reflect un-successful dieting or underreporting of an incriminated set of foods. Because of the non-significant estimates this interpretation is speculative, but women seem to be more health-conscious and more often on a diet (31% of women vs. 17% of men at M-82 had been on one or more diets, data not shown). A variable that could have shed a light on the question was 'weight at 25'. However, we found this self-reported variable too prone to bias to use it in the analysis.

Another limitation is that the study was observational and, hence, no groups were assigned predefined dietary patterns. Further, the diet could change freely over time, and the changes in factor scores, which we used to estimate such changes, may have been less specific than what could be achieved by an (controlled) intervention approach like e.g. the secondary prevention in the 'Lyon Diet Heart Study' (de Lorgeril et al., 1994). However, conducting a randomised trial to test the food intake pattern-obesity association would be very difficult, in particular as primary prevention. Blinding of the intervention diet would not be possible, and because it would be too expensive to provide all the food for a longer period, only the dietary advice could be randomised. Low motivation and compliance with such advice in the intervention group and self-initiated diet changes in the control group could also be a major concern.

7.2.1 Selection bias and dropout

The participation rate and tendency to dropout could present a problem in terms of selection bias and a lack of variation in the diet and BMI. If hypothetically, the group of subjects, who become obese by eating an unhealthy diet, leave the study prematurely, then that association may be missed. Likewise, if the variation in the diet diminishes by selective dropout in the extremes of the various 'diet distributions' and the effect of a hypothetically adipogenic diet is weak (but dose-respondent), then the association between the diet and obesity may become non-significant. The dropout analyses showed that the 'completers' were different from those who left the study, and the associations between baseline factor scores and BMI for women who left the study were less significant but similar to the 'completer' group. This association dilution was probably due to the lower standard deviation of BMI and the factor scores in the group of 'completers'. For men, on the other hand, the associations between the factors and BMI were different in the two groups (and helps to explain that the 'Green' factor score was positively associated with concurrent BMI, but inversely with BMI change although not significantly). The difference between completers and those who dropped out could indicate a problem of selection leading to overestimation of the inverse association (or in the worst case: inversion of the 'true' association). The possible problems of selection bias due to dropout should, however, be weighed against the advantages and necessity of follow-up in a (longitudinal) study of the diet-obesity relation.

7.2.2 Confounding

A number of factors associated with dietary intake including other lifestyle factors, sex and age also influence weight gain. Two strategies were used to deal with potential confounders: restriction of the analyses by gender and multivariate regression modelling including a number of covariates such as age, physical activity, smoking, education and parity. Especially the 'Green' and 'Sweet' score means were associated with a number of covariates. However, the differences were weak and in most cases, adjustment for covariates only attenuated the regression model estimates slightly. The included confounders may not have been sufficiently accurate assessed, e.g. physical activity, and there may have been some interaction not accounted for, for instance within the factors or between foods in the factors and the covariates. It is believed, however, that the most important confounders were included, where possible.

To investigate the possible role of confounding from macronutrients, we examined the correlation between the factor scores and macronutrients as well as total energy intake but found only modest correlations. Hence, we did not adjust the factor scores for energy intake, which was likely to be underreported and only available in the sub population, who filled in the diet record. The possibility of residual confounding by factors not included in the study e.g. alcohol intake, partner status, hormone replacement therapy; type of work, work related physical activity and genetic disposition to obesity cannot be excluded. However, the associations between each of the above mentioned variables, available at M-82 and controlled for age, and BMI changes from M-82 to M-87 were not significant in linear regression analyses (data not shown).

7.3 Measurement limitations

7.3.1 Validity

As discussed in Paper I there are a number of ways to analyse food intake patterns but factor analysis is useful in the situation where the 'a priori' knowledge of the diet-obesity association is missing or inconsistent, and the likelihood of finding distinguishable food patterns is present. It could be argued that factor analysis - like cluster analysis - is too explorative of nature, and involves a high amount of subjectivity (choosing input foods, number of factors, rotation, significance level and method of scoring). A justified argument against the use of food intake patterns is that there is no true validation method or golden standard. In addition, the amount of the total variation in the food intake variables that can be explained by major factors is small; typically 15-30%, maximum 59%, in exploratory factor analysis, depending on the food/factor ratio (studies reviewed in Paper I). Nevertheless, the tests of reproducibility using different dietary assessment methods and population groups supported that the given number and structure of factors identified in this study represented actual patterns of food intake. Whereas the selection of food indicators, loading significantly on the factors only, increased the robustness of the factors and the likelihood of acceptable model fit in other data, it also lowered the amount of variation explained by the factors. The low R^2 in the models with change in BMI as outcome (Paper III), however, not only reflects the low association with the factor scores, but also with the covariates (R^2 changes only moderately when the other variables are included in the adjusted models).

The use of BMI as an indicator of body fat and the cut-point of 30 kg/m^2 as a marker for obesity is debatable, but given the material and in favour of comparability with other studies, we found it appropriate. The at-location standardised measurements of height and weight prevented underreporting of BMI and increased precision. However, BMI or BMI changes may not be as important for health as changes in body *composition* - (the relation between fat and fat-free body mass, Heitmann et al., 2000a) or fat distribution (waist or hip, Lissner et al., 2001). On the other hand, weight gain in the age group studied is most often based primarily on fat tissue enlargement. With the use of 'standardised' height when computing BMI at the different examinations, the problem of age-related 'shrinkage' in height due to osteoporosis and discus-degeneration (giving false BMI increases) was solved in a way that maintained the adjustment of weight by height, but lowered the sources of error.

7.3.2 Information bias

Studies have shown a tendency for overweight or obese to considerably underreport their fat or carbohydrate intake relative to protein intake, if compared to their estimated energy expenditure (Heitmann & Lissner, 1995; Johansson et al., 1998) and the intake of under-reporters seem to better meet the dietary guidelines (Pomerleau et al., 1999; Heitmann et al., 2000b). It was not possible to assess underreporting in this study. Thus, it cannot be excluded that some of the predominantly inverse associations between the 'Sweet', 'Traditional' or 'Sweet-Traditional' factor scores, changes in factor scores and BMI/ subsequent BMI gain, observed in our study, is a consequence of underreporting. However, controlling for baseline preceding change in BMI and baseline BMI was the strength of this study and probably ruled out the effect of BMI-related underreporting. In addition, Gibson, 1996 found an inverse association between sugary foods and BMI that persisted after adjustment for underreporting. Moreover, the FFQ used without portion size as in this study, may have the advantage of being less prone to over- or underreporting than filling in the diet record. The diet record may even selectively lead to a lower intake at the end of the week (when confronted with the measuring of the diet) and thus lower total energy intake, in the (obese) participants, who are more likely to have experience in 'counting calories' from previous diets.

7.3.3 Random errors

The precision problems with the dietary assessment methods are well known. They include low accuracy, dependence on memory, and a lack of precise validation tools and each level of data reduction or other 'manipulation' has its sources of error. Specifically for this study, the missing continuity between the lower categories in the FFQ (see material and methods chapter) could pose a problem. However, except perhaps for the missing '4-6 times a week' category, it would be overshadowed by the inaccuracy of the reported intake expected for foods eaten more seldom. For ranking subjects for the analyses of this project, it was thus acceptable that subjects ticked the category in the FFQ, which was closest to their actual intake. *Theoretically*, some of the missing food intake data could be explained by the lack of appropriate categories, and the missing continuity would represent a major source of possible error if the FFQ were to be quantified (i.e. calculating intake in e.g. g/d of each of the food groups using standard portions). However, the random errors per definition, do not *distort* the potential associations between food intake patterns and obesity, but may *attenuate* them considerably.

8. Overall Conclusions

Rq 1: *How can food intake patterns be identified and applied in epidemiological studies?*

The review of the literature showed that the most widespread methods to describe food intake patterns in observational studies are the diet index method, cluster analysis and factor analysis.

Rq 2: *What have previous investigations shown, regarding the association between food intake patterns and obesity?*

A least two such patterns can be found in most populations: a Western or traditional diet and a diet containing more vegetables and less meat and fat. Only a few studies have examined the association between food intake patterns and obesity; they were all cross-sectional and gave inconsistent results.

Rq 3-5: *Which food intake patterns can be identified in a Danish population? Can food intake patterns identified using food frequency questionnaire data, be reproduced using data from diet records? Are the food intake patterns identified in one subgroup of the Danish population reproducible in another subgroup of the same population?*

By the use of factor analysis on food frequency data from the Danish MONICA 1 cohort, we identified three food intake factors for men labelled 'Green', 'Sweet' and 'Traditional', and two for women labelled 'Green' and 'Sweet-Traditional'. The factors could be reproduced using data from a seven-day diet record, and in another sample of the same population..

Rq 6/7: *Are changes in relative weight (body mass index) or obesity development associated with a particular food intake pattern or a change in these patterns in the Danish population?*

Linear regression analyses of the association between factor scores and BMI and BMI change showed that, for men, scores on the 'Sweet' and 'Traditional' factor were inversely associated with concurrent BMI and subsequent 11-year BMI-change, respectively. In women, score on the 'Sweet-Traditional' factor was inversely associated with concurrent BMI as well as subsequent 5-year BMI-change. After adjustment for previous BMI change and updated covariates, no significant associations were found.

9. Implications and Future Research

From a public health point of view, the present study and other studies reviewed give no simple answer to the primary prevention of obesity by dietary changes. The inconsistent associations found between the food intake patterns and BMI, and especially changes in BMI, as well as the sizeable number of dropouts, hamper the formulation of new guidelines. In addition, guidelines should not only be based on observational studies, of which most are cross-sectional. Further, the habitual diet intake *is* difficult to assess, and the methods used to study both food intake and related patterns also need to be improved in the future studies necessary for evidence-based obesity prevention. Finally, the inter-relationship between diet and other factors, which influence weight gain such as habitual physical activity level and genetic constitution, need further clarification.

Consequently, we recommend additional longitudinal and later experimental studies.
Optimally such studies should include:

- Repeated FFQ with a list of foods long enough to cover all aspects of food intake as specifically as possible with standard portions as reference;
- Frequent measurements of BMI or even better waist and hip circumference or bio-impedance (or other) measurements of body composition;
- Frequent assessments of physical activity habits supplemented by tests of max O_2-uptake capacity or the use of an accelerometer.
- A large proportion of subjects should followed by the doubly labelled water method to estimate energy balance
- Other relevant covariates and possibly genetic markers of obesity.

With regard to the identification of food intake patterns, there are other options as described in Paper I and the section addressing choice of method (5.3.1). However, factor analysis or structural equation modelling probably offers the most flexible models, in which, a number of associations can be modelled simultaneously and parameters be set, in order to fit the assumptions of the relation between diet and obesity. A longitudinal model with multiple parallel 'growth' processes for body mass, food intake and physical activity level would probably be the ideal approach. However, due to the character of the material available for this project, this was beyond our reach. In case a fourth follow-up of the cohort was conducted, such kind of analyses may be worth pursuing.

10. Dansk Resumé

Aktuelle tal viser at prævalensen af fedme i Danmark er stigende og omkring 300 000 voksne danskere er i dag fede (svært overvægtige). Hvis denne stigning følger udviklingen i mange andre vestlige lande, vil fedmen og de fedmerelaterede sygdomme som sukkersyge (NIDDM), forhøjet blodtryk, slidgigt og nogle kræftformer fremover udgøre en overvældende belastning af sundhedssektoren. Det habituelle kostindtag er een af mange faktorer som sættes i forbindelse med fedmeudviklingen. Tidligere observationelle studier har dog ikke kunne levere et sikkert bevis for en (energiuafhængig) effekt af fordelingen af makronæringsstoffer i kosten på efterfølgende vægtændringer. Interaktioner mellem makro- og mikronæringsstoffer og nogle kostelementers appetitregulerende effekt kan være af betydning for vægtbalancen. Ved at studere fødevarerne som de bliver spist snarere end som isolerede makronæringsstoffer, øges muligheden for at tage hensyn til denne interaktion. Derfor har vi undersøgt om kombinationen af fødevarer i kosten - kostmønstre - har en indflydelse på vægtændringer og risikoen for udvikling af fedme.

Alle tidligere publicerede studier der har inddraget associationen mellem kostmønstre og body mass index (BMI) eller udvikling af fedme blev gennemgået. Det viste sig at alle 30 studier var på tværsnitsniveau og analyser med BMI var oftest sekundære til andre mål med studierne. Resultaterne pegede ikke i samme retning og sammenligningen af studierne var hæmmet af forskelle i hvordan kostmønstrene blev identificeret.

Relationen mellem kostmønster og fedmeudvikling blev derfor undersøgt i et longitudinelt design, ved hjælp af data indsamlet af Center for Sygdomsforebyggelse ved Amtssygehuset i Glostrup i perioden mellem 1982 og 1994 i forbindelse med den såkaldte DanMONICA undersøgelse. Kostmønstre blev identificeret ved faktoranalyse af data fra kost-frekvensskemaer i den gruppe af populationen som udfyldte både kostfrekvensskemaer og syvdages kostdagbøger. De tre faktorer for mænd, blev betegnet 'Grønt', 'Sødt' og 'Traditionelt', og de to for kvinder 'Grønt' og 'Sødt-Traditionelt' efter de fødevarer som indgik i faktorerne med mest vægt, og de svarer til kostmønstre som er observeret i andre befolkninger. Faktorerne kunne reproduceres ved hjælp af data fra kostdagbøgerne og genfindes i den gruppe af populationen som kun udfyldte kostfrekvensskemaet . Hver af de tre kostfaktorscores blev associeret med samtidigt BMI og efterfølgende BMI ændring over en 11-års periode, under hensyntagen til alder, uddannelse, fysisk aktivitetsniveau, rygning, paritet, (BMI ved udgangspunktet og forudgående BMI ændringer) og de andre kostfaktorscores.

Tværsnits analyserne viste en signifikant invers association mellem score på 'Sødt' og 'Sødt-Traditionelt' faktoren og BMI for henholdsvis mænd og kvinder. Prospektive analyser viste en signifikant invers association mellem score på den 'Traditionelle' faktor og efterfølgende 11-års BMI-ændring, mens der for kvinder var en invers association mellem score på 'Sødt-Traditionelt' faktoren og efterfølgende 5-års BMI-ændring. I de fuldt justerede analyser af associationen mellem faktor scorer, faktorscore ændringer og efterfølgende BMI-ændringer eller risikoen for fedme ved 11-års opfølgnings undersøgelsen blev der ikke fundet signifikante resultater.

Det må dog bemærkes at associationerne mellem kostmønster og fedme, observeret i dette studie, er beskedne og forskellige for mænd og kvinder. Kostmønster analyser bør udvikles yderligere og bør anvendes på andre data, der inkluderer longitudinelle målinger af kost og vægtændringer, for at skabe yderligere evidens for en sådan sammenhæng.

11. REFERENCES

Alberti KG & Zimmet PZ (1998) Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabetic Medicine* **15,** 539-553.

Allison DB, Fontaine KR, Manson JE, Stevens J, & VanItallie TB (1999) Annual deaths attributable to obesity in the United States. *Journal of the American Medical Association* **282,** 1530-1538.

Altman DG (1991) *Practical Statistics for Medical Research*. First edition. London: Chapman & Hall.

Astrup A, Grunwald GK, Melanson EL, Saris WH, & Hill JO (2000) The role of low-fat diets in body weight control: a meta-analysis of ad libitum dietary intervention studies. *International Journal of Obesity and Related Metabolic Disorders* **24,** 1545-1552.

Barker ME, McClean SI, Thompson KA, & Reid NG (1990) Dietary behaviours and sociocultural demographics in Northern Ireland. *British Journal of Nutrition* **64,** 319-329.

Beaudry M, Galibois I, & Chaumette P (1998) Dietary patterns of adults in Quebec and their nutritional adequacy. *Canadian Journal of Public Health* **89,** 347-351.

Beck-Nielsen H, Henriksen JE, Hermansen K, Madsen LD, Olivarius N, Mandrup-Poulsen TR, Pedersen OB, Richelsen B, & Schmitz OE (2000) *Type 2-diabetes og det metaboliske syndrom– diagnostik og behandling* , Klaringsrapport, **6**

Bray GA (1987) Overweight is risking fate. Definition, classification, prevalence, and risks. *Annals of the New York Academy of Science* **499,** 14-28.

Brown CD, Higgins M, Donato KA, Rohde FC, Garrison R, Obarzanek E, Ernst ND, & Horan M (2000) Body Mass Index and the Prevalence of Hypertension and Dyslipidemia. *Obesity Research* **8,** 605-619.

Camus JP (1966) [Gout, diabetes, hyperlipemia: a metabolic trisyndrome]. *Revue Du Rhumatisme Et Des Maladies Osteo-Articulaires* **33,** 10-14.

Carey VJ, Walters EE, Colditz GA, Solomon CG, Willett WC, Rosner BA, Speizer FE, & Manson JE (1997) Body fat distribution and risk of non-insulin-dependent diabetes mellitus in women. The Nurses' Health Study. *American Journal of Epidemiology* **145,** 614-619.

Cogswell ME, Perry GS, Schieve LA, & Dietz WH (2001) Obesity in women of childbearing age: risks, prevention, and treatment. *Prim. Care Update. Ob Gyns.* **8,** 89-105.

Colditz GA, Willett WC, Stampfer MJ, London SJ, Segal MR, & Speizer FE (1990a) Patterns of weight change and their relation to diet in a cohort of healthy women. *American Journal of Clinical Nutrition* **51,** 1100-1105.

Colditz GA, Willett WC, Stampfer MJ, Manson JE, Hennekens CH, Arky RA, & Speizer FE (1990b) Weight as a risk factor for clinical diabetes in women. *American Journal of Epidemiology* **132,** 501-513.

de Lorgeril M, Renaud S, Mamelle N, Salen P, Martin JL, Monjaud I, Guidollet J, Touboul P, & Delaye J (1994) Mediterranean alpha-linolenic acid-rich diet in secondary prevention of coronary heart disease. *Lancet* **343,** 1454-1459.

Field AE, Coakley EH, Must A, Spadano JL, Laird N, Dietz WH, Rimm E, & Colditz GA (2001) Impact of overweight on the risk of developing common chronic diseases during a 10-year period. *Archives of Internal Medicine* **161,** 1581-1586.

Fraser GE, Welch A, Luben R, Bingham SA, & Day NE (2000) The effect of age, sex, and education on food consumption of a middle-aged English cohort-EPIC in East Anglia. *Preventive Medicine* **30,** 26-34.

French SA, Jeffery RW, Forster JL, McGovern PG, Kelder SH, & Baxter JE (1994) Predictors of weight change over two years among a population of working adults: the Healthy Worker Project. *International Journal of Obesity and Related Metabolic Disorders* **18,** 145-154.

Friedenreich CM (2001) Review of anthropometric factors and breast cancer risk. *European Journal of Cancer Prevention* **10,** 15-32.

Fung T, Hu FB, Fuchs C, Giovannucci E, Hunter DJ, Stampfer MJ, Colditz GA, & Willett WC (2003) Major dietary patterns and the risk of colorectal cancer in women. *Archives of Internal Medicine* **163,** 309-314.

Fung TT, Rimm EB, Spiegelman D, Rifai N, Tofler GH, Willett WC, & Hu FB (2001) Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *American Journal of Clinical Nutrition* **73,** 61-67.

Gex-Fabry M, Raymond L, & Jeanneret O (1988) Multivariate analysis of dietary patterns in 939 Swiss adults: sociodemographic parameters and alcohol consumption profiles. *International Journal of Epidemiology* **17,** 548-555.

Gibson SA (1996) Are high-fat, high-sugar foods and diets conducive to obesity? *International Journal of Food Sciences and Nutrition* **47,** 405-415.

Gittelsohn J, Wolever TM, Harris SB, Harris GR, Hanley AJ, & Zinman B (1998) Specific patterns of food consumption and preparation are associated with diabetes and obesity in a Native Canadian community. *Journal of Nutrition* **128,** 541-547.

Goodman MT, Hankin JH, Wilkens LR, Lyu LC, McDuffie K, Liu LQ, & Kolonel LN (1997) Diet, body size, physical activity, and the risk of endometrial cancer. *Cancer Research* **57,** 5077-5085.

Greenwood DC, Cade JE, Draper A, Barrett JH, Calvert C, & Greenhalgh A (2000) Seven unique food consumption patterns identified among women in the UK Women's Cohort Study. *European Journal of Clinical Nutrition* **54,** 314-320.

Groop L & Orho-Melander M (2001) The dysmetabolic syndrome. *Journal of Internal Medicine* **250,** 105-120.

Guthrie HA & Guthrie GM (1976) Factor analysis of nutritional status data from Ten State Nutrition Surveys. *American Journal of Clinical Nutrition* **29,** 1238-1241.

Heitmann BL, Erikson H, Ellsinger BM, Mikkelsen KL, & Larsson B (2000a) Mortality associated with body fat, fat-free mass and body mass index among 60-year-old swedish men-a 22-year follow-up. The study of men born in 1913. *International Journal of Obesity and Related Metabolic Disorders* **24,** 33-37.

Heitmann BL & Lissner L (1995) Dietary underreporting by obese individuals--is it specific or non-specific? *British Medical Journal* **311,** 986-989.

Heitmann BL, Lissner L, & Osler M (2000b) Do we eat less fat, or just report so? *International Journal of Obesity and Related Metabolic Disorders* **24,** 435-442.

Heitmann BL, Richelsen B, Hansen GL, & Hølund U (2001) *[Overweight and Obesity] Befolkningens sundhed set i relation til den øgede forekomst af fedme i Danmark*, ISBN 87-90765-51-6. Denmark: Sundhedsstyrelsen. http://www.sst.dk/publ/publ1999/fedme/Default.htm

Hill JO, Melanson EL, & Wyatt HT (2000) Dietary fat intake and regulation of energy balance: implications for obesity. *Journal of Nutrition* **130,** 284S-288S.

Howell MA (1974) Factor analysis of international cancer mortality data and per capita food consumption. *British Journal of Cancer* **29,** 328-336.

Hu FB (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. *Current Opinion in Lipidology* **13,** 3-9.

Hu FB, Manson JE, Stampfer MJ, Colditz G, Liu S, Solomon CG, & Willett WC (2001) Diet, Lifestyle, and the Risk of Type 2 Diabetes Mellitus in Women. *The New England Journal of Medicine* **345,** 790-797.

Hu FB, Rimm EB, Stampfer MJ, Ascherio A, Spiegelman D, & Willett WC (2000) Prospective study of major dietary patterns and risk of coronary heart disease in men. *American Journal of Clinical Nutrition* **72,** 912-921.

Jacobson HN & Stanton JL (1986) Pattern analysis in nutrition. *Clinical Nutrition* **5,** 249-253.

Johansson L, Solvoll K, Bjorneboe GE, & Drevon CA (1998) Under- and overreporting of energy intake related to weight status and lifestyle in a nationwide sample. *American Journal of Clinical Nutrition* **68,** 266-274.

Jørgensen LM, Sørensen TIA, Schroll M, & Larsen S (1995) Influence of dietary factors on weight change assessed by multivariate graphical models. *International Journal of Obesity and Related Metabolic Disorders* **19,** 909-915.

Jørgensen T (1987) Prevalence of gallstones in a Danish population. *American Journal of Epidemiology* **126,** 912-921.

Kant AK (1996) Indexes of overall diet quality: a review. *Journal of the American Dietetic Association* **96,** 785-791.

Kennedy ET, Ohls J, Carlson S, & Fleming K (1995) The Healthy Eating Index: design and applications. *Journal of the American Dietetic Association* **95,** 1103-1108.

Kopelman PG (2000) Obesity as a medical problem. *Nature* **404,** 635-643.

Korkeila M, Rissanen A, Kaprio J, Sørensen TIA, & Koskenvuo M (1999) Weight-loss attempts and risk of major weight gain: a prospective study in Finnish adults. *American Journal of Clinical Nutrition* **70,** 965-975.

Kroke A, Bergmann M, Klipstein GK, & Boeing H (1998) Obesity, body fat distribution and body build: their relation to blood pressure and prevalence of hypertension. *International Journal of Obesity and Related Metabolic Disorders* **22,** 1062-1070.

Kroke A, Liese AD, Schulz M, Bergmann MM, Klipstein-Grobusch K, Hoffmann K, & Boeing H (2002) Recent weight changes and weight cycling as predictors of subsequent two year weight change in a middle-aged cohort. *International Journal of Obesity and Related Metabolic Disorders* **26,** 403-409.

Lissner L & Heitmann BL (1995) Dietary fat and obesity: evidence from epidemiology. *European Journal of Clinical Nutrition* **49,** 79-90.

Lissner L, Bjorkelund C, Heitmann BL, Seidell JC, & Bengtsson C (2001) Larger Hip Circumference Independently Predicts Health and Longevity in a Swedish Female Cohort. *Obesity Research* **9,** 644-646.

Loehlin JC (1998) *Latent variable models: an introduction to factor, path, and structural analysis*. 3rd. edition. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Maison P, Byrne CD, Hales CN, Day NE, & Wareham NJ (2001) Do different dimensions of the metabolic syndrome change together over time? Evidence supporting obesity as the central feature. *Diabetes Care* **24,** 1758-1763.

McCrory MA, Fuss PJ, McCallum JE, Yao M, Vinken AG, Hays NP, & Roberts SB (1999) Dietary variety within food groups: association with energy intake and body fatness in men and women. *American Journal of Clinical Nutrition* **69,** 440-447.

Mokdad AH, Bowman BA, Ford ES, Vinicor F, Marks JS, & Koplan JP (2001) The Continuing Epidemics of Obesity and Diabetes in the United States. *Journal of the American Medical Association* **286,** 1195-1200.

MONICA data centre (2000) The WHO MONICA project. *Http://Www Ktl Fi/Monica/* .

Must A, Spadano J, Coakley EH, Field AE, Colditz G, & Dietz WH (1999) The disease burden associated with overweight and obesity. *Journal of the American Medical Association* **282,** 1523-1529.

Muthen LK & Muthen BO (2001a) Mplus for Windows. Muthen & Muthen.

Muthen LK & Muthen BO (2001b) *Mplus User's Guide*. Version 2. Los Angeles, CA: Muthen & Muthen.

Mørup I (1986) Micro-CAMP® Computer Assisted Menu Planning APS.

Narbro K, Jonsson E, Larsson B, Waaler H, Wedel H, & Sjostrom L (1996) Economic consequences of sick-leave and early retirement in obese Swedish women. *International Journal of Obesity and Related Metabolic Disorders* **20,** 895-903.

Oliveria SA, Felson DT, Cirillo PA, Reed JI, & Walker AM (1999) Body weight, body mass index, and incident symptomatic osteoarthritis of the hand, hip, and knee. *Epidemiology* **10,** 161-166.

Pomerleau J, Ostbye T, & Bright-See E (1999) Potential underreporting of energy intake in the Ontario Health Survey and its relationship with nutrient and food intakes. *European Journal of Epidemiology* **15,** 553-557.

Pryer JA, Nichols R, Elliott P, Thakrar B, Brunner E, & Marmot M (2001) Dietary patterns among a national random sample of British adults. *Journal of Epidemiology and Community Health* **55,** 29-37.

Randall DE, Marshall JR, Brasure J, & Graham S (1991) Patterns in food use and compliance with NCI dietary guidelines. *Nutrition and Cancer* **15,** 141-158.

Randall E, Marshall JR, Brasure J, & Graham S (1992) Dietary patterns and colon cancer in western New York. *Nutrition and Cancer* **18,** 265-276.

Reaven GM (1988) Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes* **37,** 1595-1607.

Sanchez-Villegas A, Delgado-Rodriguez M, Martinez-Gonzalez MA, & Irala-Estevez J (2003) Gender, age, socio-demographic and lifestyle factors associated with major dietary patterns in the Spanish Project SUN (Seguimiento Universidad de Navarra). *European Journal of Clinical Nutrition* **57,** 285-292.

Shapiro JA, Williams MA, & Weiss NS (1999) Body mass index and risk of renal cell carcinoma. *Epidemiology* **10,** 188-191.

Shaw GM, Nelson V, & Moore CA (2002) Prepregnancy body mass index and risk of multiple congenital anomalies. *American Journal of Medical Genetics* **107,** 253-255.

Slattery ML, Boucher KM, Caan BJ, Potter JD, & Ma KN (1998) Eating patterns and risk of colon cancer. *American Journal of Epidemiology* **148,** 4-16.

Statens Institut for Folkesundhed (2002) Pressenotat. Forekomst af undervægt i den danske befolkning år 2000. (8/1 2002). *Http://Www Si-Folkesundhed Dk/Nyheder* .

Svendsen OL, Heitmann BL, Mikkelsen KL, Raben A, Ryttig KR, Sørensen TIA, Breum L, Tauber-Lassen E, Hansen GL, Ovesen L, Richelsen B, Søndergaard K, & Toubro S (2001) *Fedme i Danmark. En rapport fra 'Danish Task Force on Obesity' under Dansk Selskab for Adipositas Forskning (in Danish)*, Klaringsrapport i Ugeskrift for Læger., Nr. 8. Denmark. http://www.laegeforeningen.dk/lf/pics/pdf/klar8-2001.pdf

Terry P, Hu FB, Hansen H, & Wolk A (2001) Prospective Study of Major Dietary Patterns and Colorectal Cancer Risk in Women. *American Journal of Epidemiology* **154,** 1143-1149.

The Committee of Public Accounts (17-1-2002) *Tackling obesity in England*, Ninth report. London: British Parliament. http://www.publications.parliament.uk/pa/cm200102/cmselect/cmpubacc/421/42103.htm

Thompson D, Brown JB, Nichols GA, Elmer PJ, & Oster G (2001) Body mass index and future healthcare costs: a retrospective cohort study. *Obesity Research* **9,** 210-218.

Togo P, Osler M, Sørensen TIA, & Heitmann BL (2001) Food intake patterns and body mass index in observational studies. *International Journal of Obesity and Related Metabolic Disorders* **25,** 1741-1751.

Tucker KL, Dallal GE, & Rush D (1992) Dietary patterns of elderly Boston-area residents defined by cluster analysis. *Journal of the American Dietetic Association* **92,** 1487-1491.

Vass A (2002) Obesity causes 30 000 deaths a year, report says. *British Medical Journal* **324,** 192c.

Visscher TL & Seidell JC (2001) The public health impact of obesity. *Annual Review of Public Health* **22,** 355-375.

Willett WC (1998) Is dietary fat a major determinant of body fat? *American Journal of Clinical Nutrition* **67,** 556S-562S.

Willett WC (2002) Dietary fat plays a major role in obesity: no. *Obesity Reviews* **3,** 59-68.

Wirfalt AK & Jeffery RW (1997) Using cluster analysis to examine dietary patterns: nutrient intakes, gender, and weight status differ across food pattern clusters. *Journal of the American Dietetic Association* **97,** 272-279.

Wolf AM & Colditz GA (1998) Current estimates of the economic cost of obesity in the United States. *Obesity Research* **6,** 97-106.

Wolff CB & Wolff HK (1995) Maternal eating patterns and birth weight of Mexican American infants. *Nutrition and Health (Berkhamsted, Hertfordshire)* **10,** 121-134.

12. APPENDICES

1. ABBREVIATIONS AND DEFINITIONS

Diet record = A database of foods eaten over a given amount of time. In the diet record used, the daily food intake was weighed (or given standard portion weight) and noted for each of seven days, using a list more than 100 foods and drinks including space for additional foods. The weights were entered and summed by a computer program to obtain nutrient and energy estimates (see Paper II).

BMI = Body mass index = measured body weight / height^2 in kg/m^2

Factor, Food Factor, Food intake factor = the estimated variable representing the latent variable based on a set of correlated indicators e.g. of a (part of a) food intake pattern

FFQ = Food Frequency Questionnaire. The FFQ used in this study included 26 food groups.

Food or food group: The distinction between foods or food items and food groups is not well defined. The macronutrient contents can vary considerably *within* a common food group. At the same time foods from presumably very different food groups, can be alike with regard to macronutrients.

Food Intake Pattern (or 'Food pattern' in text)= the distribution (by frequency and/or amount) of foods in the habitual diet. The food intake pattern is described by a number of factors (i.e. the score computed on the factors) in the analyses for this project. The term Food Intake Pattern' was chosen to stress that the pattern is based on combinations of *foods,* the intake of which are correlated. The main alternative: 'Dietary Pattern' has been used to describe the diet more generally including by the content of macronutrients, which has not been the focus of this thesis. Other less frequently used alternatives are: 'Eating pattern', 'Eating Habits' and 'Meal patterns', which probably more reflects the process and timing of the intake of foods rather then the actual foods.

Indicators/ Items = Measurable variables (i.e. used to describe latent variables) e.g. food frequencies

Latent variable = A concept not directly measurable more like a feature or trait which can be described by a number of measurable indicators or items. As related examples, intelligence, alcoholism and depression have been modelled as a latent variable in psychology/ psychometric studies.

Loading; factor loading = the correlation between the indicator and the factor (an indicator "loads on the factor) – a high loading means that an indicator is particularly representative of the factor/latent variable

M-82, M-87, M-93 = MONICA 1 survey and follow-ups in 1987/88 an 1993/94

MONICA = MONitoring of trends and determinants in CArdiovascular disease = multi centre study initiated by the World Health Organization

Obesity = BMI 30 kg/m^2 or more,

Overweight = BMI between 25-29.99 kg/m^2

Rotation = 'A series of iterative searches for transformation matrices that will change initial factor solutions into final ones that account just as well for the original correlations but are simpler in other ways' Loehlin, 1998. Simpler in terms of a more clear clustering of items in high and low loadings on different factors.

Score; factor score = the value computed for each individual on each factor (the individual has a score on the factor). In the case of food factor scores, the score is the combination of the individuals' intake of $food_{i-n}$ and the loading of $food_{i-n}$ on the factor. The individual can be ranked by score value for each factor, implying a higher or lower agreement with the given factor. The score can be roughly approximated by a weighed sum (of loadings x frequency category) but is estimated using minimisation techniques.

Scree plot: The eigenvalues from e.g. a principal component analysis plotted on the number of components (1 to number of indicators in the analysis), see Figure 1 in the Appendix, and Paper II, Figure 1. An appropriate number of factors are found where the curve changes its slope to the asymptotic level (additional factors are unlikely to add useful information). The term scree is lend from geology and refers to the rubble at the base of the mountain and the cut-point is supposed to be where the slope is so steep that the loose stones (the less important factors) will roll off… (Loehlin, 1998 refers to Cattel (1966) The scree test for the number of factors. *Mulitivariate Behavioral Research* **1,** 245-276 for this information.

2. Translations

Appendix Table 1 Food groups on the food frequency questionnaire (FFQ) and foods from 7-day diet records merged into summary variables, sorted by factors for men as in Paper II, Table 2.

Food group variables, FFQ [a]	In Danish	Foods listed in the diet record (in Danish; English names in Paper II)
Wheat bread with whole grains or bran	Franskbrød, kerne eller klidbrød	Knækbrød; franskbrød (med kerner)
Raw vegetables	Rå grøntsager	Gulerødder; oliven (grønne); Syltede grøntsager (assorterede); rå grønt-sager (assorterede); tomater
Boiled vegetables	Kogte grøntsager	Kogte grøntsager (assorterede); majs; grønne ærter; sojabønner; grønne bønner
Rye bread with whole grains or bran	Rugbrød (fuldkorn, klid-brød)	Fuldkorns rugbrød
Fruit	Frugt	Æbler pærer; avocado; bananer; frugt (kogt); frugt (syltet); frugt (rå); frugt grød; appelsiner; rosiner; jordbær
Fish	Fisk	Torsk eller forel; torskerogn; kaviar; ål; fiskefilet; hornfisk; hellefisk; sild; makrel; fiskefars; muslinger; rødspætte; laks; rejer; blæksprutte; ørred; tunfisk
Rice	Ris	Ris
Cheese	Ost	30% ost; 40% ost; 45% ost
Milk or yoghurt	Mælk, yoghurt	Skummet mælk; let mælk; kærne mælk; kærnemælks-koldskål; kvark 5% sødmælk; kaffefløde; piskefløde; kakaomælk; frugt yoghurt; ymer; creme fraiche (18%)
Cake, biscuits or other baked goods	Kager, kiks	Kiks; småkager; Wienerbrød; pandekager
Candy or chocolate	Slik, chokolade	Chokolade; konfekt; bolsjer eller lakrids
Soft drink or ice-cream	Sodavand, is	Is; sodavand (m. sukker)
Jam/marmalade or honey	Marmelade, honing	Honning; marmelade eller syltetøj
Meat[b]	Kød	Oksekød (fedt); oksekød (magert); lam; svinekød (paneret); svinekød (fedt); svinekød (magert); svinekød (medium fedt); bacon; frikadeller; kalvekød. Lever; nyre; hjerte. Kylling; and eller hare; gås; høne; kalkun
Pâté and meat for bread	Leverpostej, kødpålæg	Oksebryst; levepostej; spegepølse; saltkød
Potatoes	Kartofler	Nye kartofler; gamle kartofler; kartoffel mos;
Sausages	Pølser, medisterpølse	Pølser, medisterpølse
White (wheat) bread	Franskbrød (alm.)	Rundstykker; krydder; franskbrød
Butter, lard and hard margarine	Smør, fedt, margarine	Smør og olivenolie; margarine; svinefedt
Eggs	Æg	Æggehvide; hele æg og røræg; æggeblomme
Vegetable fat marga-rine/spread	Plantemargarine	Majs olie; soja olie; solsikke olie; 'plante' margarine
Juice	Juice	Juice
Oatmeal porridge or hulled grain	Havregrød eller gryn	Byggryn; müesli; havregryn
Low-fat spread/margarine	Diætmargarine, minarine	Diætmargarine; minarine
Rye bread ordinary	Rugbrød (lyst el. alm. mørkt)	Rugbrød alm. mørkt; lyst rugbrød; øllebrød;
Spaghetti/pasta	Spaghetti	Spaghetti, Pasta
Omitted foods	Rå løg; sukker; sennep; jævnet sovs; opbagt sovs; sauce Bearnaise; mayonnaise; dressing; nødder og mandler; solsikkekerner; Chips; Franske kartofler; pizza; sammenkogt ret; klar suppe; tomat suppe; mayonnaise salater; stuvede grøntsager; gul budding; citronfromage; risengrød. Kaffe; te; vand; danskvand; øl (alm., let eller stærk); vin; spiritus; hedvin.	

[a] **Boldfaced** food groups were included in the final factors for men and/or women (Table 2) other foods were excluded

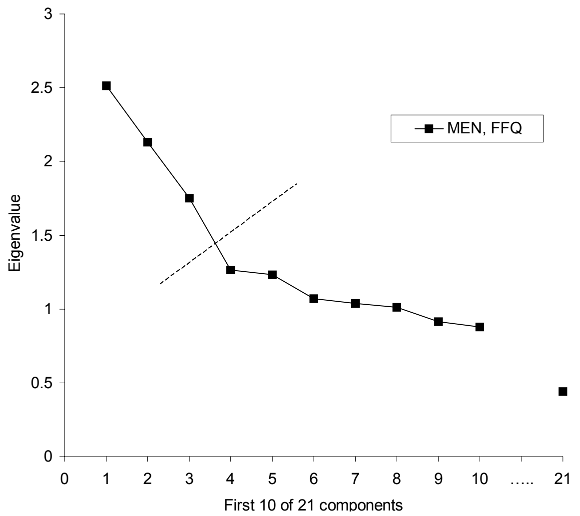
[b] Meat was not subdivided in the FFQ except for pâté, meat for bread and sausages. Poultry and organs were added to the new 'meat' variable for the diet record data.

3. Factor analysis

3.1. Exploratory factor analysis

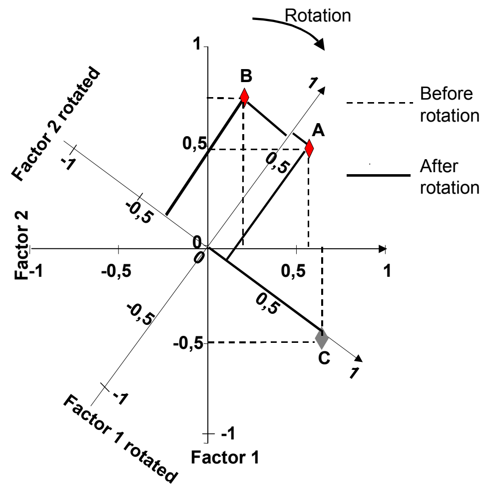
The exploratory factor analysis is based on an analysis of the covariance /correlation matrix of the food indicators under study. Typically, a principal component analysis is applied and the eigenvalues plotted on the number of components (1 to n indicators) as the scree plot (Figure 1 in the Appendix, and Paper II, Figure 1). An appropriate number of factors are found where the curve changes its slope to the asymptotic level, as illustrated by helping lines (additional factors are unlikely to add useful information). The selected factors are then rotated to simplify the factor structure and improve interpretation of the factors. In the example, (Figure 2 in the Appendix) the axes measure the loadings of variables A and B on Factors 1 (y-axis) and 2 (x-axis). Before rotation loadings were: A(0.5;0.5), B(0.75;0.2) and after rotation A(0.73;0.1), B(0.73;-0.3).

Appendix Figure 1 Scree plot (MEN, FFQ data)



See Paper II, Figure 1 for more details

Appendix Figure 2 Rotation of factors (orthogonal)



This principle is applied to all the variables at once for each possible pair of factors entered in the rotation. Various algorithms have been developed - e.g. "Varimax" for "orthogonal" (axes at 90° after rotation) and "Promax" for "oblique" rotation (factors allowed to correlate, axes ≠ 90 °), which was used in this study.

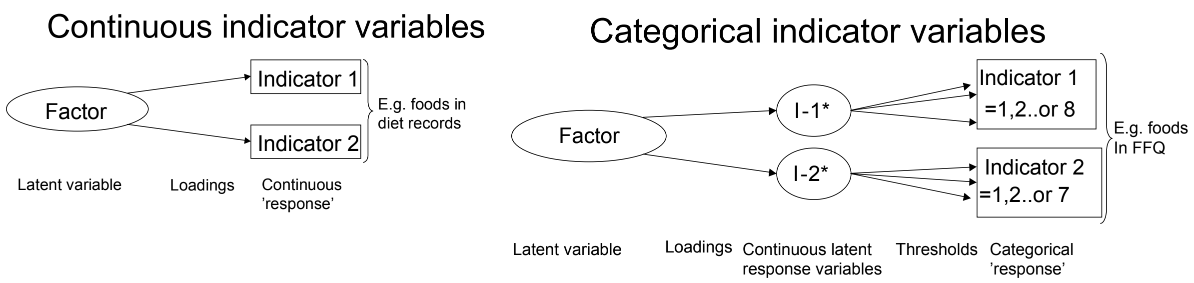
Based on each subjects intake of the foods used as indicators a score can be computed for each factor using the final set of loadings on the given factor as “weights”. A number of algorithms can be used to compute the score, in this study the scores based on the exploratory factor analysis were computed in SPSS using the regression method (‘The scores produced have mean of 0 and a variance equal to the squared multiple correlation between the estimated factor scores and the true factor values. The scores may be correlated even when factors are orthogonal’, quote from SPSS help-file). This new reduced food factor variable can be associated with the outcome in question. Exploratory factor analysis was applied on a sub-sample of the population as a pilot for the confirmatory factor analysis in the remainder of the sample (Paper II). Most studies assembled in the review (Paper I) used exploratory factor analysis only.

3.2. Confirmatory factor analysis and Structural Equations Modelling (SEM)

When SEM is applied, the model usually has two or more interrelated factors each defined by a number of indicators and optionally exogenous covariates are included. Confirmatory factor analysis can be viewed as a simple variant of the SEM family. The most important difference between exploratory factor analysis and confirmatory factor analysis/SEM is that the latter allows formal testing of hypotheses. To do that, one has to take an a priori decision on the number of factors and the indicators. With the fixation of one or more parameters (typically the indicators specifically load on one factor only i.e. their loadings on the other factors are fixed at zero), it is possible to estimate the remainder as well as measures of model fit. In addition, covariates and their direct effects (indicators regressed on covariates), as well as indirect effects (factors regressed on covariates), can be estimated/tested. This allows differential item functioning (one or more indicator (= item) perform differently in various subgroups defined by the covariates) and local dependence (two or more indicators are correlated *given* the factor) to be modelled and tested. For simplicity this was not applied in the Papers on which the thesis was based. Confirmatory factor analysis was used to compute factor scores used in the second and third paper (and compared with the scores computed using *exploratory* factor analysis/SPSS in Paper II).

Since the FFQ contains only categorical information, which was not normally distributed, a specialised program Mplus® was used (Muthen & Muthen, 2001a). The estimation of the model with categorical indicators in Mplus is based on the assumption of a continuous latent response variable interposed between the factor and the categorical variable (Figure 3 in the Appendix). The latent response variable has a number of thresholds corresponding to the number of categories in the actual variable. Depending on the 'true' latent response the subject will have a propensity to tick one of the categories in the FFQ. The latent *response* variable is not the same as the latent variable, which is approximated by the factor, as described in the Material and Methods chapter.

Appendix Figure 3 Continuous and categorical indicators in factor analysis



3.3. Estimation of factor scores in Mplus

Adapted from APPENDIX 2 and 11 in the Mplus manual Muthen & Muthen, 2001b

y is the observed (FFQ) variables
x is an independent background variable (not modelled in this study)
T is a threshold vector of length equal to the number of thresholds in the model, where a categorical *y* variable with *C* categories contributes *C* - 1 thresholds

A categorical y-variable y_j (j = 1,2,3,…p) with *C* ordered categories is defined as

$$y_{ij} = \mathrm{c},\ if\quad \mathrm{T}_{j,c} < y_{ij}{*} \leq \mathrm{T}_{j,c+1}$$

for categories c = 0,1,2,…,C-1 and $\mathrm{T}_0 = -\infty,\ \ \mathrm{T}_c = \infty$

The continuous latent variable structural equation model used in the Mplus framework is expressed in two parts: a measurement part and a structural part (the structural part = the latent variables regressed on each other and the *q*-dimensional vector *x* of independent variables is not used for the current analyses, and therefore not considered further).

The measurement part of the model is defined in terms of the *p*-dimensional latent response variable vector *y**,

$$y_i^* = v + \Lambda\eta_i + \varepsilon_i \tag{175}$$

where η is an *m*-dimensional vector of latent variables (constructs or factors), ε is a p-dimensional vector of residual or measurement errors which is uncorrelated with other variables, *v* is a p-dimensional parameter vector of measurement intercepts, Λ is a *p* x *m* parameter matrix of measurement slopes or factor loadings. The covariance matrix of ε is denoted Θ.

The parameter arrays *v*, Λ, and Θ correspond to the programs 'BY' statements.

The Mplus framework assumes conditional normality for *y** given *x*, so that it suffices to consider the conditional expectation and conditional variance.

Consider Bayesian estimation of η_i given y_i and x_i. The posterior distribution of η_i is

$$g(\eta_i \mid y_i, x_i) \propto \o(\eta_i \mid x_i) f(y_i \mid \eta_i, x_i), \tag{179}$$

where the prior $\O(\eta_i | x_i)$ is multivariate normal with mean vector μ_i and covariance matrix Σ.

In the case of binary (or categorical) *y's* (indicators), conditional independence is assumed

$$f(y_i \mid \eta_i, x_i) = \Pi\prod_{j=1}^{p} f_j(y_{ij} \mid \eta_i, x_i) \tag{181}$$

In the context of factor score estimation with binary *y* variables, θ is therefore assumed to be diagonal.

The restriction on the residual variances of θ is a natural consequence of the latent response variables of *y** being measured by binary or categorical *y* variables, while the assumption of zero residual covariances is made to simplify the factor score estimation. With binary *y* variables

$$f_j(y_{ij} = 1 \mid \eta_i) = \Phi[(-\mathrm{T}_j + \lambda'_j\,\eta_i)\theta_{jj}^{-1/2}], \tag{183}$$

where $\lambda_j^{'}$ is the jth row of Λ and $\theta_{jj}^{1/2}$ is the *j*th diagonal element of Θ.

The factor score estimate $\hat{\eta}_i$ is obtained from the mode of the posterior (179) of η_i by minimizing the function *F* with respect to η_i,

$$F = (\eta_i - \mu_i)'\Sigma^{-1}(\eta_i - \mu_i) - \sum_{j=1}^{p} \ln f_j(y_{ij} \mid \eta_i). \tag{184}$$

To find model estimates, iterative techniques from numerical analysis are used to optimize the fitting function corresponding to a particular estimator. In most cases, a quasi-Newton technique is used where only first-order derivatives of *F* are needed. Following a few initial steps using a gradient method, the method uses an approximation to the second-order derivative matrix built up during iterations. The default number of iterations is 1000 and the default convergence criterion is that the absolute value of each first-order derivative has to be less than 0.00005.

4. MODEL SPECIFICATION FOR THE COMPUTATION OF FACTOR SCORES

[Remarks in italics]
Mplus VERSION 2.01 MUTHEN & MUTHEN

```
TITLE: confirmatory factor analysis with 1 follow-up MEN
DATA:
        FILE IS c:\dokumenter\spss-sav\alligned variables\ffq.dat;
        FORMAT IS 83f9.2;
VARIABLE:
        NAMES ARE
            subject gender
            82aeg 82fisk 82fransa 82fransk 82frugt 82gronko 82gronra 82gryn 82juice
            82kageki 82kartof 82koed 82leverp 82maelkp 82marmel 82minari 82ost
            82plntma 82poelse 82ris 82rugbra 82rugbrk 82slikch 82smorfe 82sodavi
            82spaghe
            87aeg 87fisk 87fransa 87fransk 87frugt 87gronko 87gronra 87gryn 87juice
            87kageki 87kartof 87koed 87leverp 87maelkp 87marmel 87minari 87ost
            87plntma 87poelse 87ris 87rugbra 87rugbrk 87slikch 87smorfe 87sodavi
            87spaghe
            93aeg 93.....……… …….;
                                          [.etc other variables]
       USEVARIABLES ARE
       [M-82 FFQ-variables]
            82fransk 82rugbrk 82frugt 82gronko 82gronra 82ris
            82fransa 82kartof 82koed 82leverp
            82kageki 82slikch 82sodavi 82ost
            82maelkp 82fisk 82poelse 82aeg 82smorfe 82marmel
       [M-87 FFQ-variables]
            87fransk 87rugbrk 87frugt 87gronko 87gronra 87ris
            87fransa 87kartof 87koed 87leverp
            87kageki 87slikch 87sodavi 87ost
            87maelkp 87fisk 87poelse 87aeg 87smorfe 87marmel;

[Setting categorical variable option, deleting participants with missing data and selecting men,
only:]
        CATEGORICAL ARE 82fransk-87marmel;
        MISSING = BLANK;
        USEOBS ARE (gender EQ 1);
       [Identification for score computation]
        IDVARIABLE = subject;

ANALYSIS:
       [Allowing difference in mean score over time:]
        TYPE IS MEANSTRUCTURE;

       [Defining the model:]
MODEL:
[Defining the three factors by FFQ foods in M-82, number marks the loading parameter]

            82green BY     82fransk     (1)
                           82rugbrk     (2)
                           82frugt      (3)
                           82gronko     (4)
                           82gronra     (5)
                           82ris        (6)
                           82ost        (7)
                           82maelkp     (8)
                           82fisk       (9)
                           82fransa     (10);

            82sweet BY     82kageki     (11)
                           82slikch     (12)
                           82sodavi     (13)
                           82marmel     (14);

            82tradi BY     82fransa     (15)
                           82kartof     (16)
                           82koed       (17)
                           82leverp     (18)
                           82smorfe     (19)
                           82poelse     (20)
                           82aeg        (21);
```

[Model input continued]

[Defining the three factors by FFQ foods in M-87, the same loadings as in M-87]

```
87green BY    87fransk   (1)
              87rugbrk   (2)
              87frugt    (3)
              87gronko   (4)
              87gronra   (5)
              87ris      (6)
              87ost      (7)
              87maelkp   (8)
              87fisk     (9)
              87fransa   (10);

87sweet BY    87kageki   (11)
              87slikch   (12)
              87sodavi   (13)
              87marmel   (14);

87tradi BY    87fransa   (15)
              87kartof   (16)
              87koed     (17)
              87leverp   (18)
              87smorfe   (19)
              87poelse   (20)
              87aeg      (21);
```

[Thresholds set to be equal over time ($1=threshold between first and second FFQ category)]

```
[82fransk$1    87fransk$1]    (22);
[82rugbrk$1    87rugbrk$1]    (23);
[82frugt$1     87frugt$1]     (24);
[82gronko$1    87gronko$1]    (25);
[82gronra$1    87gronra$1]    (26);
[82ris$1       87ris$1]       (27);
[82ost$1       87ost$1]       (28);
[82maelkp$1    87maelkp$1]    (29);
[82fisk$1      87fisk$1]      (30);
[82fransa$1    87fransa$1]    (31);
[82kartof$1    87kartof$1]    (32);
[82koed$1      87koed$1]      (33);
[82leverp$1    87leverp$1]    (35);
[82smorfe$1    87smorfe$1]    (36);
[82poelse$1    87poelse$1]    (37);
[82aeg$1       87aeg$1]       (38);
[82kageki$1    87kageki$1]    (39);
[82slikch$1    87slikch$1]    (40);
[82sodavi$1    87sodavi$1]    (41);
[82marmel$1    87marmel$1]    (42);
```

[Repeated for other thresholds ($2=threshold between second and third FFQ category)

```
[82fransk$2     87fransk$2]    (43);
[82rugbrk$2     87rugbrk$2]    (44);
[82frugt$2      87frugt$2]     (45);
….
```

.[etc. For all thresholds between categories of FFQ variables]

```
[82marmel$7     87marmel$7]    (168);
```

[Setting means to 0 in M-82 and free in M-87]

```
[82green-82tradi@0 87green-87tradi*];
```

[Estimating correlations between same foods over time]

```
82fransk-82marmel PWITH 87fransk-87marmel;
```

[Keeping factor-factor correlations equal over time]

```
82green WITH 82sweet (169);
82green WITH 82tradi (170);
82sweet WITH 82tradi (171);
87green WITH 87sweet (169);
87green WITH 87tradi (170);
87sweet WITH 87tradi (171);
```

```
OUTPUT: STANDARDIZED;
```

[Saving factor-scores for factors at M-82 and M_87]

```
SAVEDATA:
  FILE IS c:\dokumenter\scores\3F f-u m.dat;
  SAVE = FSCORES;
  FORMAT IS F9.2;
```

13. PUBLICATIONS, LIST OF TABLES AND FIGURES

PAPER I:

Food Intake Patterns and Body Mass Index in Observational Studies. Review.

P Togo, M Osler, TIA Sørensen, & BL Heitmann (2001) *International Journal of Obesity* **25;** 12:1741-1751

PAPER II:

Consistency of Food Intake Factors by Different Dietary Assessment Method and Population Groups

P Togo, BL Heitmann, TIA Sørensen & M Osler (2003) *British Journal of Nutrition* **90**, 667–678

PAPER III

A Longitudinal Study of Food Intake Pattern and Obesity in Adult Danish Men and Women

P Togo, M Osler, TIA Sørensen & BL Heitmann (2003) Submitted, International Journal of Obesity

Paper I-III