24-hour dietary recalls as reference calibration measurements in EPIC: from statistical theory to epidemiological application

Nadia Slimani

Promotor:
Prof. dr. W.A. van Staveren
Hoogleraar in de Voeding van de oudere mens
Wageningen Universiteit

Co-promotor:
Dr. E. Riboli
Chief of the Unit of Nutrition and Cancer,
International Agency for Research on Cancer
Lyon, France

Promotiecommissie:
Prof. dr. ir. D. Kromhout (RIVM, Bilthoven)
Prof. dr. ir. P. van't Veer (Wageningen Universiteit)
Prof. dr. C. West (Wageningen Universiteit)
Prof. dr. DAT Southgate (Nutrition, Diet and Health Group of the Institute of Food Research, Norwich and University of Nottingham, United Kingdom)

Stellingen

1. The 24-hour diet recall is a cost-effective method for obtaining standardized dietary measurements in multi-centre (multi-ethnic) studies (*this thesis*).

2. Although measurement errors cannot be wholly prevented, the standardisation of 24-hour diet recalls ensures that their order of magnitude is comparable across EPIC countries (*this thesis*).

3. Living at the time when competition is believed to be the only engine of progress, it is far from rhetoric to think that intellectual synergism can still play a role in the growth of scientific knowledge, and this also authorizes a dose of realistic optimism for our future research. *Riboli, Am J Epidemiol, 2000.*

4. Multivariate analysis methods will predominate in the future and will result in drastic changes in the manner in which research workers think about problems and how they design their research. *Hair et al. Multivariate data analysis -with readings-, third edition, 1992*

5. The role and contribution of women to science is not much debated. Whereas the pioneers were frequently forced to choose between their private and professional life, new generations strive to combine both.

6. Considering the increasing contribution to our diet of processed and functional foods, the time has come to consider the food industry as a natural partner of nutritional scientists to estimate and interpret nutritional exposures properly.

7. The only way to judge individuals is to judge them individually. This may be the best way to combat any kind of racism and discrimination.

8. Credibility is the most difficult thing to achieve but also the most easy thing to lose.

24-hour dietary recalls as reference calibration measurements in EPIC: from statistical theory to epidemiological application

Nadia Slimani

Proefschrift
Ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
prof.dr.ir. L Speelman,
in het openbaar te verdedigen
op woensdag 4 december 2002
des namiddags om vier uur in de Aula

24-hour diet recalls as reference calibration measurements in EPIC: from statistical theory to epidemiological application

PhD-thesis Wageningen University - With references- With summary in Dutch.
Slimani N, 2002
ISBN: 90- 5808-710-7

Abstract

24-hour diet recalls as reference calibration measurements in EPIC: from statistical theory to epidemiological application

PhD Thesis. Wageningen University, the Netherlands and the International Agency for Research on Cancer, Unit of Nutrition and Cancer, Lyon, France
Nadia Slimani

Large multi-centre cohort studies have been set up with the aim of increasing the statistical power to detect an association between diet and disease by including study populations varying both in dietary exposures and outcome diseases. However, such studies raise also new statistical and methodological challenges to improve the comparability of dietary measurements collected from populations from different geographical and socio-cultural origins. In the European Prospective Investigation into Cancer and Nutrition (EPIC), involving half a million subjects from 23 centres in 10 European countries, a calibration approach was adopted to correct for systematic errors in individual dietary measurements obtained by means of different dietary assessment methods across centres, by re-expressing them according to the same reference scale (between-cohort calibration). 24-hour diet recall was selected as the reference calibration method in EPIC. In order to meet the theoretical requirements, such a reference calibration method must be highly standardized across study centres, provide accurate dietary estimates at the population level and should not be an obstacle to the collection of data from a representative sub-sample of the entire cohort. This thesis deals with translating the theoretical statistical concept of the calibration into an epidemiological application in the context of a large multi-centre study on diet and cancer.

The thesis presents first some theoretical concepts of the main sources of measurement errors associated with the 24-hour diet recall method and the standardisation procedures used to prevent or minimize them in the context of a large multi-centre study. An ad hoc face-to-face computerized 24-hour diet recall interview programme (EPIC-SOFT) is then described as the end product used as reference calibration method across the EPIC centres. The rationale and a new theoretical concept of standardizing food composition tables in large multi-centre studies such as EPIC is discussed separately. In a second part, the design and the populations of the calibration sub-studies set up in each EPIC centre are described, as well as the representativeness of the calibration sub-sample compared to the whole cohort.

In order to estimate the degree of standardisation of the 24-hour diet recalls across interviewers an analysis of variance was performed using mean total energy per interviewer as dependent variable. Furthermore, a validation study using urinary nitrogen as independent reference measurements was conducted to estimate the reliability of mean 24-hour diet recall measurements of protein (and total energy) for between-cohort calibration. Finally, the 36,900 24-hour recalls collected in EPIC were used to study the large heterogeneity in dietary patterns existing between the EPIC centres, using the same detailed dietary methodology.

This was the first time that calibration sub-studies using the same dietary methodology has been set up in a large multi-centre European study. Despite inherent methodological and practical problems, the calibration worked well in a large multi-centre setting, and the calibration sub-sample can reasonably be considered to be representative of the entire EPIC cohort in most centres. The standardisation across interviewers and the validation of 24-hour diet recalls for between-cohort calibration shown encouraging results. However, measurement errors, particularly conscious or unconscious behaviour of respondents and/or interviewers, cannot be prevented entirely, and the problem of measurement errors and its determinants need to be considered in the application of the calibration, as well as day of the week and seasonal coverage. This first experience should also lead to additional improvements in the reference method and setting up nested calibration sub-studies in the future. Furthermore, it demonstrates that standardized dietary methodologies can be developed in an international context and opens perspectives for using the EPIC-SOFT programme or other approaches in large European nutritional studies.

Contents

Chapter 1

General introduction

Background

In the complex field of investigation of the role of lifestyle in the etiology of chronic diseases, diet is certainly one of the individual characteristics most difficult to measure. This is particularly true when the exposure of interest is not the diet over a particularly short time period but rather long-term individual dietary intakes. After decades of intensive research in the field of nutritional epidemiology, the controversial results which have emerged so far on the relationships between diet and cancer have led to a critical consideration of traditional research methods and to the elaboration of new methodological concepts and study designs for investigating the association between diet and disease outcomes (1,2). Until the end of the 1980s, the dominant study design in diet and cancer research was the case-control study involving patients with clinically diagnosed cancers (cases) and cancer-free subjects representative of the same underlying population from which the cases originated (controls). These studies, however, present potential drawbacks. The cases and controls may, for example, report their past diet differently, particularly if the disease has lead to conscious or unconscious changes in diet, introducing systematic differential bias in dietary measurements between cases and controls and distorting the analysis of association between diet and cancer. Further, the limited variations in dietary exposure in studies conducted on population groups with relatively homogeneous diet is an additional concern and could account also for the modest association reported so far between diet and cancer (3).

During the 80s and 90s several large cohort studies were initiated with the aim of overcoming the principal limitations associated with case-control studies. In cohort studies, subjects free of the disease of interest (i.e. cancer) are recruited then followed up over a long time period during which new incident cancer cases are recorded. Furthermore, two of these projects, the European Prospective Investigation into Cancer and Nutrition (EPIC) (4) and the Multiethnic cohort in Hawaii and Los Angeles (5), were designed in order to increase the heterogeneity of both the exposure (diet) and the disease incidence (e.g. cancer), with the ultimate goal of increasing the statistical power to detect an association between diet and diseases, if it actually exists.

In parallel to the improvement of study designs and of dietary assessment methods to estimate usual individual dietary intakes during the 70s-80s, greater attention was paid to the methodology for identification, quantification and correction of errors associated with dietary measurements.

A new generation of sophisticated validation/calibration study designs using several dietary and/or biological measurements was proposed to estimate the reliability and improve the comparability of dietary exposures. Before the 1980s, the "relative" validity of dietary measurements, designed to assess the magnitude of dietary assessment errors, was essentially estimated using, as the reference measurement, another dietary method assumed to be of a greater validity. During the 80s-90s, two biological markers of absolute quantitative dietary intakes, urinary nitrogen and doubly-labelled water for estimating total nitrogen (protein) (6) and energy intakes (7), respectively, emerged as the most promising independent reference measurements to be used in validation (or calibration) studies (8). Further, the use of these two biomarkers contributed to breaking the myth, which persisted until the end of 80s, of the existence of a gold standard dietary assessment method. All the validation or calibration studies conducted so far using biological markers have shown consistently that all existing dietary assessment methods contain measurement errors, although these can vary in order and magnitude (9-11). This stimulated statistical research on the nature and magnitude of dietary measurement errors and on how to model them in mathematical equations (i.e. structural equation models) (12-14).

The large multi-centre or multi-ethnic cohort studies raised further new statistical and methodological issues related to the comparability, at the food and nutrient levels, of dietary measurements collected from populations with different eating habits and geographical and ethnic origin. The inherent problem of obtaining accurate measurements of individuals' long-term intakes, associated with all dietary assessment methods (15-17), is amplified in multi-centre studies where different dietary questionnaires are used to capture the heterogeneity of dietary patterns existing across study populations. The magnitude and nature of systematic and random errors in dietary measurements may thus vary across study populations and distort the estimation and interpretation of the overall relationship between diet and disease when all cohorts are combined together.

Rosner et al. (18) initially developed the concept of calibration, and others (19-21) then extended it to multi-centre study designs in order to achieve two main objectives 1) at the population level to adjust for systematic over- and under-estimation of the true mean dietary intakes in each cohort; 2) at the individual level, to attempt to correct for attenuation bias in relative risk due to random errors in dietary measurements. In its simplest form, the calibration concept involves re-scaling the original individual dietary measurements (Q) in Q', using a linear regression model defined as $Q' = a + bT$, where a and b are, respectively, the intercept and the slope of the regression line of the true unknown measurement (T). As the true long-term individual intakes (T) cannot be measured, the regression factors, a and b, can only be determined by regressing a substitute reference to approximate T (e.g. 24-hour diet recalls or food records) (R) on the main dietary assessment questionnaire (Q).

Theoretically, the calibration concept in nutritional epidemiology relies on three major statistical requirements:

1) *The reference measurements (R) should provide an unbiased estimate of mean dietary intakes at the population level.* However, in the absence of a totally unbiased reference measurement, this statistical requirement is difficult to satisfy in practice. However, if the direction and magnitude of systematic dietary measurement errors are constant across study populations, the reference method can be used for between-cohort calibration. The questionnaires are therefore calibrated against a dietary method which only has *relative* validity but which provides comparable measurements across study populations. In order to satisfy this objective, the principal practical implication is that the reference dietary method must be highly standardized across study populations.

2) *The reference dietary measurements should be obtained in a representative sub-sample of the study population.* The practical implication is that the study should be designed in order to obtain a high participation rate from all sub-categories of the study population.

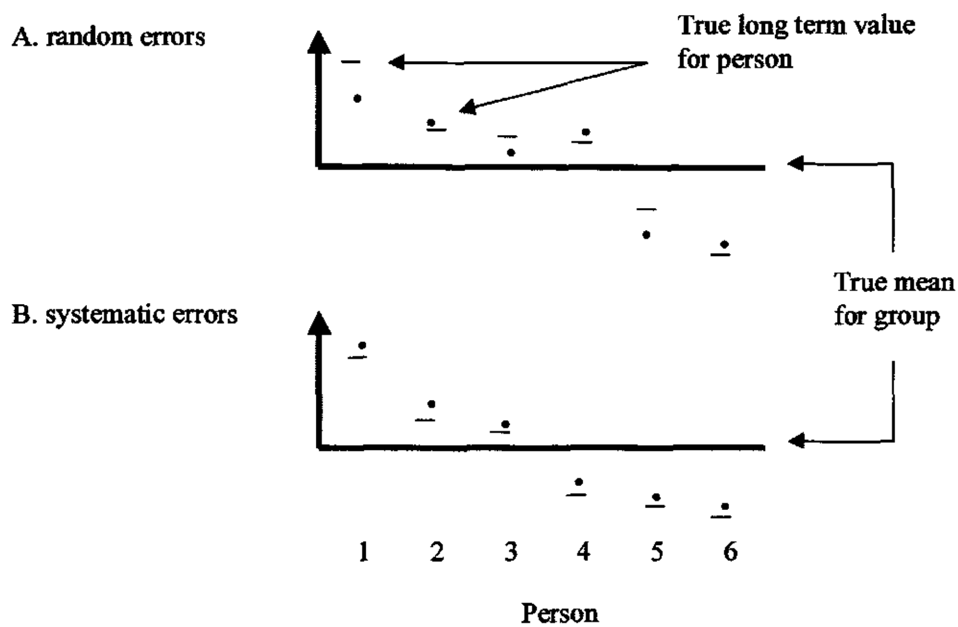
3) *There should be no correlation between the random errors of the two sets of measurements obtained with the reference method and the main dietary assessment method used in the whole cohort.* This requirement is particularly important for within-cohort calibration, but is not addressed in this thesis.

Theoretical framework for the standardization of dietary measurements

The principal aim of the *standardization* of a dietary method is to prevent or minimize measurement errors and ensure that the any systematic over- or underestimation would apply with equal magnitude to the different study populations. In practical terms, this means developing an approach which ensures that the same procedure for collecting, and consequently managing and analysing the dietary data is approximately the same everywhere. It is therefore crucial to have a good understanding of the type of errors which may be associated with dietary measurements.

When, as it is the case in the EPIC calibration study, one single dietary measurement per subject is used to estimate the mean usual intakes of a population group (i.e one 24-hour diet recall), two types of errors can theoretically be identified : *random between-person errors* and *systematic between-person errors* (12) (see figure 1).

Figure 1: Type of between-person measurement errors (adapted from reference 12)



Random errors, including both the true between-subject variation in food (or nutrient) intakes around the true mean intake of the group and the remaining sources of errors due to the dietary method used, reflect the precision (*reproducibility*) of the method for estimating long-term intakes.

This error may be due to insufficient sample size to cover a high within-/between-person variation or due to systematic within person error which are distributed randomly among subjects, but can be corrected by increasing the number of measurements to estimate correctly, the mean population intakes, in our context. *Systematic errors* (or group-level bias), which may be *additive* (or constant) or *multiplicative* (proportional) to the individual's true intake level, affect the subject non-randomly, and the mean group estimate is therefore biased. This error reflects the «*accuracy* » of the measurements.

In contrast to random errors, systematic errors can not be corrected by increasing the sample size, and they can be measured only by having a superior and independent measurement of diet in validation or calibration study design.

Several authors have proposed mathematic models to define the association between the true unknown dietary intakes called « latent variable » (T) and the dietary estimate containing the errors (R) (19, 20). As for the calibration equation model described above, the following linear equation was initially proposed to model this association: $R = \alpha_R + \beta_R T + \varepsilon_R$, where α_R = *constant bias*, β_R = *proportional bias*, ε_R = *random error* and where the *total or net error* (e_R), corresponding to the sum of all errors, is the systematic deviation (over- or under-estimation) of R compared to T ($R = T + e_R$).

Therefore, in theory, good standardisation of the reference dietary measurements containing errors, is obtained when the additional bias tends to 0, the multiplicative bias tends to 1 and the total error tends to 0. According to certain authors, the method does not need to be perfectly reliable (residual random error tends to 0) when, the method is used for estimating dietary intakes at the population group level, as is the case in EPIC. However, in practice, the situation is different. It is not possible, at the stage of the definition and standardisation of dietary measurements to identify particularly the *net measurement error*. This is frequently the result of interactions of several *sources of error* (e.g. recalled bias, food quantification, food composition tables, etc..) which may vary according to a number of parameters. Therefore, in practice, the *a priori* standardisation of dietary measurements consists, as far as possible, in clearly identifying, understanding, preventing or minimizing its different potential sources of errors, and making sure that they are applied equally across study populations. The final *net measurement error* resulting from the additional effect of these different sources of error, however, always remains unpredictable and unquantifiable at the development stage of the standardised dietary interview procedure. The nature, magnitude, direction and possible interactions of these different measurement errors can only be estimated, *a posteriori*, in a well defined validation/calibration study designs and/or using structural mathematical equation models.

Rationale of the thesis

The theoretical and statistical basis and requirements for implementing calibration sub-studies in large nutritional multi-centre studies, as summarized above, was the main topic of Kaaks' PhD thesis on "*Efficiency aspects of design and analysis of prospective cohort studies on diet, nutrition and cancer*" (22). This thesis presented methodological approaches for improving the design and analysis of prospective cohort studies, and proposed particularly the use of the calibration method in EPIC.

The present thesis logically follows that defended by Kaaks and deals with translating the theoretical concepts of the calibration into an epidemiological application. The present thesis aims to answer the following questions:

1) How can an *ad hoc* common reference method be standardized across populations with great differences in dietary patterns, socio-cultural habits and languages?
2) How can national food composition tables be standardized for international use and enable calibration at the nutrient level?
3) How can nested calibration studies be set up in practice, in the context of large multi-centre studies, and what are the degree and the determinants of its success in meeting the statistical requirements of representativeness of the calibration sub-sample compared to the whole cohort ?
4) To what extent can the common method be standardized across the interviewers involved in the field work and what is its validity as a reference measurements for between-cohort calibration?
5) For which kind of further applications, in addition to the calibration exercise, can such a standardized validated reference method be used? For example, can it be used to describe the dietary patterns existing across European geographical regions?

Since this thesis was conducted within the framework of EPIC, a short description of this study will be given. This is followed by the outline of the thesis.

European Prospective Investigation Into Cancer And Nutrition (EPIC)

EPIC is currently the largest epidemiological cohort study in the world investigating the relationship between diet, cancer and other chronic diseases such as cardio-vascular diseases. The project, initiated at the beginning of 90s, involves about half a million participants from 23 administrative centres in 10 Western European countries (Denmark, France, Germany, Greece, Italy, Spain, the Netherlands, Norway, Sweden and the UK). EPIC was set up to investigate the association between dietary and other lifestyle, endogeneous and genetic factors (23). The recruitment of study populations, representing heterogeneous groups, was mainly determined by practical and logistic considerations in order to obtain high participation and long-term follow-up from the study participants. Data on diet, anthropometry and lifestyle factors, and a blood sample were collected from each study participant. Dietary information was collected using validated country-specific dietary assessment methods to estimate individual long-term usual intakes. In addition, a second dietary method based on a single computerized 24-hour diet recall interview, collected from a sub-sample of the entire EPIC cohort, was used as a reference calibration measurement. The calibration population was defined as a random sample from each of these cohorts, weighted according to the cumulative numbers of cancer cases expected over 10 years of follow-up per sex and 5- year age strata. The sample sizes were chosen to provide calibration at both the individual and the population level, although it was recognised that the 24-hour diet recall and the main dietary assessment instrument would not have fully independent error structures.

Outline of the thesis

Chapters 2 and 3 discuss the theoretical concept of standardization and practical development of an *ad hoc* computerized 24-hour diet recall programme (EPIC-SOFT) used as reference calibration assessment method across the 10 European countries participating in EPIC. Furthermore, in chapter 2, a first insight into the level of standardisation of 24-hour diet recalls across the interviewers participating in the field work study is estimated by an analysis of variance, using total energy intakes as the dependent variable. Another important aspect is the standardisation of the estimation of nutrient intakes in the context of multi-centre studies addressed in *chapter 4*. In this chapter, the rationale and *theoretical* concept of standardizing food composition tables to be used in large multi-centre studies such as EPIC were developed. However, its practical realisation is not a purpose of the present thesis. *Chapter 5* presents the practical setting for collecting 36,900 24-hour diet recalls using the EPIC-SOFT programme from a sub-sample of each of the 23 EPIC cohort.

The rationale, study design of the calibration sub-studies set up in EPIC and a description of the population characteristics are given. Furthermore, in order to estimate the successfulness of the calibration approach in practice, the optimal logistical setting, the response rates and the representativeness of the calibration sub-sample compared to the rest of the entire EPIC cohort are discussed. *In chapter 6*, the actual validity of 24-hour diet recalls, as reference measurements for between-cohort calibration, is estimated in a validation study comparing mean dietary intakes of nitrogen (and total energy) with an independent quantitative biological marker, mean urinary nitrogen. Furthermore, in order to estimate the usefulness of the calibration in EPIC, mean intakes of nitrogen and total energy obtained from the non-calibrated EPIC questionnaires were also compared both to the 24-hour dietary recall and urinary measurements. In *chapter 7* we make use of the 36,900 24-hour diet recall measurements to study the actual heterogeneity in dietary patterns existing between the centres involved in EPIC, using a common standardized dietary methodology. Finally, the general discussion addresses the issue of the extent to which the standardized 24-hour diet recall methodology and calibration study design set up in EPIC, which is described in this thesis, satisfy the first two theoretical statistical requirements of the calibration concept, i.e. that the reference dietary measurements should be standardized and valid and collected from a representative samples of the entire EPIC cohort. The third requirement, regarding the independence of random errors between the 24-hour diet recalls and baseline dietary questionnaires, is not addressed in this thesis.

References

1. Willett WC. *Nutritional Epidemiology*, 1st ed. NewYork: Oxford University Press, 1990.
2. Kaaks R, Riboli E. The role of multi-centre cohort studies in studying the relation between diet and cancer. Cancer Lett 1997;114:263-70.
3. Wynder EL & Hebert JR. Homogeneity and nutritional exposure: an impediment in cancer epidemiology ?. J Natl Cancer Inst 1987; 79:605-7.
4. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, Stram DO, Monroe KR, Earle ME, Nagamine FS. A multiethnic cohort in Hawaii and LosAngeles: baseline characteristics. Am J Epidemiol 2000, 15: 358-70
5. Riboli E & Kaaks R: The EPIC project: Rationale and study design. *Int. J. Epidemiol.1997*, 26, Suppl. 1: S6-14.
6. Bingham SA, Cummings JH. Urine nitrogen as an independent validatory measure of dietary intake: a study of nitrogen balance in individuals consuming their normal diet. Am J Clin Nutr 1985;42:1276-89.
7. Schoeller DA. Measurement of energy expenditure in free-living human using doubly labeled water. J Nutr 1988;118:1278-89.

8. Kaaks R, Riboli E, Sinha R. Biochemical markers of dietary intakes. In: Toniolo P, Boffetta P, Shuker D, Rothman N, Hulka B, Pearce N, eds. Application of biomarkers in cancer epidemiology (IARC Sci Publ No 142). Lyon: International Agency for Research on Cancer, 1997:103-26.
9. Black AE, Bingham SA, Johansson G, Coward WA. Validation of dietary intakes of protein and energy against 24 hour urinary N and DLW energy expenditure in middle-aged women, retired men and post-obese subjects: comparisons with validation against presumed energy requirements. Eur J Clin Nutr. 1997; 51 (6):405-13.
10. Black AE, Cole TJ. Biased over- or under-reporting is characteristic of individuals whether over time or by different assessment methods. J Am Diet Assoc, 2001 101 (1):70-80.
11. Kipnis V, Midthune D, Freedman LS, Bingham S, Schatzkin A, Subar A, Carroll RJ. Empirical evidence of correlated biases in dietary assessment instruments and its implications. Am J Epidemiol, 2001, 153: 394-403
12. Willett W. An overview of issues related to the correction of non-differential exposure measurement error in epidemiologic studies. Stat Med, 1989, 8: 1031-1040
13. Plummer M, Clayton D. Measurement error in dietary assessment: an investigation using covariance structure models. Part I. Stat Med. 1993 May 30;12(10):925-35.
14. Kaaks R, Riboli E, Estève J, van Kappel AL, van Staveren WA. Estimating the accuracy of dietary questionnaire assessments: validation in terms of structural equation models. Stat Med, 1994, 13: 127-142
15. Bingham SA. The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. Nutr Abstr Rev (Series A) 1987;57,705-42.
16. Cameron ME & van Staveren WA (Eds). Manual of Methodology for food consumption studies, Oxford University Press, Oxford, 1988
17. Black AE, Goldberg GR, Jebb SA, Livingstone MB, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Eur J Clin Nutr. 1991, 45:583-99.
18. Rosner B, Willett WC. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. Am J Epidemiol. 1988, 127(2):377-86.
19. Plummer M, Clayton D, Kaaks R. Calibration in multi-centre cohort studies. Int J Epidemiol 1994;23:419-26.
20. Kaaks R, Plummer M, Riboli E, Estève J, van Staveren WA. Adjustment for bias due to errors in exposure assessments in multi-centre cohort studies on diet and cancer: a calibration approach. Am J Clin Nutr 1994;59:S245-50.
21. Kipnis V, Carroll RJ, Freedman LS, Li L. Implications of a new dietary measurement error model for estimation of relative risk: application to four calibration studies. Am J Epidemiol. 1999 Sep 15;150 (6):642-51.
22. Kaaks R. Efficiency aspects of design and analysis of prospective cohort studies on diet, nutrition and cancer. Thesis Wageningen, The Netherlands, 1994.
23. Riboli E, Hunt K, Slimani N et al. EPIC: Study populations and data collection. Publ Health Nutr (in press).

Chapter 2

Standardization of the 24-hour diet recall calibration method used In the European Prospective Investigation into Cancer and Nutrition (EPIC): General concepts and preliminary results

Nadia Slimani, Pietro Ferrari, Marga Ocké, Ailsa Welch, Heiner Boeing Marti van Liere, Valeria Pala, Pilar Amiano, Areti Lagiou, Irene Mattisson, Connie Stripp, Dagrun Engeset, Ruth Charrondière, Marilyn Buzzard, Wija van Staveren and Elio Riboli

Eur J Clin Nutr 2000; 54: 900-917

Abstract

Despite increasing interest in the concept of calibration in dietary surveys, there is still little experience in the use and standardisation of a common reference dietary method especially in international studies. In this paper, we present the general theoretical framework and the approaches developed to standardize the computer-assisted 24-hour diet recall method (EPIC-SOFT) used to collect about 37,000 24-hour dietary recall measurements (24-HDR) from the 10 countries participating in the European Prospective Investigation into Cancer and Nutrition (EPIC). An analysis of variance was performed to examine the level of standardization of EPIC-SOFT across the 90 interviewers involved in the study. The analysis used a random effects model in which mean energy intake per interviewer was used as the dependent variable, while age, BMI, energy requirement, week day, season, special diet, special day, physical activity and the EPIC-SOFT version were used as independent variables. The analysis was performed separately for men and women. The results show no statistical difference between interviewers in all countries for men and five out of eight countries for women, after adjustment for physical activity and the EPIC-SOFT programme version used, and the exclusion of one interviewer in Germany (for men), and one in Denmark (for women). These results showed an interviewer effect in certain countries and a significant difference between gender suggesting an underlying respondent's effect due to higher under-reporting among women, consistently observed in EPIC. However, the actual difference between interviewer and country mean energy intakes is about 10%. Furthermore, no statistical differences in mean energy intakes were observed across centres from the same country, except in Italy and Germany for men, and France and Spain for women, where the populations were recruited from areas scattered throughout the countries.

Despite these encouraging results and the efforts to standardize the 24-HDR interview method, conscious or unconscious behaviour of respondents and/or interviewer bias cannot be prevented entirely. Further evaluation of the reliability of EPIC-SOFT measurements will be conducted through validation against independent biological markers (e.g.nitrogen).

Introduction

The European Prospective Investigation into Cancer and Nutrition (EPIC) is a cohort study on diet and cancer that includes about 480,000 subjects from nine European countries (Denmark, France, Germany, Greece, Italy, Spain, The Netherlands, Sweden and United Kingdom) (1). Norway recently joined the project. Data on diet, anthropometry and lifestyle factors as well as blood samples were collected from each study subject. Information on usual dietary intakes was assessed using dietary history or food frequency questionnaires developed and validated in each of the participating countries (2-10). In order to adjust, at the group level, for systematic over- or under-estimation of the true mean food (or nutrient) intakes estimated by the various EPIC questionnaires, it was decided to perform an additional common dietary measurement in a subsample of the study subjects (11). The principal statistical requirements of this calibration approach are that the reference dietary method has measurement errors independent of those of the dietary assessment method to be calibrated, that the validity of the method at the population level is equally good across study centres, and that the reference measurements are obtained from a representative sample of each study cohort (12-14).

Although weighed or estimated food records were considered for the calibration method, a single 24-hour diet recall interview (24-HDR) was chosen for the EPIC project, and collected from a representative sample of approximately 1200 to 6000 subjects in each cohort. In the context of large epidemiological studies, the 24-HDR offers several advantages. It does not require literacy, it is possible to obtain a high response rate from subjects with different educational and socio-economic levels, it is relatively cheap, and the interview time is short (15-21). However, like any dietary assessment method, 24-HDR is not free of error, and it tends to underestimate food and nutrient intakes (22-25). In the absence of a completely unbiased dietary reference method, *a priori* standardization of the 24-HDR appeared the only means of preventing differential measurement errors among the study population groups. A software programme (EPIC-SOFT) was developed to ensure the highest possible level of standardization of the 24-HDR interviews collected in the 23 European centres involved in the EPIC project. The structure and functions of EPIC-SOFT are described in detail elsewhere (26).

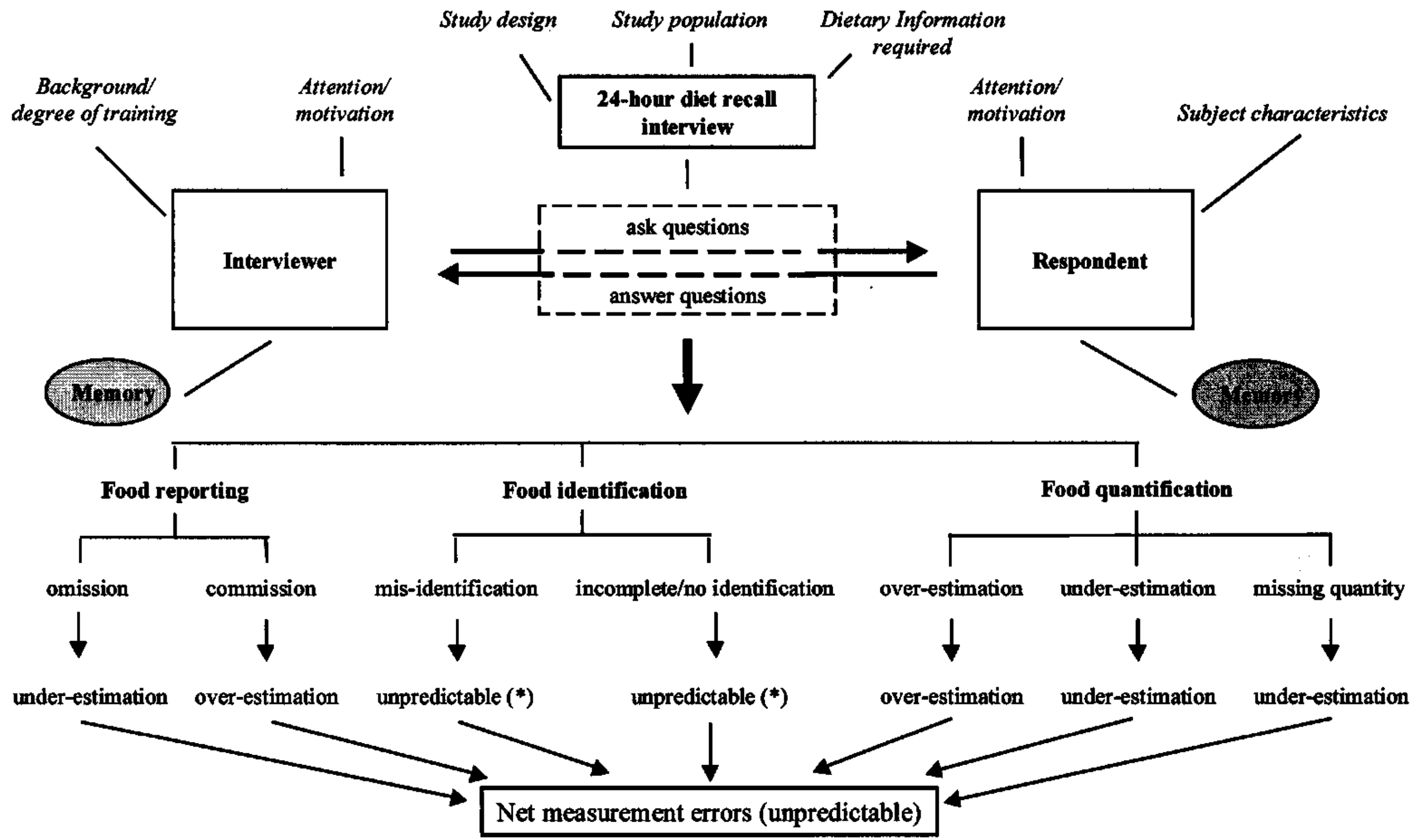
In this paper, we present the theoretical framework and approaches used to standardize 24-HDR measurements within and between the EPIC centres. To provide an insight into the level of standardization of the 24-HDR measurements in EPIC, we report and discuss an analysis of variance among interviewers, using a random effects model in which mean energy intake per interviewer is the dependent variable.

Theoretical framework for the standardization

The principal aim of the *standardization* of a dietary method is to minimize (differences in) measurement errors. In practice, the *net measurement error* is the result of interactions of several *sources of error*. The work of standardising a dietary measurement should therefore be based on the correct identification and prevention of its potential *sources of error*. In **Figure 1** we summarize the potential sources of error associated with 24-HDR measurements that can be categorized in three groups.

The *interviewer* (who asks the questions), the *respondent* (who answers the questions) and the *dietary method* (24-HDR interview procedure) used as the interviewer's tool to collect and subsequently process and analyse the information obtained. The interactions within this triangular system (*respondent, interviewer, method*) can in theory affect three main components of the 24-HDR measurements: *food reporting, food identification* and *food quantification*. Depending on the type of error introduced, over- or under-estimation of the food or nutrient intakes may result. For some types of error the impact is unpredictable. Similarly, the final *net measurement error* resulting from all the effects of these different sources of error always remains unpredictable and unquantifiable at the development stage of the standardised dietary interview procedure. Among the potential sources of error associated with retrospective dietary assessment methods such as 24-HDR, the respondent's memory remains the central inherent problem (27-31). Theoretically, memory problems affect food reporting in two ways. The subject may either forget to recall foods that were actually consumed (*errors of omission*) or may report foods that were not consumed during the recalled day (*errors of commission*).

Figure 1: *Overview of potential sources of error associated with the 24-hour diet recall interview measurement*

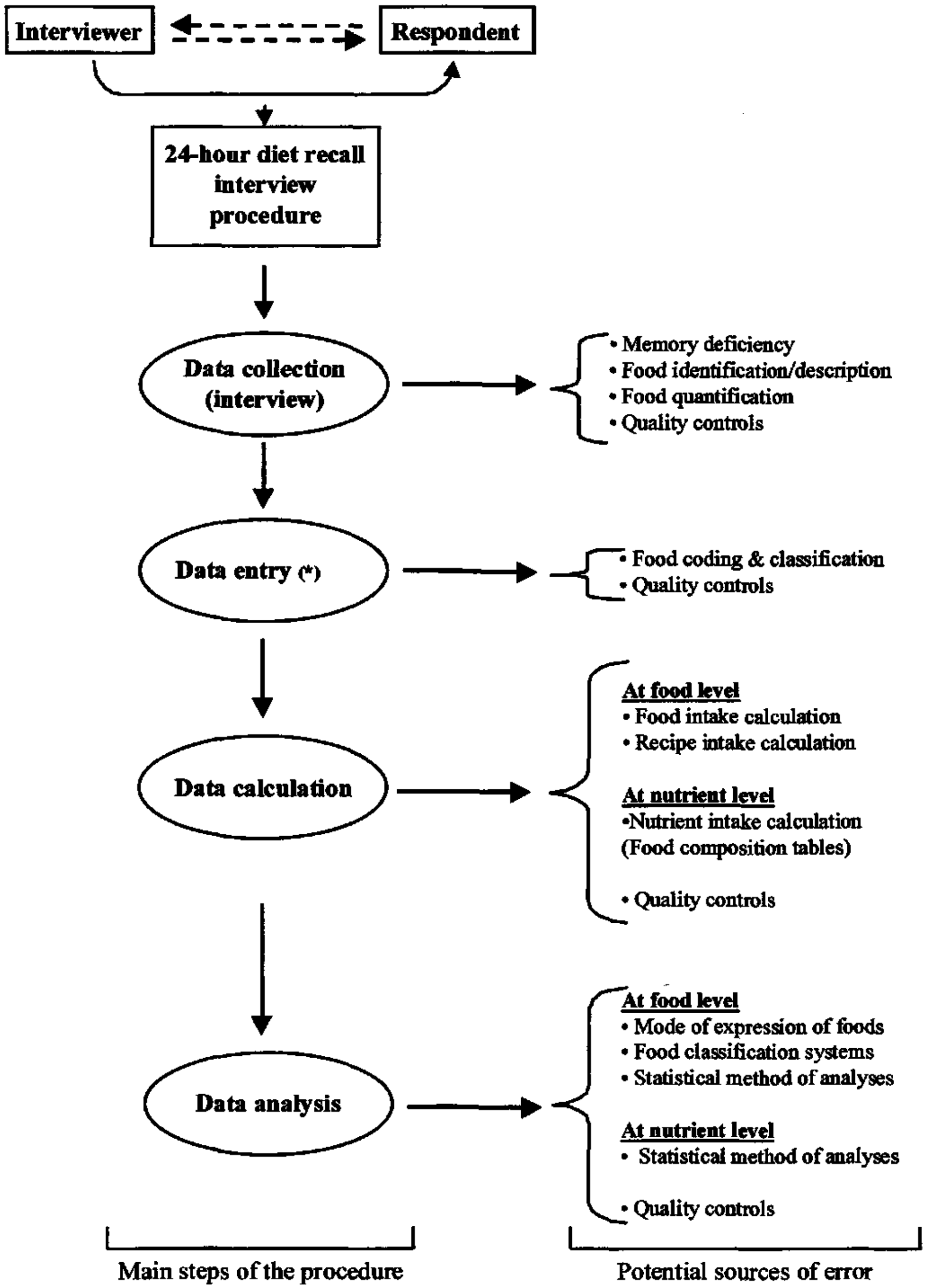


Although the *omission* of foods is frequently associated with the open-ended 24-HDR method, the reporting of foods not consumed was also observed in studies where 24-HDRs were compared with unobtrusive pre-weighed or recorded measurements collected on the same day (32-35). Additionally, various factors may interfere with the cognitive process of retrieving and recalling information. They include the characteristics of the study subjects such as gender (34, 36-39), age (34, 40-42), level of education or ethnic group (38), and the general setting of the interview (27). The perception of what is a "healthy diet" also leads to social desirability bias. Some subjects tend to over- or under-report what they consider "good" or "bad" foods (15, 31). More recent studies suggest that under-reporting of food consumption may not be nutrient- or food-neutral (43-45). Other studies show that obese people tend to under-estimate their dietary intakes systematically (46-48).

The interviewer is also a potential source of error (49) because of the open-ended nature of the 24-HDR method and the subsequent problem of the interviewer's memory. Since it is not known beforehand what the subject will report, the interviewer needs to know all the rules to identify correctly, describe, quantify and check the thousands of foods or recipes that may be reported by different subjects. These problems are particularly true for the traditional 24-HDR assessment method that used a blank sheet of paper for the interview (50). The difficulty of learning by heart and retrieving all these rules during the interview adds a supplementary memory burden for the interviewers. Various other individual behavioural factors such as question wording, incorrect use of probing questions, verbal or other reactions to the subject's answers, and the inability to establish a good relationship with the respondent (51-52) can also become sources of error, although they are difficult to document and control.

The method of collecting, managing and analysing the 24-HDR data can also introduce random and systematic errors (**Figure** 2). The approaches used in EPIC for *preventing memory deficiency*, and the standardization of *identifying and describing foods*, *quantifying portion sizes and handling recipes* in the 24-HDR interview method are described in the next section. The overall quality controls, represented at each step of figure 2, will be combined together and presented in a separate section.

Figure 2 : *Flow of the 24-hour diet recall interview procedure*



Specific approaches used in EPIC-SOFT to prevent or minimize potential sources of error

Reducing error due to memory

A series of procedures was included in EPIC-SOFT to improve the recall of foods consumed the day before, and to check the completeness and consistency of the answers. The dietary recall was structured in two main steps. In the first step (called the « quick list »), the subject is asked to remember in generic terms all the foods eaten at each meal during the previous day (e.g. meat, bread, fish, cake, cheese, salad). Only in a second step, once the quick list is completed, are additional questions asked to describe and quantify the foods in detail. The rationale for this approach, derived from previous studies (53-54), was that subjects have less difficulty remembering what they ate (e.g. bread for lunch), if they are not asked to describe and quantify each food at the same time (e.g. wholemeal, 2 slices). This approach avoids interruptions in the memory retrieval mechanism and loss of concentration due to subsequent detailed questions on the nature of the foods consumed.

To standardize the cognitive memory aids used in the first phase, the quick list is divided into eleven common food consumption occasions during the day from « before breakfast » to « after dinner, during the night ». For each food consumption occasion prompted on the screen, questions on place and time are asked as additional probes to help the study subject remember what he/she did the day before and subsequently what he/she consumed. Thereafter, the subject is systematically asked what foods were consumed. The interviewer uses a checklist adapted to the local dietary habits to ensure that no major component of the food consumption occasion is forgotten. At the end of the « quick list », a summary of the individual meal pattern reported by the subject is displayed on the screen as a check.

Within EPIC-SOFT, most of the rules to describe, quantify and check pre-entered foods and recipes are automatically prompted during the interview. At each stage of the procedure, the information entered by the interviewer is systematically checked for completeness and consistency (e.g. outlier values, empty fields, incomplete food description and quantification). In this way, errors due to memory and differences in competence, background and dietary knowledge among the interviewers are minimized (26).

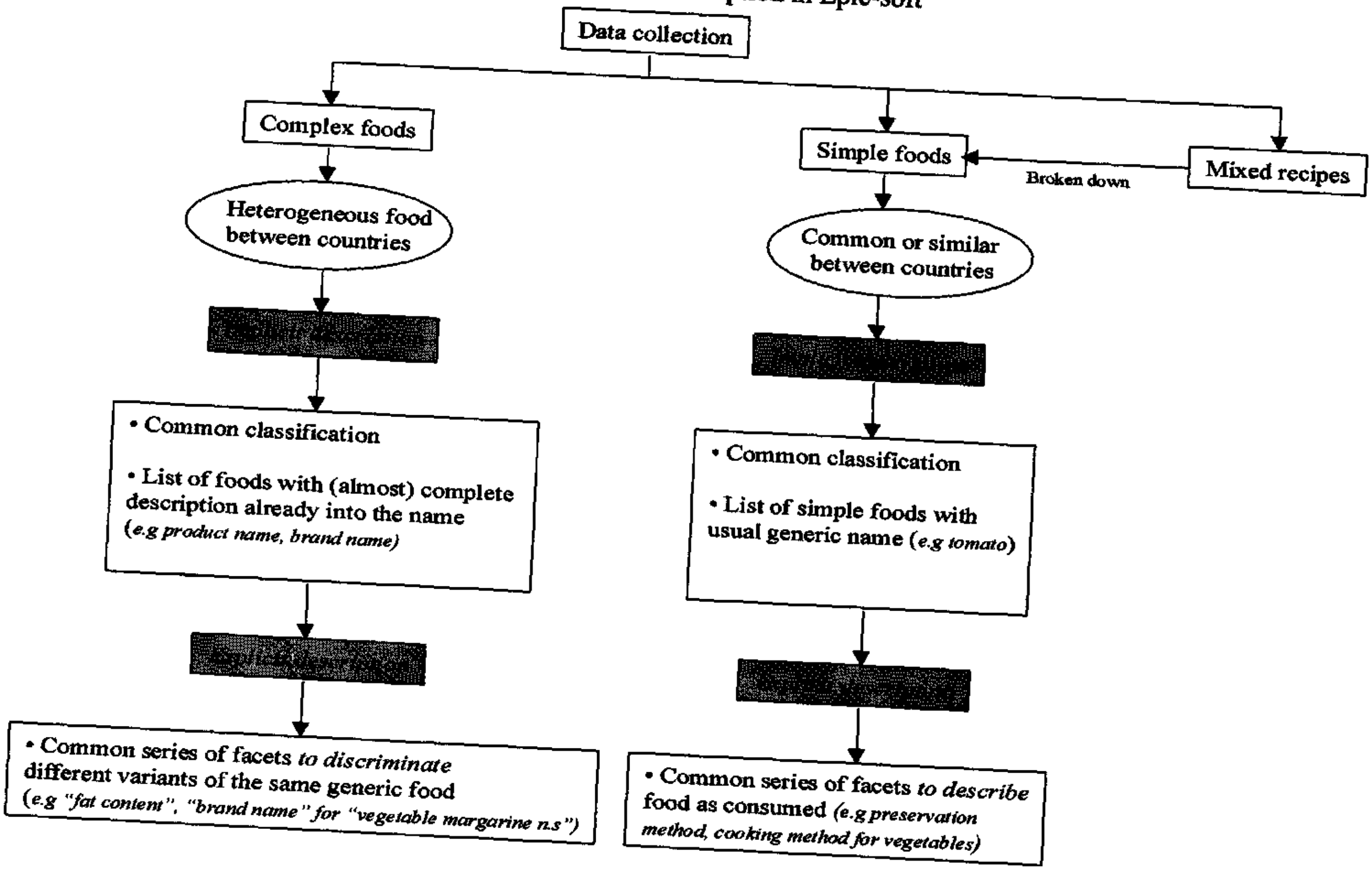
Identifying and describing foods and recipes

Open-ended methods such as 24-HDRs cause specific difficulties because the interviewee can report an extremely wide range of foods at various levels of detail. The standardised approach for identifying and describing foods developed in EPIC-SOFT was based on two underlying principles : first, to ensure that *the foods likely to be reported* were pre-entered in the software and, second, to standardize the *level of detail for describing these foods* required for the purpose of the EPIC study.

Country-specific lists of foods and recipes were available in EPIC-SOFT. The lists were based on the EPIC pilot studies and other sources of national food consumption data available. In addition, EPIC-SOFT includes the possibility to enter as "temporarily missing" any item not found in the lists. After the interview, it can be decided whether the food list should be updated with a new food or whether an existing food within the database is similar enough to select as the relevant food item. A series of "not specified" generic foods can be used when the subject is not able to provide the expected level of food description (e.g. beef rather than the specific cut of the beef). This approach minimizes arbitrary decisions by the interviewers. All the food and recipe items available in these pre-defined food or recipe lists are already classified according to common groups and sub-groups. This *ad hoc* classification (**Annex 1**) grouped foods that shared similar nutritional characteristics and that, more specifically, could be described and quantified in a similar way across regions.

In order to standardize the food identification and description methods, we developed two complementary food description systems ("*implicit description*" and "*explicit description*") which take into account the similarities and divergences of foods consumed within and between countries (**Figure 3**). The "*Implicit food description*" assumes that the food name alone is sufficient to identify and describe the food (e.g. apple, Kellogg's cornflakes). The "*Explicit food description*" means that the study subject must be asked additional questions in order to identify and describe completely how the food was actually consumed (e.g. different cooking methods used for vegetables, fish, meat). The «explicit food description» system in EPIC-SOFT was developed using the general concept of "*facets*" and "*descriptors*" from the multi-factorial Langual coding system (55) initially used to describe technological and toxicological food characteristics. The "*facets*" used in Langual to code different characteristics of the food (e.g. cooking method, preservation method) are used in EPIC-SOFT as a series of systematic questions to describe similar foods.

Figure 3: General concept of standardisation of food description in Epic-soft



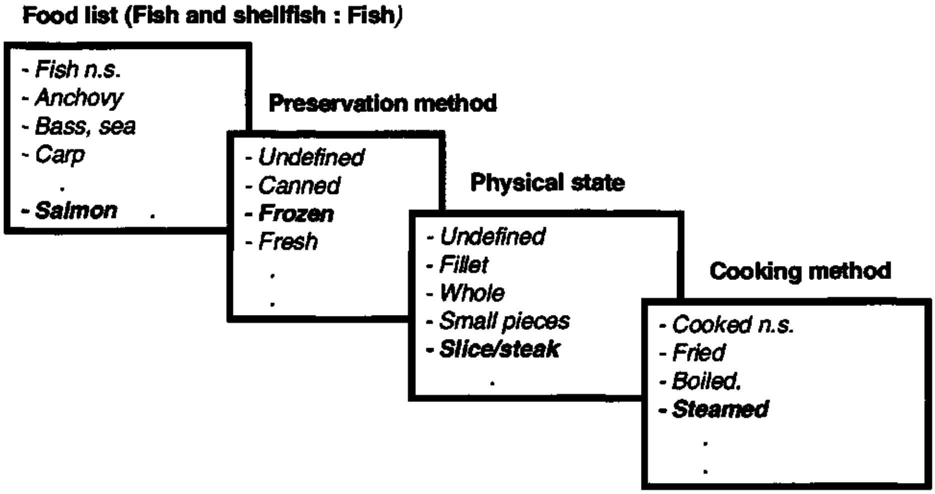


Figure 4: Food *description in EPIC-SOFT using the facet/descriptor approach (e.g "salmon")*

The "*descriptors*", which are the terms associated with each facet, are used as pre-defined potential answers (e.g. boiled, fried, steamed for the "cooking method" facet). Similar to the generic items in the food list, the descriptors always include the option « unknown » that can be used when the subject is not able to provide the expected level of detail (e.g. for yoghurt : "undefined" descriptor for the "fat content" facet). **Figure 4** shows an example to describe « salmon » using facets and descriptors. This approach ensures that in all study centres the same type of food is described by asking the same sequence of questions (facets), and that the answers given by the study subjects (descriptors) are recorded in the same way, according to common pre-defined lists of facets/descriptors.

For recipes, the descriptive system asks only for the brand name of commercial recipes. The recipe description is standardized mainly by naming and classifying the recipe correctly and obtaining detailed information on its ingredients in the same way as for foods. When the subject is not able to report individual recipes, about 150 to 350 recipes are available as standards, depending on the EPIC-SOFT version. The use of pre-defined food lists that are already coded and combined with the facet/descriptor approach in EPIC-SOFT provides a high level of food specificity that tends to minimize subjective interpretations and coding errors (56-57).

Quantifying foods and recipes
The quantification methods in the 24-HDR interviews were selected to cover both the wide variety of food types consumed by the subjects, and the large range of portion sizes, including both main dishes and recipe ingredients. Since study subjects have difficulty estimating their food portions (58-59), and because of the possibility that different errors may be introduced when different methods are used for quantifying portion sizes, food quantity estimation is standardized in EPIC-SOFT by ensuring that the same methods are offered as choices to quantify the same types of foods. This approach is expected to increase the likelihood that, if errors exist, they will apply equally for estimations of comparable foods within and between countries.

Overall, five types of quantification methods were used in a standardized way. For each specific food, the number and type of quantification methods were pre-defined according to four main parameters : the *type of food*, its *physical state* (powder or liquid), whether it is *estimated raw or cooked*, and *with or without inedible part*. Depending on the combination of these different parameters, the specific quantification method(s) to be used for each food and their corresponding comments are presented on the screen, and the interviewer can select the most appropriate option.

Photographs of household measurements
Household measurements such as glasses, spoons, bowls, etc. are used to estimate the volume of portions for beverages, sauces and foods such as flour or sugar. To facilitate the field work it was decided that it was more practical to use pictures of household measures rather than the actual utensils. Once the measure is identified, the interviewer uses a ruler to indicate the real dimensions (e.g. height and diameter of a glass) and to check it corresponds to what the subject actually used. The forms of the household measurements varried according to local habits, but the number of portions (e.g. spoons, glasses) and the fractions (1/2, 1/3, 1/4, etc.) were the same for all study centres.

Standard units
Standard units were used for foods such as fruit, vegetables, and commercial or processed products prepared and/or packaged per unit. For fruit and vegetables, in particular, it was decided to use three standard units per food (small, medium and large) to prevent systematic overestimation by high consumers and underestimation by low consumers when one single standard is available (60). During the interview, the diameter of the medium unit, measured by means of a ruler, is used as the reference to determine with the subject whether the food consumed was equal to, smaller or bigger than the

medium one. The subject may indicate a specified fraction or a multiple of the standard unit selected.

Food-specific photographs

Portions of foods or recipes that cannot be estimated by either household measurements or by the standard unit method were generally estimated by photos. For EPIC as a whole, a picture book containing sets of photos of 140 foods and recipes was produced (61). Each set of photos contains four to six pictures with portions arranged in increasing size, and with a difference between them of about 25% to allow a real visual perception of differences in size. The choice of the smallest and largest portions of each series was based on data collected during the EPIC pilot phase. In addition, fractional and multiple servings were available as standard answer options, which increases the range of selectable portion sizes without increasing the number of photos and means that food left on the plate can be checked for (e.g. the subject can indicate that only one third of the selected portion size was actually consumed). In addition, this function tends to minimize the flat-slope syndrome which can occur when subjects are not able to select portion sizes above or below the smallest or largest ones provided (62).

It was shown, for photographs, that food quantification errors can be attributable to problems of visual *perception*[1] or *conceptualization*[2] independent of memory errors (63-65). Particular attention was therefore paid to standardising the technical methods used to produce the photos using the results and the experience of two validation studies conducted during the EPIC pilot phase in France and Italy (62,63). The pictures were taken from the same angle, the same food presentation was used on the dish, the number of pieces was consistent with increasing portion sizes, and good visual representation of foods was obtained.

Specifying a known amount

When the subject knows the exact quantity of the food consumed, a weight or volume can be entered manually in a blank field available on the quantification screen. This method is used particularly to quantify the ingredients of recipes known by the subject.

---

[1] Ability to relate an amount of food which is present in reality to an amount depicted in a photograph (see Nelson *et al*, 1994)

[2] Ability to make a mental construction of an amount of food which is not present in reality, and to relate that to a photograph (see Nelson *et al*, 1994)

System calculation and treatment of unknown amounts
When a subject is unable to report a portion size, an "*unknown portion*" option can be selected during the interview, meaning that the quantity is then considered missing. However, default portions were foreseen for small quantities of foods (e.g. melted cheese on a dish) frequently unknown by the subjects. In addition, for some additions to foods (e.g. fat, sauce or sweetener) the missing quantities are replaced by standard amounts calculated automatically based on the weight of the food consumed. The amount of fat used during cooking, which is often unknown to the subject, is also systematically calculated using standard coefficients. All these «estimated standard quantities» are tracked by the system and can be replaced later on by more recent and accurate values.

In addition, different algorithms were developed to convert systematically the original form of expression of foods (as estimated during the interview) to their final common mode of consumption (as finally consumed by the subject). The conversions from volume to weight (grams), from raw to cooked weights and from foods with to without inedible parts are calculated automatically by the software using standardized food-specific conversion factors. These can, however, always be recalculated using other coefficients if necessary, by changing the original information automatically stored by the system.

Handling recipes
Management of recipes is crucial for the standardization of 24-HDR, since ingredients and methods of preparation differ widely, and since a large proportion of food items reported is derived from recipes. Traditionally, recipes are assigned to a food group on the basis of the primary ingredients. This approach has been shown to overestimate the contribution of certain food groups (e.g. meat, fish, grain products) and underestimate others (e.g. milk products, fat and oils) (66). The fat and oil group, for example, doubles its energy contribution from 5% to 12% when the mixed dishes are broken down. In the approach adopted for EPIC-SOFT, all recipes are broken down into simple ingredients, so that the problems of comparison, classification and management of recipes within and between countries are overcome because their basic ingredients (single foods) are considered rather than the recipes themselves. To standardize the classification into food groups, it was necessary to assign independently which types of products were handled systematically as recipes and which as foods, whether or not the subject knew the ingredients. For example bread, cake, biscuits, soup, and drinks are systematically seen as foods and not broken down into ingredients.

The break-down into ingredients and calculation of their separate quantities as consumed is accomplished in two steps. First, the total recipe amount as finally consumed by the subject is estimated using one of the pre-defined quantification methods in the software (e.g. photo, standard unit). In parallel, each ingredient of a standard or individual recipe reported by the subject, most frequently indicated in its raw form, is converted into its final « as consumed » form using reference adjustment factors for weight change due to cooking, edible part measured or derived from the literature (67). The « as consumed » proportion of each ingredient is then calculated using the recipe quantity estimated by the subject as the common denominator.

Quality controls

The objective of data quality controls is to prevent and identify errors during the interview or in the subsequent data processing steps, and to check that the procedure used to collect and handle the dietary data is well standardised between interviewers. Although it has been shown that well-trained nutritionists can ensure good manual data collection, coding and calculation (68), a computerized 24-HDR interview programme offers the technical possibility to implement a large series of systematic quality controls to check the information generated while the subject is still present and limit *a posteriori* arbitrary decisions on outlier values or unlikely food data. The EPIC-SOFT system checks for missing information and outlier values. Although a single 24-HDR is not enough to estimate accurately an individual's long-term intake, the energy and macro-nutrient intakes estimated by 24-HDR are compared with standard energy requirements based on the subject's age, sex, weight and height, in order to correct, with the subject, for gross or suspicious errors identified as too high or too low compared to the standards. In the current EPIC-SOFT versions, temporary and country-specific nutrient databases are used for intermediate estimates of crude nutrient intakes, and to correct for any gross possible under- or over-estimation while the study subject is still present. The pooled analyses of the calibration data at the nutrient level will be recalculated using the final standardised EPIC nutrient databases over the next two years. The overall rationale and concept for standardising these nutrient databases is discussed elsewhere (69,70). Additional software and guidelines were developed to ensure a high level of control and standardization in the maintenance of the EPIC-SOFT database versions over time, and the automatic recalculation and export of the interview data performed after the interview.

Interviewer training on how to use the EPIC-SOFT programme and its related tools requires 3-4 days. An interview is generally completed in about 30-35 minutes regardless of the country.

Evaluation of the standardization of the 24-HDR measurements

In the previous sections we presented the approaches developed in EPIC-SOFT to maximize the standardization of 24-HDR measurements. As a preliminary evaluation of this standardization, we used data from the calibration study to examine differences across interviewers in the mean energy intakes obtained. We hypothesised that after adjustment for potential confounders such as gender, age and physical activity of the subjects, our standardization of dietary measurements would result in no differences in mean energy intakes by interviewer.

Methods

For the EPIC calibration study 1200 to 6000 24-HDR interviews were collected in each of ten countries by 90 female interviewers. Energy intake was calculated in EPIC-SOFT using temporary country-specific food composition tables. As there are possible systematic differences between the national food composition tables (69), we decided to study the differences among interviewers separately for each country. At a later stage, when a standardized food composition table is available, we will extend the analysis to between-country comparisons.

Data from two countries (Greece and Norway) were not included in the analysis because their fieldwork or data processing was still ongoing. In addition, we excluded the data from 20 interviewers who performed fewer than 30 interviews in a given centre. Three interviewers in France and one in The Netherlands who conducted interviews in more than one region (11_1, 11_7 and 11_8; 51_4/52_3) have, however, been considered separately in the analysis to avoid confounding the interviewer and the centre effect. In total, energy intakes of 32,063 subjects collected by 70 interviewers were included in this statistical analysis.

A random effects model was used in which energy intake was the dependent variable and the "interviewer" variable was considered as a random effect. Under this assumption, the mean energy intake corresponding to each interviewer is considered as randomly selected from a theoretical distribution of interviewers. To achieve asymptotic normality of the dependent variable, the individual measurements of energy intake were log-transformed. The model included the following potential confounding factors as fixed effects: respondent's age, body mass index, predicted energy requirement (reference), week day and season of recalled day, whether the study subject was on a special diet during the recalled day and whether it was a special day (travel, celebration, illness). To study the centre effect, "centre" was introduced as a fixed effect in the model. The analysis was performed separately for men and women.

The hypothesis to be tested was whether the variance of the random effects equals zero, indicating that there is no significant difference between the energy intake measurements collected by the interviewers within a given country. The significance of the variable "interviewer" was tested by comparing the deviance of the model with and without the "interviewer" variable.

For countries where the "interviewer" variable was found to be significant, physical activity was subsequently added to the model when data were available, as an important additional possible confounder. Since much individual data is missing for physical activity, the values were considered at the interviewer level in the model. Two separate variables were used: occupational physical activity and leisure activities. The former was entered into the model as the percentage of subjects with manual work per interviewer. The latter (e.g. walking, cycling, gardening) were recorded as the number of hours of physical activity per week and converted into energy expenditure using the standard energy cost of the activity (71), and summed up to obtain a continuous score.

In three countries (France, Spain, The Netherlands) a preliminary version of EPIC-SOFT was used to collect the first 24-HDR interviews. A variable indicating the EPIC-SOFT version used to collect the interviews was then introduced into the model to explain possible training or methodological effects in the countries where systematic differences were observed between interviewers. In addition, in order to estimate the magnitude of the differences between interviewers, we measured the differences in percentages between the arithmetic mean of subject energy intake per interviewer and the country mean energy intake as well as the differences in mean energy intakes between centres and countries.

Results

Table 1 a & b presents the geometric means of the subjects' energy intake per interviewer, centre and country. P values, indicating the significance of difference of mean energy intake (log-transformed) estimated by interviewers are reported for comparison between interviewers (p value 1) and between centres (p value 3).

Table 1a : Test of difference of mean energy intake between interviewers and centres from the same country using a mixed model : Men

Country/centre	Interviewer codes	No.of interviews	Geometric Mean of energy intake (*)			p value (1)	p value (2)	p value (3)
			2.5th percentile	Mean	97.5th percentile			
ITALY						0.396	-	0.023
Florence	21_1	162	1478	2616	4619			
Florence	21_3	105	1312	2407	4510			
Varese	22_1	149	989	2428	5428			
Varese	22_3	42	1716	2618	3966			
Varese	22_5	117	1323	2651	4832			
Ragusa	23_1	45	979	2379	4621			
Ragusa	23_2	55	1156	2412	4765			
Ragusa	23_4	60	1229	2409	4706			
Turin	24_1	122	1299	2411	4177			
Turin	24_2	119	1014	2328	4387			
Turin	24_3	308	1314	2468	4506			
Turin	24_4	128	1243	2442	4990			
SPAIN						0.000	0.09	0.198
Oviedo	31_1	189	1242	2443	4810			
Oviedo	31_4	197	1553	2838	5398			
Granada	32_2	149	1203	2568	4461			
Granada	32_5	65	1222	2417	4154			
Murcia	33_4	244	1029	2498	4947			
Pamplona	34_1	444	1437	2594	4429			
San Sebastian	35_1	52	1772	2861	4771			
San Sebastian	35_2	143	1577	3213	5578			
San Sebastian	35_3	135	1588	2845	5114			
San Sebastian	35_4	160	1263	2781	4671			
UK						0.824	-	0.200
Cambridge	41_4	93	1124	2252	3829			
Cambridge	41_10	53	1170	2355	4216			
Cambridge	41_11	42	1434	2147	3223			
Oxford	42_1	190	1237	2333	4031			
NETHERLANDS						0.001	0.004	-
Bilthoven	51_1	233	1136	2489	5317			
Bilthoven	51_2	413	1233	2679	5554			
Bilthoven	51_3	426	1392	2744	5009			
Bilthoven	51_4	207	1227	2531	5486			

Table 1a (continued)

Country/centre	Interviewer codes	No.of interviews	Geometric Mean of energy intake (*)			p value (1)	p value (2)	p value (3)
			2.5th percentile	Mean	97.5th percentile			
GERMANY						0.000	0.000	0.045
Heidelberg	71_1	36	963	2176	5968			
Heidelberg	71_2	177	896	2104	3960			
Heidelberg	71_3	598	1210	2475	4626			
Heidelberg	71_4	35	957	2124	3471			
Heidelberg	71_5	134	1050	2197	4635			
Heidelberg	71_6	53	940	2232	3363			
Potsdam	72_2	844	1089	2352	4159			
Potsdam	72_5	202	1285	2382	4286			
Potsdam	72_6	166	1369	2397	4543			
SWEDEN						0.000	0.179 (~)	0.102
Malmo	81_1	551	969	2093	3823			
Malmo	81_2	584	1091	2217	4087			
Malmo	81_3	286	948	2270	4265			
Umea	82_1	232	1205	2429	4435			
Umea	82_2	317	1197	2341	4368			
Umea	82_3	272	1182	2241	3898			
Umea	82_4	337	1217	2456	4495			
Umea	82_5	218	1472	2496	4079			
DENMARK						0.777	-	0.288
Aarhus	91_2	564	1335	2608	4590			
Copenhagen	92_1	89	1281	2602	4583			
Copenhagen	92_2	587	1215	2546	4652			
Copenhagen	92_4	328	1241	2418	4166			
Copenhagen	92_5	352	1310	2407	4326			

(1) : p value for difference in mean total energy intake between *interviewers* nested by centre from the same country

(2) : p value for difference in mean total energy intake between interviewers needed by centre from the same country, and adjusted for physical activity

(3) : p value for difference in mean total energy intake between *centres* from the same country

(*) : Total *energy intake* adjusted for age, body mass index, energy requirement, seasons, week days, special diet, special day

(~) : Only physical activity at work has been considered in the model (information on leasure time activity was not available for that country)

Table 1b : Test of difference of mean energy intake between interviewers and centres from the same country using a mixed model : Women

Country/centre	Interviewer codes	No.of interviews	Geometric Mean of energy intake (*)			p value (1)	p value (2)	p value (3)
			2.5th percentile	Mean	97.5th percentile			
FRANCE						0.345	-	0.009
Ile-de-France	11_3	706	988	1905	3595			
Ile-de-France	11_7	495	926	1829	3410			
Nord-Pas-Calais	11_4	452	1021	1939	3246			
Alsace-Lorraine	11_1	23	1047	2055	3293			
Alsace-Lorraine	11_5	455	1071	1918	3412			
Rhone-Alpes	11_1	413	942	1909	3539			
Rhone-Alpes	11_6	605	954	1787	3140			
Lang./Roussillon	11_7	388	971	1809	3374			
Lang./Roussillon	11_8	237	925	1781	3104			
Aquitaine	11_8	443	997	1888	3323			
Bretagne/P.Loire	11_9	324	1095	1853	2964			
Bretagne/P.Loire	11_10	311	1011	1866	3302			
ITALY						0.139	-	0.166
Florence	21_1	443	820	1740	3011			
Florence	21_3	325	839	1696	3416			
Varese	22_1	419	713	1642	3036			
Varese	22_3	135	945	1751	3109			
Varese	22_5	227	776	1819	3149			
Ragusa	23_1	39	853	1702	4025			
Ragusa	23_2	44	1018	1832	3895			
Ragusa	23_4	52	733	1590	3687			
Turin	24_1	57	739	1474	2493			
Turin	24_2	74	779	1658	3527			
Turin	24_3	188	745	1691	3263			
Turin	24_4	73	619	1808	3707			
Naples	25_1	258	710	1652	3241			
Naples	25_2	145	842	1760	3298			

Table 1b (Continued)

Country/centre	Interviewer codes	No.of interviews	Geometric Mean of energy intake (*) 2.5th percentile	Mean	97.5th percentile	p value (1)	p value (2)	p value (3)
SPAIN						0.889	1.000	0.000
Oviedo	31_1	162	917	1714	3077			
Oviedo	31_4	162	893	1881	3794			
Granada	32_2	195	792	1663	3085			
Granada	32_5	105	728	1576	2982			
Murcia	33_4	304	830	1756	3590			
Pamplona	34_1	271	887	1815	3120			
San Sebastian	35_1	33	774	1789	3127			
San Sebastian	35_2	65	864	2130	3976			
San Sebastian	35_3	57	1171	2000	3634			
San Sebastian	35_4	89	1035	1939	3661			
UK						0.189	-	0.707
Cambridge	41_2	53	768.	1603	2600			
Cambridge	41_3	37	687	1680	4342			
Cambridge	41_4	109	838	1764	3095			
Cambridge	41_10	63	909	1575	3188			
Cambridge	41_11	36	688	1620	4039			
Oxford	42_1	379	895	1675	3089			
NETHERLANDS						0.000	0.000	0.358
Bilthoven	51_1	317	767	1779	3607			
Bilthoven	51_2	394	928	1902	3349			
Bilthoven	51_3	469	856	1930	3832			
Bilthoven	51_4	234	841	1834	3549			
Utrecht	52_1	1154	951	1900	3301			
Utrecht	52_2	373	801	1806	3184			
Utrecht	52_3	341	780	1705	3192			
GERMANY						0.000	0.004	0.707
Heidelberg	71_1	45	909	1859	3229			
Heidelberg	71_2	200	850	1636	3162			
Heidelberg	71_3	666	849	1835	3346			
Heidelberg	71_4	47	741	1608	2717			
Heidelberg	71_5	90	819	1592	3009			
Heidelberg	71_6	40	736	1782	3108			
Potsdam	72_2	591	796	1658	3104			
Potsdam	72_5	267	868	1777	3160			
Potsdam	72_6	196	733	1725	3110			

Table 1b (Continued)

Country/centre	Interviewer codes	No.of interviews	Geometric Mean of energy intake (*) 2.5th percentile	Mean	97.5th percentile	p value (1)	p value (2)	p value (3)
SWEDEN						0.000	0.000 (~)	0.110
Malmo	81_1	685	883	1701	3162			
Malmo	81_2	684	819	1695	3212			
Malmo	81_3	342	763	1720	3200			
Umea	82_1	299	1016	1880	3071			
Umea	82_2	345	811	1743	3096			
Umea	82_3	323	813	1683	2956			
Umea	82_4	450	778	1755	3264			
Umea	82_5	207	994	1995	3479			
DENMARK						0.000	-	0.231
Aarhus	91_2	506	1027	1959	3765			
Copenhagen	92_1	91	640	1827	3700			
Copenhagen	92_2	636	870	1809	3507			
Copenhagen	92_4	352	847	1693	3300			
Copenhagen	92_5	406	873	1825	3385			

(1) : p value for difference in mean total energy intake between *interviewers* nested by centre from the same country

(2) p value for difference in mean total energy intake between interviewers needed by centre from the same country, and adjusted for physical activity

(3) : p value for difference in mean total energy intake between *centres* from the same country

(*) : Total *energy intake* adjusted for age, body mass index, energy requirement, seasons, week days, special diet, special day

(~) : Only physical activity at work has been considered in the model (information on leasure time activity was not available for that country)

For men, the results show that after including physical activity in the model for Spain and Sweden, five countries (Italy, Spain, UK, Sweden and Denmark) out of seven do not differ in the mean energy intakes across interviewers. For the two other countries (The Netherlands and Germany), the adjustment for physical activity did not change the results. In The Netherlands, the differences across interviewers were no longer statistically significant, after adjustment for the EPIC-SOFT version used (p=0.0528). For Germany the difference in mean energy intake across interviewers was entirely due to one interviewer (No 71_3) out of the nine. When she was excluded, no statistical difference was observed between the remaining interviewers. For women the results are less consistent. In four countries (France, Italy, UK, Spain) out of eight there were no statistically significant differences among interviewers. The adjustment for physical activity of data from the Netherlands, Germany and Sweden did not change the results (physical activity was not available for Denmark). For Denmark, the difference was no longer statistically significant after the exclusion of one interviewer (No. 92_4) out of five. For women, no explanation was found for the differences between interviewers observed in The Netherlands and Sweden. In contrast to men, the adjustment for the EPIC-SOFT version did not change the results in Dutch women. The differences in mean energy intake according to interviewer were confirmed by similar differences in the mean number of items reported per 24-HDR interview or the percentage of low energy intake values. The difference in mean energy intake across centres from the same country is generally not statistically different, except for men in Italy and Germany, and for women in Spain and France. It is important to note that in Italy, Spain and France the EPIC centres are scattered all over the country, and that in Germany, one centre is located in the Eastern part (Potsdam) the other in the Western part (Heidelberg). True regional differences in energy intakes may explain the variations observed across centres. However it should be noted that the differences in mean energy intakes in such countries are not consistent across gender (only women are involved in France). The mean energy intake for each interviewer and the country mean, range from 90% to 105% for men and 86% to 114% for women (**Figures 5 a & b**), with the exception of interviewer No. 35_2 from Spain, who was shown to provide a mean energy intake comparable to the other interviewers after adjustment for physical activity (Table 1a & b). The percentage of interviewers with a mean energy intake within ±5% and ±10% of the country mean energy intake represents, respectively, 71% and 98% for men and 74% and 94% for women, whereas the difference between mean centre energy intakes and country mean energy intakes is within ± 5% for 82% of the centres for men and 89% for women (**Figures 6 a & b**).

Figure 5a: Difference (%) between the mean subject energy intake and the country mean energy intakes, per interviewer : Men

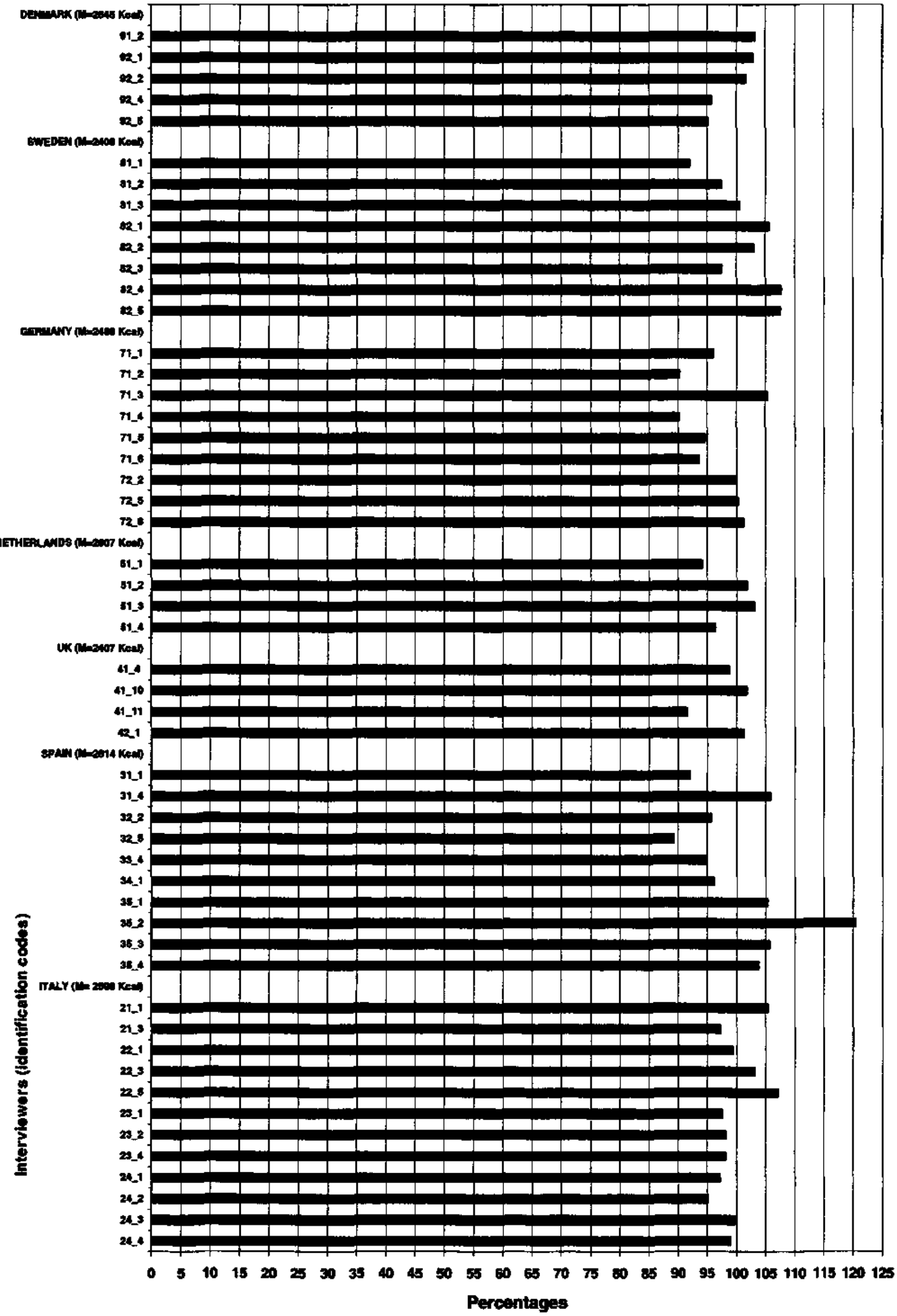


Figure 5b : Difference (%) between the mean subject energy intake and the country mean energy intakes, per interviewer : Women

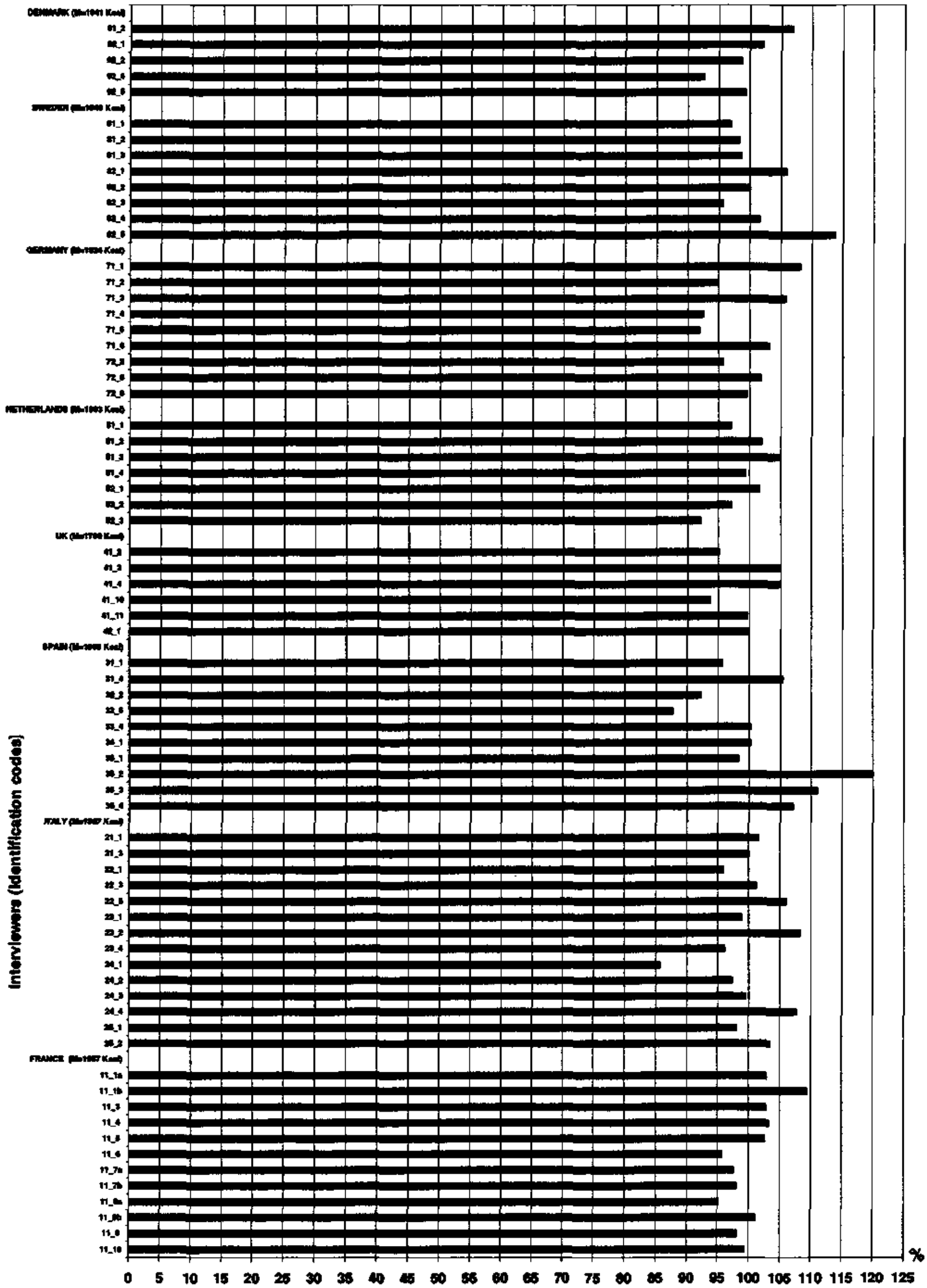


Figure 6a : Differences (%) between center and country mean energy intakes : Men

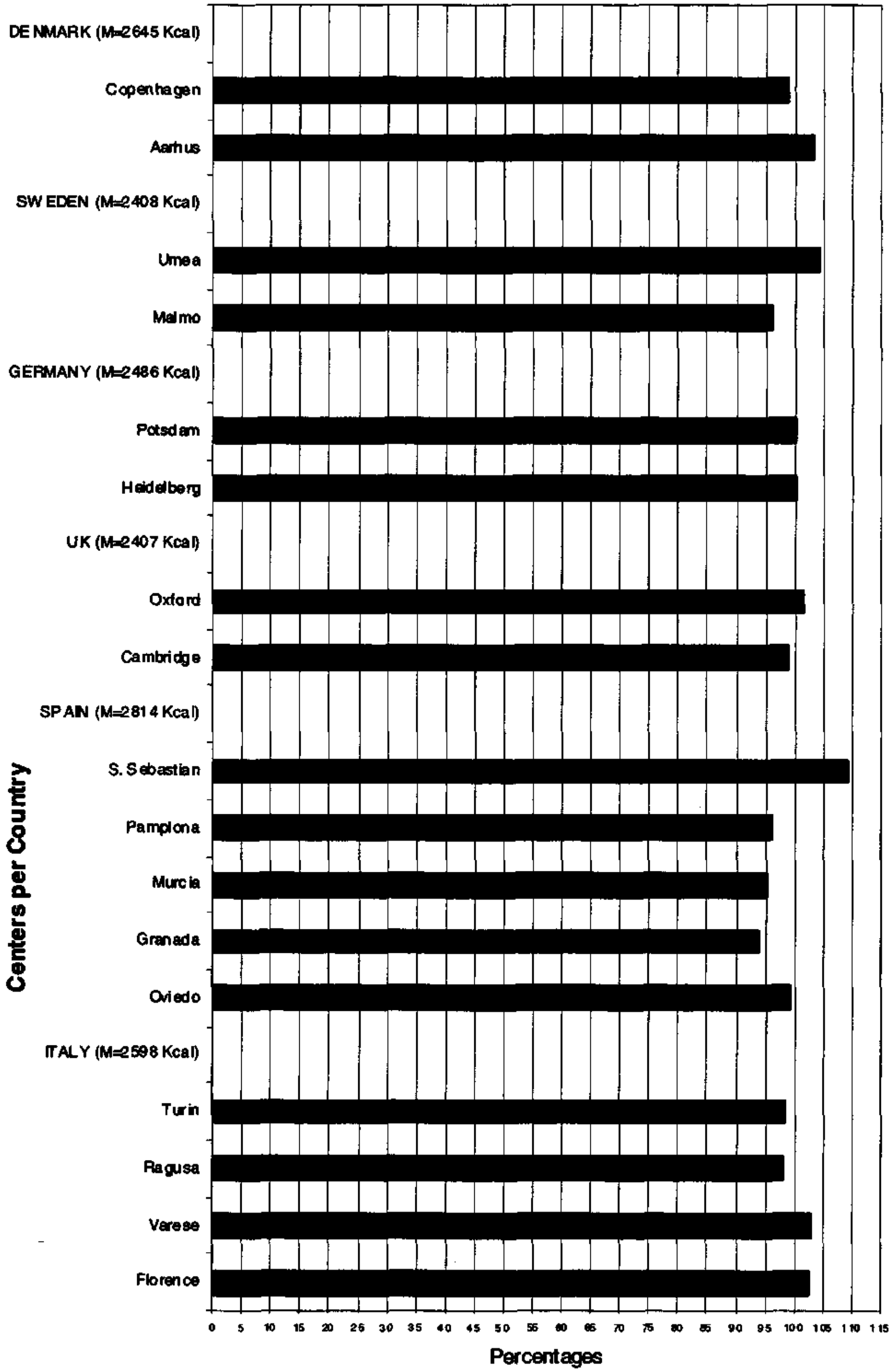
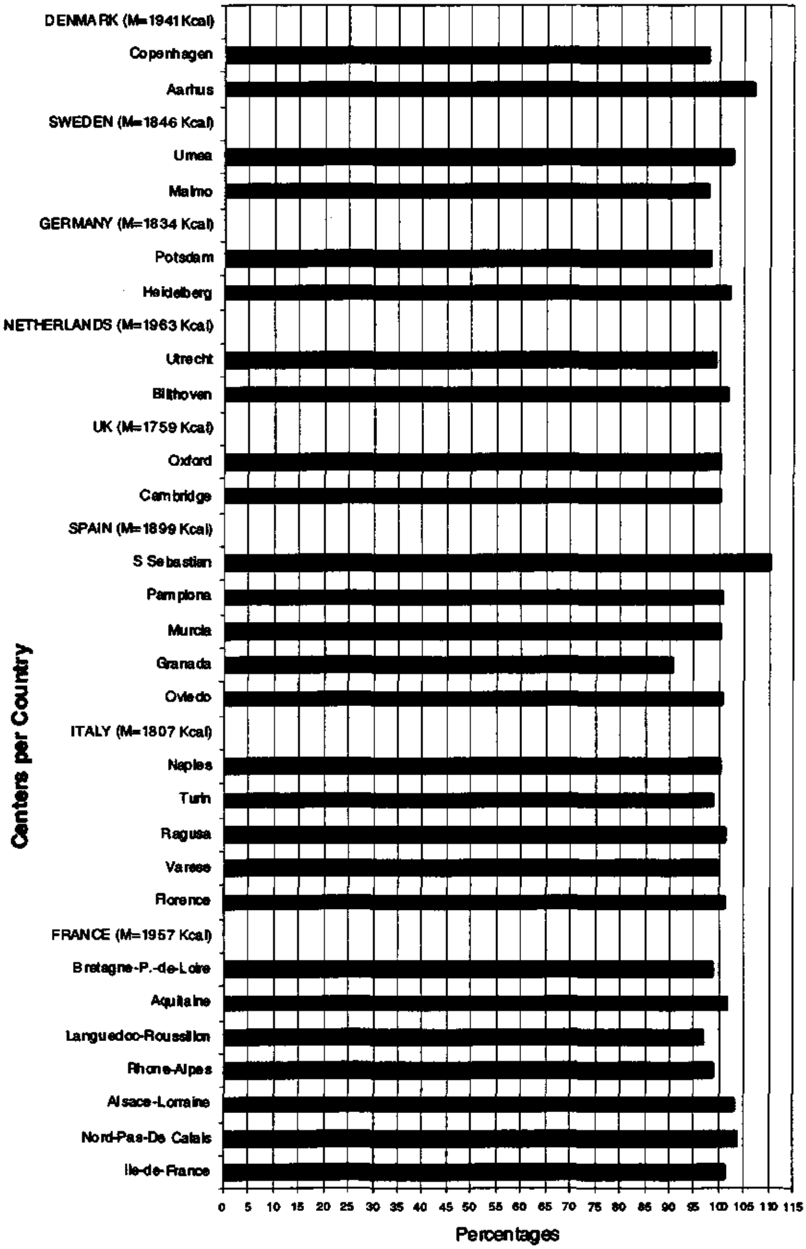


Figure 6b : Differences (%) between center and country mean energy intake : Women



Discussion

In this paper we have considered a number of potential sources of error associated with 24-HDR measurements, and we have discussed how to minimize many of them, particularly if a computerized interview approach is used. Other errors, however, such as those due to respondents' conscious or unconscious behaviour bias, remain unpredictable and difficult to prevent. These errors may have substantial implications for the analysis and interpretation of dietary intake measurements, particularly if they introduce differential bias in food reporting (72,73). As a preliminary evaluation of the standardization of the 24-h recall interviews collected with the EPIC-SOFT programme, we have presented a comparison of the mean total energy intakes across the 70 interviewers involved in the EPIC calibration study. Total energy intake was used as a dependent variable in a mixed model, after adjustment for a series of potential confounders. We felt that total energy intake was the only variable which could possibly disentangle the random and systematic measurements errors from the true and large natural variations in dietary intake existing across EPIC centres. Indeed, compared to other nutrients, mean energy intakes are expected to vary in a relatively narrow *physiological* range, relatively independently of food intake patterns, if correctly adjusted for. In our sex-specific analysis we adjusted total energy intake for age, BMI, predicted energy requirement, special diet, special day, and, when available, physical activity. Overall, the absence of large difference in mean energy intake across centres from the same country observed in our analysis is reassuring and confirms our physiological assumption. It may indicate that, at the group level, centres from the same country are quite comparable in energy intakes, after appropriate adjustments. The EPIC populations are mainly middle-aged, and physical activity is expected to be modest at the group level, particularly among women. This observation needs, however, to be confirmed by the ongoing EPIC studies on physical activity and other planned studies using doubly labeled water measurements. In addition, with the two variables (one discrete and the other continuous) used in our model it is not possible to compare the level of physical activity according to one absolute scale, which may explain some imprecision in the adjustment made. This study has shown that EPIC-SOFT provides comparable results across interviewers particularly for men. This is despite the high likelihood of detecting statistically significant differences due to the large number of interviewers, the heterogeneity of the study populations and the fact that no extreme energy values were excluded. In five out of seven countries (for men) and four out of eight countries (for women) no statistically significant differences were observed in mean energy intakes across interviewers nested by centre. The interviewer effect among men in the Netherlands disappeared after adjustment for the EPIC-SOFT version. However, as this effect was observed neither for women nor in the other countries where the same preliminary EPIC-SOFT programme version was used, we cannot eliminate an interviewer effect. For men in Germany and women in Denmark, the differences in energy intake observed among interviewers were entirely attributable to one interviewer. However, at least for Denmark we cannot exclude the possibility of a confounding effect, since it was not possible to adjust for physical activity, which was not available for the analyses. In the remaining countries (Netherlands, Sweden) for women, no explanation was found for the differences observed across interviewers, which may be due to a lack of standardization in the 24-HDR interviews across interviewers. However, the differences between interviewers, when compared with country mean energy intakes did not exceed 10% in most cases, despite the fact that no extreme energy intake values were excluded. The differences are even smaller (~5%) in both genders when the centres and country mean energy intakes are compared. The differences observed across interviewers may also be the result of a lack of control for all possible confounders, or other underlying problems. The example of the San Sebastian centre shows the importance of confounders in such a complex analysis. In this centre, the mean energy intakes estimated by one interviewer were about 400 Kcal higher than those of the others for both sexes. She was, however, operating in a rural area whereas the other interviewers were working in towns. The differences observed across interviewers disappeared when physical activity was entered in the model.

Other possible explanations for systematic differences across interviewers may be associated with underlying methodological problems. A surprising observation in this analysis is that a better standardization of the 24-HDR appears to have been obtained for men than women. It seems unlikely that these gender differences are attributable only to a lack of standardization across interviewers or to interviewers' effects. This observation may be related to a greater degree of under-reporting of food intake by women compared to men, which was consistently observed in all EPIC centres (unpublished data).

Although, overall, these results give a first encouraging appraisal regarding the level of standardization across the interviewers involved in EPIC, they do not prove the actual *absolute* reliability of the 24-HDR measurements obtained with EPIC-SOFT. The absence of statistical differences may reflect systematic errors equally applied across interviewers and centres from the same country. Further complementary analyses are therefore planned to compare the actual validity of 24-HDR intakes across centres using urinary nitrogen and potassium as gold standards, to estimate the magnitude of systematic measurement errors across the EPIC centres.

Conclusion

In general, the methodological issues related to the standardization of 24-HDR measurements described in this paper are not specific to the international study design of EPIC. Most of the problems discussed can be extrapolated to other dietary methods or study contexts. The results of our analysis of variance showed that, despite our efforts to standardize the method, problems of standardization still remain. Gender differences indicated that these problems are probably influenced by individual characteristics among women rather than by methodological factors. In addition, we demonstrated the strong potential role of confounders in such a heterogeneous study population, which need to be considered carefully in any methodological or etiological pooled data analysis. Further studies are needed to investigate the specific type and nature of measurement errors, particularly those related to under-reporting, frequently associated with 24-HDR or other open-ended methods. In addition, more comprehensive statistical models should be developed to correct for dietary measurement errors attributable to the interviewer, respondent and the dietary method, so that the overall understanding, analysis and interpretation of dietary data can be improved (74-77).

Acknowledgement

The work described in this paper was carried out with the support of the "Europe Against Cancer" Programme of the European Commission. The authors wish to thank Corinne Casagrande and Bertrand Hémon (IARC) for their contribution to the technical and programming work, and all the interviewers who participated to the field work studies in each EPIC centres. The authors would like also to thank Drs Loïc Le Marchand and Christine Friedenreich for their precious comments on the manuscript, and Dr. Antonio Ciampi for his contribution to the statistical analysis.

Institutes collaborating in the European Prospective Investigation into Cancer and Nutrition (EPIC)

Coordination	Lyon, International Agency for Research on Cancer (IARC)
France	Villejuif (Paris), INSERM, U 351, Institut Gustave Roussy
Germany	Heidelberg, German Cancer Research Centre
	Potsdam- Rehbrücke, German Institute for Human Nutrition
Denmark	Copenhagen, Danish Cancer Society
	Aarhus, Institute of Epidemiology and Social Medicine, University of Aarhus

Spain	Barcelona, Catalan Institute of Oncology (ICO)
	Granada, Andalusian School of Public Health
	Oviedo, Council for Health and Social Affairs of Asturia
	San Sebastian, Health Administration of Guipuzcoa (Basque Country)
	Murcia, Council for Health and Social Affairs of Murcia
	Pamplona, Institute of Public Health, Regional Government of Navarra
UK	Cambridge, Institute of Public Health, University of Cambridge
	Dunn Clinical Nutrition Centre, School of Clinical Medicine,Addenbrookes Hospital University of Cambridge
	Oxford, Cancer Epidemiology Unit, Imperial Cancer Research Fund, University of Oxford
Greece	Athens, University of Athens School of Medicine,
Italy	Milan, National Cancer Institute
	Florence, Centre for Cancer Research and Prevention (CSPO)
	Turin, Cancer Epidemiology Unit, University of Turin
	Ragusa, Ragusa Cancer Registry
	Naples, Department of Biomedical Sciences, and Human Oncology, University of Naples
Netherlands	Utrecht, Julius Center for Patient Oriented Research, Utrecht University Medical School
	Bilthoven, National Institute of Public Health & Environmental Protection (RIVM)
Sweden	Malmö, Departments of Medicine and Community Medicine, Lund University
	Umeå, Department of Nutritional Research, University of Umeå
Norway	Tromsø Institute for Community Medicine, University of Tromsø

References

1. Riboli E & Kaaks R (1997): The EPIC project: Rationale and study design. *Int. J. Epidemiol.* 26, Suppl. 1, S6-14.
2. Overvad K, Tjønneland A, Haraldsdóttir J, Bang S, Ewertz M & Møller-Jensen O (1991): Development of a semi-quantitative food frequency questionnaire to assess food, energy and nutrient intake in Denmark. *Int. J. Epidemiol.* 20, 906-912.
3. Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, Sneyd MI, Key TJA, Roe I & Day NE (1994): Comparison of dietary assessment methods in nutritional epidemiology : weighted records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. *Br. J. Nutr.* 72, 619-643.
4. Ocké MC, Bueno de Mesquita HB, Goddijn HE, Jansen A, Pols MA, van Staveren WA & Kromhout D (1997): The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire and validity and reproducibility for food groups. *Int. J. Epidemiol.* 27, Suppl. 1, S37-48.
5. Bohlscheid-Thomas S, Hoting I, Boeing H & Wahrendorf J (1997): The reproducibility and relative validity of a food frequency questionnaire newly developed for the German part of the EPIC project. I. Food group intake. *Int. J. Epidemiol.* 26, Suppl. 1, S59-70.
6. EPIC Group of Spain (1997): Validity and reproducibility of a diet history questionnaire in Spain. I. Foods. *Int. J. Epidemiol.* 26, Suppl. 1, S91-99.
7. Katsouyanni K, Rimm EB, Gnardellis C, Trichopoulos D, Polychronopoulos E & Trichopoulou A. (1997): Reproducibility and validity of an extensive semi-quantitative food frequency questionnaire using dietary records and biochemical markers among Greek school teachers. *Int. J. Epidemiol.* 26, Suppl. 1, S118-127.
8. van Liere MJ, Lucas F, Clavel F, Slimani N & Villeminot S (1997): The validity and reproducibility of a French dietary history questionnaire. *Int. J. Epidemiol.* 26, Suppl. 1, S128-136.
9. Pisani P, Vineis P, Krogh V & Berrino F (1997): Validity and reproducibility of the dietary questionnaire for the Italian component of EPIC. *Int. J. Epidemiol.* 27, Suppl. 1, S152-160.
10. Riboli E, Elmståhl S, Saracci R, Gullberg B & Lindgärde F (1997): The Malmö Food Study: Validity of two dietary assessment methods for measuring Nutrient intakes. *Int. J. Epidemiol.* 26, Suppl. 1, S161-171.
11. Kaaks R & Riboli E (1997): Validation and calibration of dietary intake in measurements in the EPIC project. *Int. J. Epidemiol.* 26, Suppl. 1, S15-25.
12. Plummer M, Clayton D & Kaaks R (1994): Calibration in multi-centre cohort studies. *Int. J. Epidemiol.* 23, 419-426.
13. Kaaks R, Plummer M, Riboli E, Estève J & van Staveren WA (1994). Adjustment for bias due to errors in exposure assessments in multi-center cohort studies on diet and cancer : a calibration approach. *Am J Clin Nutr* : 59, Suppl., S245S-250.
14. Kaaks R, Riboli E & van Staveren WA (1995): Calibration of dietary intake measurements in prospective cohort studies. *Am. J. Epidemiol.* 142, 548-556.

15. Madden JP, Goodman SJ & Guthrie HA (1976): Validity of the 24-hour recall. *J. Am. Diet. Assoc.* 68, 143-147.
16. Gersovitz M, Madden JP & Smiciklas-Wright H (1978): Validity of the 24-hr. dietary recall and seven-day record for group comparisons. *J. Am. Diet. Assoc.* 73, 48-53.
17. Block G (1982): A review of validations of dietary assessment methods. *Am. J. Epidemiol.* 115, 492-505.
18. Pietinen P (1985): An overview of the dietary methods used in epidemiologic studies on nutrition and cancer. In *Diet and Human Carcinogenesis,* eds. J Joossens, M Hill & J Geboers, pp 235-244. Amsterdam: Elsevier.
19. Fanelli MT & Stevenhagen KJ (1986): Consistency of energy and nutrient intakes of olders adults : 24-hour recall vs 1-day food record. *J. Am. Diet. Assoc.* 86, 665-667.
20. Morgan KJ, Johnson SR, Risek RL, Reese R & Stampley GL (1987): Collection of food intake data: an evaluation of methods. *J. Am. Diet. Assoc.* 87, 888-896.
21. Cameron ME & van Staveren WA (1988): *Manual on Methodology for Food Consumption Studies*. Oxford: Oxford Medical Publications.
22. Bingham SA (1987): The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutrition Abstracts and Reviews (series A)* 57, 705-742.
23. Black AE, Goldboerg GR, Jebb SA, Livingstone MBE, Cole TJ & Prentice AM (1991): Critical evaluation of energy intake data using fundamental principles of energy physiology : 2. Evaluating the results of published surveys. *Eur. J. Clin. Nutr.* 45, 583-599.
24. Black AE. Physical activity levels from a meta-analysis of doubly labeled water studies for validating energy intake as measured by dietary assessment (1996). *Nutr. Rev.* 54, 170-174.
25. Buzzard M (1998). 24-hour dietary recall and food record methods. ed. W Willett, *Monogr. Epidemiology Biostatist.* 30 50-73. New York, Oxford: Oxford University Press.
26. Slimani N, Deharveng G, Charrondière RU, van Kappel AL, Ocké MC, Welch A, Lagiou A, van Liere M, Agudo A, Pala V, Brandstetter B, Andrén C, Stripp C, van Staveren WA & Riboli E (1999): Structure of the standardized computerized 24-hour diet recall interview used as reference method in the 22 centres participating in the EPIC project. *Comput. Methods Programs Biomed,* 53, 251-266.
27. Krall EA, Dwyer JT & Coleman KA (1988): Factors influencing accuracy of dietary recall. *Nutr. Res.* 8, 929-941.
28. Jobe JB & Mingay DJ (1989): Cognitive research improves questionnaires. *Am. J. Public Health* 79, 1053-1055.
29. Jabine TB, Straf ML, Tanur JM & Tourangeau R, eds (1984): *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*. Washington: National Academy Press.
30. Fienberg SE, Loftus EF & Tanur JM (1985): Cognitive aspects of health survey methodology: an overview. *Milbank Mem. Fund Q.* 63, 547-564.

31. Dwyer JT, Krall EA & Coleman KA (1987): The problem of memory in nutritional epidemiological research. *J. Am. Diet. Assoc.* 87, 1509-1512.
32. Greger JL & Etnyre GM (1978): Validity of 24-hour dietary recalls by adolescent females. *Am. J. Public Health* 68, 70-72.
33. Krantzler NJ, Mullen BJ, Schutz HG, Grivetti LE, Holden CA & Meiselman HL (1982): Validity of telephoned diet recalls and records for assessment of individual food intakes. *Am. J. Clin. Nutr.* 36, 1234-1242.
34. Karvetti RL & Knuts LR (1985): Validity of the 24-hour dietary recall. *J. Am. Diet. Assoc.* 85, 1437-1442.
35. Brown JE, Tharp TM, Dahlberg-Luby EM, Snowdon DA, Ostwald SK, Buzzard IM, Rysavy SDM, Wieser SMA: Videotape dietary assessment : validity, reliability, and comparison of results with 24-hour dietary recalls from elderly women in a retirement home. *J. Am. Diet. Assoc.* 90, 1675-1679, 1990.
36. Johnson RK, Goran MI & Poehlman ET (1994): Correlates of over- and underreporting of healthy older men and women. *Am. J. Clin. Nutr.* 59, 1286-1290.
37. Smith WT, Webb KL & Heywood PF (1994): The implication of underreporting in dietary intakes. *Aust. J. Public Health*, 18, 311-314.
38. Klesges RC, Eck LH & Ray JW (1995): Who underreports dietary intake in a dietary recall? Evidence from the second National Health and Nutrition Examination Survey. *J. Consult. Clin. Psychol.* 63, 438-444.
39. Briefel RR, Sempos CT, Mc Doweel MA, Chien S, & Alaimo K (1997): Dietary method research in the Third National Health and Nutrition Examination Survey: underreporting of energy intake. *Am. J. Clin. Nutr.* 65, Suppl., S1203-1209.
40. van Staveren WA, de Groot LGPGM, Blauw YH & van der Wielen RPJ (1994): Assessing diets of elderly people: Problems and approaches. *Am. J. Clin. Nutr.* 59, Suppl., 221S-223S.
41. Poon LW (1985): Differences in human memory with aging: nature, causes, and clinical implications. In *Handbook of the Psychology of Aging (Ed. 2)* eds. JE Birren & KW Schaie, pp 427-461. New York: van Nostrand-Reinhold.
42. Nolan KA, & Blass JP (1992): Preventing cognitive decline. *Clin. Geriatr. Med.* 8, 19-34.
43. Pryer JA, Vrijheid M, Nichols R, Kiggins M & Elliott P (1997): Who are the low energy reporters in the dietary and nutritional survey of British adults? *Int. J. Epidemiol.* 26, 146-154.
44. Macdiarmid JI & Blundell JE (1997): Dietary under-reporting: what people say about recording their food intake. *Eur. J. Clin. Nutr.* 51, 199-200.
45. Johansson L, Solvoll K, Bjøneboe GEA & Drevon C (1998): Under- and overreporting of energy intake related to weight status and lifestyle in a nationwide sample. *Am. J. Clin. Nutr.* 68, 2666-2674.
46. Heitmann BL & Lissner L (1995): Dietary underreporting by obese individuals. Is it specific or non specific? *Br. Med. J.* 311, 986-989.

47. Lafay L, Basdevant A, Charles MA, Vray M, Balkau B, Borys JM, Eschwège E & Romon M (1997): Determinants and nature of dietary underreporting in a free-living population: The Fleurbaix Laventie Ville Santé (FLVS) study. *Int. J. Obes.* 21, 567-573.
48. Heerstrass DW, Ocké MC, Bueno-de Mesquita HB, Peeters PHM & Seidell JC (1998): Underreporting of energy, protein, and potassium intake in relation to body mass index. *Int. J. Epidemiol.* 27, 186-93.
49. Clark HH & Scholer MF (1992): Asking questions and influencing answers. In *Questions About Questions. Inquires into Cognitive Bases of Surveys*, ed. JM Tanur, pp 15-48. New York: Russel Sage Foundation.
50. Callmer E, Hagman U, Haraldsdóttir J, Bjørge Løken E, Seppänen R & Trygg K (1986): Standardisation of 24-hour recall, *Vår Föda* 38, Suppl. 4, 201-298.
51. Witschi JC (1990): Short-term dietary recall and recording methods. In *Nutritional Epidemiology*, ed. W Willett, *Monogr. Epidemiology Biostatist.* 15, 52-68. New York, Oxford: Oxford University Press
52. Fowler FJ & Mangione TW (1990): *Standardized Survey Interviewing: Minimizing Interviewer-Related Error.* Two chapters: "What is interviewer-related error?", pp 25-31, and "Interviewer selection and interviewer-related error", pp 96-105. Newbury Park, CA: Sage Publications.
53. Westat (1992): *NHANES-III Dietary Interviewer's Manual.* Rockville, MD: Westat Inc.
54. Feskanich D, Sielaff BH, Chong K & Buzzard M (1989): Computerized collection and analysis of dietary intake information. *Comput. Methods Programs Biomed.* 31, 47-57.
55. USDA (1992): *Langual Users Manual.* Washington DC: Centre for Food Safety and Applied Nutrition.
56. Hoover LW (1983): Computerized nutrient data bases: I. Comparison of nutrient analysis systems. *J. Am. Diet. Assoc.* 82, 501-505.
57. Frank GC, Hollatz AT, Webber LS & Berenson GS (1984): Effect of interviewer recording practices on nutrient intake- Bogalusa Heart Study. *J. Am. Diet. Assoc.* 84, 1432-1439.
58. Young LR & Nestle M (1995): Portion sizes in dietary assessment: Issues and policy implications. *Nutr. Rev.* 53, 149-158.
59. Cypel YS, Guenther PM, Petot GJ (1997): Validity of portion-size measurement aids: A review. *J. Am. Diet. Assoc.* 97, 289-292.
60. Haraldsdottir J (1993): Minimizing error in the field: quality control in dietary surveys. Eur. J. Clin. Nutr. 47(suppl 2), S19-S24.
61. van Kappel AL, Amoyel J, Slimani N, Vozar B & Riboli E (1994): *EPIC-SOFT Picture Book for Estimation of Food Portion Sizes.* Lyon: International Agency for Research on Cancer.
62. Faggiano F, Vineis P, Cravanzola D, Pisani P, Xompero G, Riboli E & Kaaks R (1992): Validation of a method for the estimation of food portion size. *Epidemiology* 3, 379-382.
63. Lucas F, Niravong M, Villeminot S, Kaaks R & Clavel-Chapelon F (1995): Estimation of food portion size using photographs: Validity, strengths, weaknesses and recommendations. *J. Hum. Nutr. Diet.* 8, 65-75.

64. Nelson M, Atkinson M & Darbyshire S (1994): Food photography I: the perception of food portion size from photographs. *Br. J. Nutr.* 72, 649-663.
65. Nelson M, Atkinson M & Darbyshire S (1996): Food photography II: use of food photographs for estimating portion size and the nutrient content of meals. *Br. J. Nutr.* 76, 31-49.
66. Krebs-Smith SM, Cronin FJ, Haytowitz DB, Cook A (1990): Contributions of food groups to intake of energy, nutrients, cholesterol, and fiber in women's diet: effect of method of classifying food mixtures. *J Am. Diet. Assoc.* 90, 1541-1546.
67. Bergström L (1994): *Nutrient Losses and Gains in Preparation of Foods.* Uppsala, Sweden: Livsmedelsverket (National Food Administration). Rapport 32/1994.
68. Beaton GH, Milner BA, Corey P, McGuire V, Cousins M, Stewart E, de Ramos M, Hewitt D, Grambsch PV, Kassim N & Little Ja (1979). Sources of variance in 24-hour dietary recall data : Implications for nutrition study design and interpretation. *Am. J. Clin. Nutr.* 32, 2546-2559.
69. Deharveng G, Charrondière R, Slimani N, Riboli E & Southgate DAT (1999): Comparison of food composition tables available in the nine European countries participating in EPIC. *Eur. J. Clin. Nutr.* 53, 60-79.
70. Slimani N, Charrondière UR, van Staveren W, Riboli E (2000). Standardization of food composition databases for the European Prospective Investigation into Cancer and Nutrition (EPIC): General theoretical concept. J. Food Comp. Anal. 13, 567-584.
71. James WPT & Schofield EC (1990). Human Energy requirements. A manual for planners and nutritionists. FAO. New York, Oxford: Oxford University Press.
72. Prentice RL (1996): Measurement error and results from analytic epidemiology : dietary fat and breast cancer. *J. Natl. Cancer Inst.* 88, 1738-1747
73. Paeratakul S, Popkin BM, Kohlmeier L, Hertz-Picciotto I, Guo X, Edwards LJ. Measurement error in dietary data: implications for the epidemiologic study of the diet-disease relationship (1998). *Eur. J. Clin. Nutr.* 52, 10, 722-727.
74. Beaton GH, Burema J & Ritenbaugh C (1997): Errors in the interpretation of dietary assessments. *Am. J. Clin. Nutr.* 65, Suppl., S1100-1107.
75. Kipnis V, Freedman LS, Brown CC, Hartman AM & Schatzkin A (1997):. Effect of measurement error on energy-adjustment models in nutritional epidemiology. *Am. J. Epidemiol.* 146, 842-851.
76. Freedman L (1998): Challenges for statistical approaches to dietary assessment. *Eur. J. Clin. Nutr.* 52, Suppl. 2, S6
77. Willett W (1989): An overview of issues related to the correction of non-differential exposure measurement error in epidemiological studies. *Stat. Med.* 8, 1031-1040.

Annex 1 : Common EPIC-SOFT food and mixed recipe classification systems

1. Common Food Classification System

00 Unclassified

01 Potatoes And Other Tubers
01 00 Unclassified
01 01 Potatoes
01 02 Other Tubers

02 Vegetables
02 00 Unclassified
02 01 Leafy Vegetables (Except Cabbages)
02 02 Fruiting Vegetables
02 03 Root Vegetables
02 04 Cabbages
02 05 Mushrooms
02 06 Grain and Pod Vegetables
02 07 Onion, Garlic
02 08 Stalk Vegetables, Sprouts
02 09 Mixed Salad, Mixed Vegetables

03 Legumes
03 00 Unclassified
03 01 Legumes

04 Fruits
04 00 Unclassified
04 01 Fruits
04 02 Nuts and Seeds (+ Nut Spread)
04 03 Mixed Fruits
04 04 Olives

05 Dairy Products
05 00 Unclassified
05 01 Milk
05 02 Milk Beverages
05 03 Yogurt
05 04 Fromage Blanc, Petits Suisses
05 05 Cheeses (Including Fresh Cheeses)
05 06 Cream Desserts, Puddings (Milk Based)
05 07 Dairy and Non Dairy Creams
05 07 00 Unclassified
05 07 01 Dairy Creams
05 07 02 Non Dairy Creams
05 08 Milk for Coffee and Creamers

06 Cereals and Cereal Products
06 00 Unclassified
06 01 Flour, Flakes, Starches, Semolina
06 02 Pasta, Rice, Other Grain
06 03 Bread, Crispbread, Rusks
06 03 00 Unclassified
06 03 01 Bread
06 03 02 Crispbread, Rusks
06 04 Breakfast Cereals
06 05 Salty Biscuits, Aperitif Biscuits, etc..
06 06 Dough and Pastry

1. Common Food Classification System (continued)

07 Meat and Meat Products
07 00 Unclassified
07 01 Fresh Meat
07 01 00 Unclassified
07 01 01 Beef
07 01 02 Veal
07 01 03 Pork
07 01 04 Mutton/Lamb
07 01 05 Horse
07 01 06 Goat
07 02 Poultry
07 02 00 Unclassified and Other Poultry
07 02 01 Chicken, Hen
07 02 02 Turkey, Young Turkey
07 02 03 Duck
07 02 04 Goose
07 02 05 Rabbit (Domestic)
07 03 Game
07 04 Processed Meat
07 05 Offals

08 Fish And Shellfish
08 00 Unclassified
08 01 Fish
08 02 Crustaceans, Molluscs
08 03 Fish Products, Fish In Crumbs

09 Eggs And Egg Products
09 00 Unclassified
09 01 Egg

10 Fat
10 00 Unclassified
10 01 Vegetable Oils
10 02 Butter
10 03 Margarines
10 04 Deep Frying Fats
10 05 Marine Oil
10 06 Other Animal Fat

11 Sugar and Confectionery
11 00 Unclassified
11 01 Sugar, Honey, Jam
11 02 Chocolate, Candy Bars, Paste, ..
11 03 Confectionery non Chocolate
11 04 Syrup
11 05 Ice Cream, Water Ice
11 05 00 Unclassified
11 05 01 Ice Cream
11 05 02 Sorbet
11 05 03 Water Ice

12 Cakes
12 00 Unclassified
12 01 Cakes, Pies, Pastries, etc…
12 02 Dry Cakes, Biscuits

13 Non Alcoholic Beverages
13 00 Unclassified
13 01 Fruit and Vegetable Juices
13 02 Carbonated/Soft/Isotonic etc.
13 03 Coffee, Tea and Herbal Teas
13 03 00 Unclassified
13 03 01 Coffee
13 03 02 Tea
13 03 03 Herbal Tea
13 03 04 Chicory, Substitutes
13 04 Waters

1. Common Food Classification System (continued)

14 Alcoholic Beverages
14 00 Unclassified
14 01 Wine
14 02 Fortified Wines (Cherry,Porto,..)
14 03 Beer, Cider
14 04 Spirits, Brandy
14 05 Aniseed Drinks (Pastis,..)
14 06 Liqueurs
14 07 Cocktails, Punches

15 Condiments And Sauces
15 00 Unclassified
15 01 Sauces
15 01 00 Unclassified And Other Sauces
15 01 01 Tomato Sauces
15 01 02 Dressing Sauces
15 01 03 Mayonnaises And Similars
15 01 04 Dessert Sauces
15 02 Yeast
15 03 Spices, Herbs and Flavourings
15 04 Condiments

16 Soups, Bouillon
16 00 Unclassified
16 01 Soups
16 02 Bouillon

17 Miscellaneous
17 00 Unclassified
17 01 Soya Products
17 02 Dietetic Products
17 02 00 Unclassified
17 02 01 Artificial Sweeteners
17 03 Snacks

2. Common Mixed Recipe Classification System

Only the main groups are common between countries, the sub-groups are country specific.

01 Based On Meat And Meat Products
02 Based On Fish, Crustaceans
03 Based On Eggs
04 Based On Cereals And Cereal Products
05 Vegetables And Vegetable Products (Including Potatoes)
06 Mixed Beverages
07 Other

Chapter 3

Structure of the standardized computerized 24-h diet recall Interview used as reference method in the 22 centres participating in the EPIC Project

Nadia Slimani, Geneviève Deharveng, Ruth U. Charrondière, Anne Linda van Kappel, Marga C. Ocké, Ailsa Welch, Areti Lagiou, Marti van Liere, Antonio Agudo, Valeria Pala, Birgit Brandstetter, Caren Andrén, Connie Stripp, Wija A. van Staveren, Elio Riboli

Comput Meth Programs Biomed 1999; 53: 251-266

Abstract

A computerized 24-hour diet recall interview programme (EPIC-SOFT) was developed for use as a common reference calibration method between 9 European countries involved in a large prospective cohort study, the European Prospective Investigation into Cancer and Nutrition (EPIC). The EPIC-SOFT programme, which was adapted for each participating country and translated into 9 languages (Danish, Dutch, English, French, German, Greek, Italian, Sweden and Spanish), was developed to perform highly standardized interactive interviews in 22 centres involved in the study. The present paper describes the general conceptual structure and presents the overall interview procedure of EPIC-SOFT.

In this software, common rules were pre-entered into the system to describe, quantify and check automatically about 1500 to 2200 foods and 150 to 350 recipes of the different EPIC-SOFT versions. These criteria were adapted to the way and level of detail with which subjects tend spontaneously to report their diet. A picture book, containing 140 sets of colored photos, was also prepared to help subjects estimate their food portions. The photos were critically selected to cover the needs of the different EPIC centres, and produced under standardized conditions (i.e. picture taken under the same angle, same food representation, etc.). The dietary data collected with national versions of EPIC-SOFT are pre-coded and classified according to a common classification system. In addition, accompanying software programmes are available to recalculate the 24-hour diet recall data regularly according to the most up-to-date EPIC-SOFT version, and export them according to a common format, which facilitates their exchange, comparison and combined analyses.

So far, EPIC-SOFT is the only available computerized 24-hour diet recall system providing comparable food consumption data between several European countries. About 30,000 24-hour diet recalls have been collected so far in the nine EPIC countries and it takes about 30-35 minutes to complete an interview, whatever the country. Statistical analyses are planned to evaluate the level of standardization achieved within and between the EPIC centres. The 24-hour diet recall data will also be validated using chemical biological markers as independent reference measurements (e.g. total urine nitrogen).

Introduction

The European Prospective Investigation into Cancer and Nutrition (EPIC) is a large multi-centre prospective cohort study involving about 460,000 middle-aged men and women from nine European countries (Spain, Italy, France, Greece, Germany, The Netherlands, United Kingdom, Sweden and Denmark) (1). Within each country, a dietary questionnaire was developed and validated to estimate the usual individual food intakes of the study subjects (2-10). An internal calibration method was adopted, based on a single 24-hour diet recall interview collected from a representative sample of about 1200 to 5000 individuals per cohort. The rationale for this approach was to use the 24-hour diet recall interview as a common reference method to adjust, at the group level, for systematic over- or under-estimation of the mean food (or nutrient) intakes measured by the country-specific baseline questionnaires (11-13). At the same time, this calibration method will be used to correct for attenuation bias when measuring the relative risk between diet and disease. The statistical power, to detect a real relationship between the dietary exposure and cancer or other chronic diseases, could thereby be increased.

One of the main statistical conditions for using the calibration approach in the EPIC project was that 24-hour diet recall is highly standardized within and between countries. Despite some recent interest in the calibration concept in studies on diet (14), there is still little experience with the methodological and practical problems of standardizing dietary assessment methods within the context of large international multi-centre studies. A software programme named "EPIC-SOFT" was therefore developed at the International Agency for Research on Cancer (IARC), in collaboration with the EPIC centres, to ensure the highest possible level of standardization of the 24-hour diet recall interviews collected in the 22 EPIC centres.

The development of EPIC-SOFT was conducted in three main steps. The first stage was to identify the potential source of errors associated with the 24-hour diet recall measurements by carrying out a literature review (15-18) and by analyzing a large pool of 24-hour diet recalls collected during the EPIC pilot phases. The second stage was to develop the approaches to prevent or minimize each identified potential source of error associated with the 24-hour diet recall in the international context of the EPIC study. The third stage was to develop the EPIC-SOFT programme and country-specific versions prepared for each of the nine participating countries.

In this paper, we present the EPIC-SOFT programme and its common structure and interview interface adopted to optimize the standardization of the dietary interview procedure within and between the EPIC centres.

The rationale and methodological concepts developed for standardizing the computerized 24-hour diet recall interviews are presented in another paper (19).

Main objectives and functions of the software

The main functions and logical structures implemented in the EPIC-SOFT programme were designed to conduct interactive dietary interviews following a strictly standardized procedure within and between the EPIC centres:

a. Information on all the foods and beverages consumed during the recall day needed to be collected, entered, and coded automatically according to common rules,
b. The software had to be user-friendly and convenient for use in large populations of different linguistic, socio-cultural, ethnic and geographical origins,
c. The quantity of the food as finally consumed (e.g. cooked and/or without inedible part), had to be calculated automatically by the system whatever way the food/recipe was quantified (e.g. raw, cooked, with or without inedible part),
d. Systematic quality controls had to be performed throughout the interview procedure on the information reported by the subject and entered by the interviewer so that possible errors or suspicious answers could be detected and clarified with the subject during the interview,
e. The procedures for maintaining and updating the common and country-specific EPIC-SOFT databases (e.g. food or recipe lists and portion sizes) had to be standardized using additional tools and specially designed software,
f. The storage, method of retrieval and export of dietary data had to be standardized to facilitate the use, exchange and (pooled) analyses of the 24-hour diet recalls.

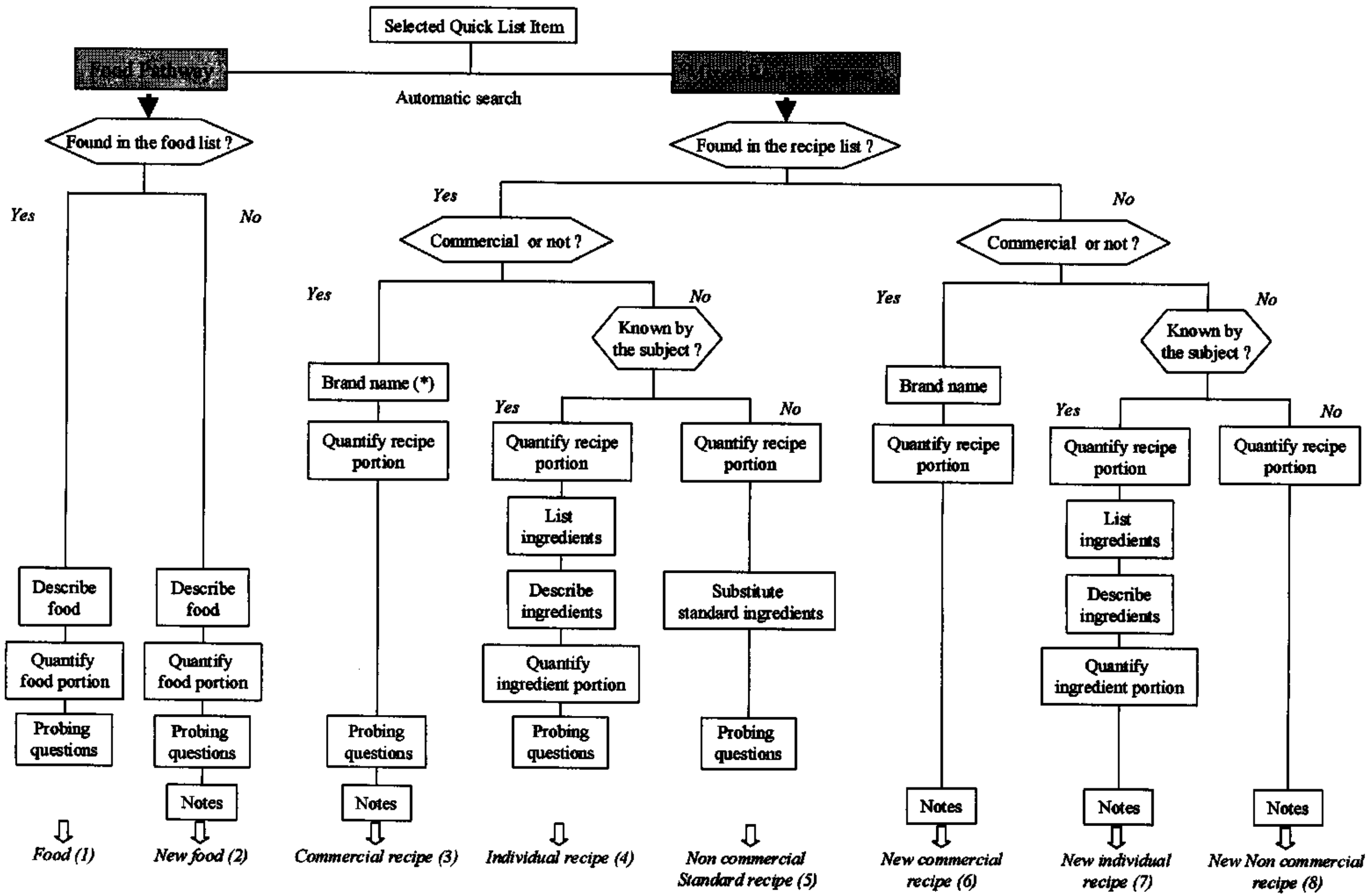
General conceptual structure of EPIC-SOFT

The general structure of EPIC-SOFT (**Figure 1**) was designed to pre-define and standardize the pathways to be followed during the 24-hour diet recalls across all countries. This overall structure relies first on a common definition of "foods", or "mixed recipes" that might be reported by the subjects: a "*food*" is a single item which is perceived and quantified as such (e.g. fruits, vegetables, fish, cakes, bread, sweets). A "*mixed recipe*" is a complex dish where all the ingredients cannot be distinguished visually and quantified separately (e.g. moussaka). To facilitate the distinction between these two terms for the interviewers, most of the frequently consumed foods or mixed recipes in the different EPIC countries were pre-entered in two separate lists containing respectively about 1500 to 2200 foods and 150 to 350 mixed recipes depending on the country. During the interview, each item recalled is automatically searched for in the pre-defined food and recipe lists. Depending on whether the item is a "food" (or ingredient of mixed recipe) or a "mixed recipe", one of two main pathways is followed. The subsequent division of the hierarchical tree into several n-branches is obtained by several conditional key questions (i.e. "item available or not in the food or recipe list"; "recipe known or not by the subject", "commercial or non commercial recipe") which are automatically managed by the system or manually entered by the interviewer.

The interview procedure is standardized whatever the pathway followed. For example, the principal steps of the interview procedure (description, quantification, checks using probing questions) remain the same, and in the same sequence for all the pathways. In addition, the initial automatic search in the pre-defined recipe list forces the interviewer to use a standard pre-entered recipe, if available, as an entry point into the system. This approach ensures that for the same recipe, the same method of quantifying the portion sizes and the same probing questions are used in all situations, whether the recipe is homemade or commercial or whether or not the subject knows the recipe (*pathway 3, 4, 5*). If the subject knows the recipe, the standard recipe is used only as an aid, and the individual ingredients reported are listed and successively described and quantified (*pathway 4*). If not, the pre-defined standard recipe is used as reference (*pathway 5*). In the latter situation, it is possible to substitute only *qualitatively* certain standard ingredients by ingredients known by the subject and actually consumed (e.g. replace fat not specified by olive oil). This approach minimizes the unavoidable differences in the precision reported between subjects who know the recipe consumed and those who do not (for example, between women and men).

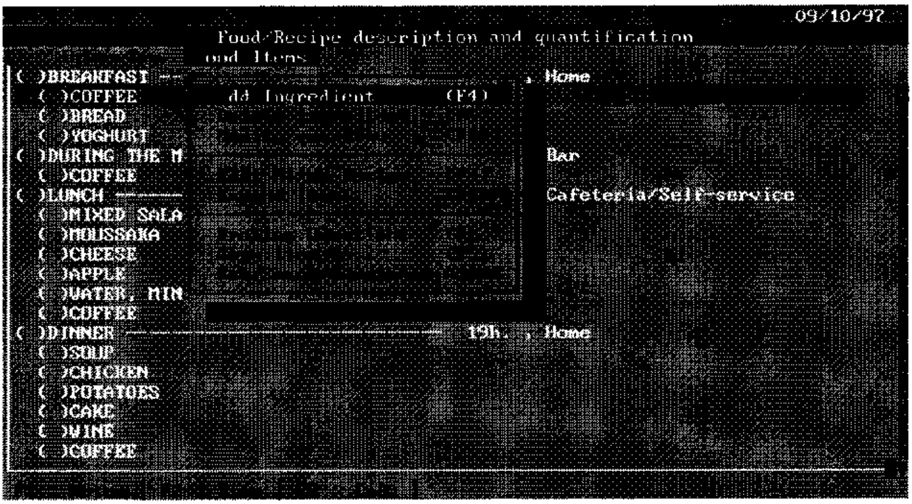
When a food or recipe is missing from the pre-defined lists, default options are proposed to the interviewer for describing, quantifying and checking the new item reported (*pathway 2, 6, 7, 8*). Additionally, at the end of each of these pathways, a free text file comes up automatically on the screen, to enter as many notes as possible on the missing items, which can be used afterwards to update the missing information in the EPIC-SOFT file according to a standardized procedure described later in this paper.

Figure 1 : Flow diagram of the EPIC-SOFT interview procedure.



Presentation of the EPIC-SOFT interface

The EPIC-SOFT interface was developed using the same basic design common to all country versions. This interface was conceived to guide equally all the interviewers during the whole 24-hour diet recall interview and minimize potential errors which might be due to varying degrees of knowledge of nutrition or computing. All common functions, in a given screen, are summarized in a pull down menu and can always be activated through the mouse, hot keys or using the keyboard (**Screen 1**).

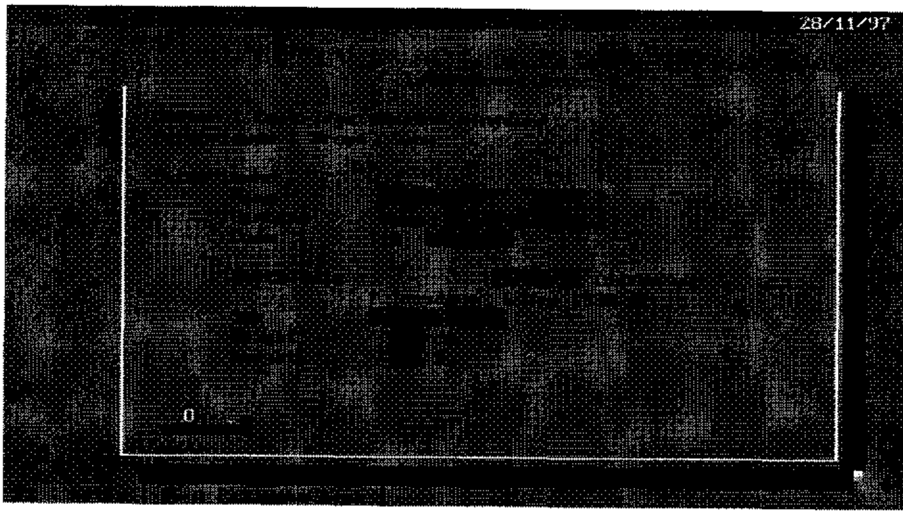


***Screen 1** : Example of pull down menu used*

The 24-hour diet recall interview performed with EPIC-SOFT is divided into four main steps: 1) *general non dietary information, 2) quick list*, 3) *description and quantification of foods and recipes, 4) quality controls at nutrient level*. Each of these stages is discussed below:

General non-dietary information

General information is collected systematically on the *interviewer, interviewee* and *recalled day* so that systematic checks can be carried out during and after the interview (**Screen 2**). For example, the information on the subject's age, sex, height and weight is used at the end of the interview to calculate the energy requirements and compare them with energy intakes estimated from the 24-hour diet recall. Information on whether the recalled day was a *special day* (for instance religious celebration, travel) or whether a *special diet* was followed (diabetes, hypertension, vegetarian) is also collected to explain and/or exclude outlier values (e.g. very low total energy intake due to illness), at the stage of the statistical analysis.



Screen 2 : General information on the interviewee

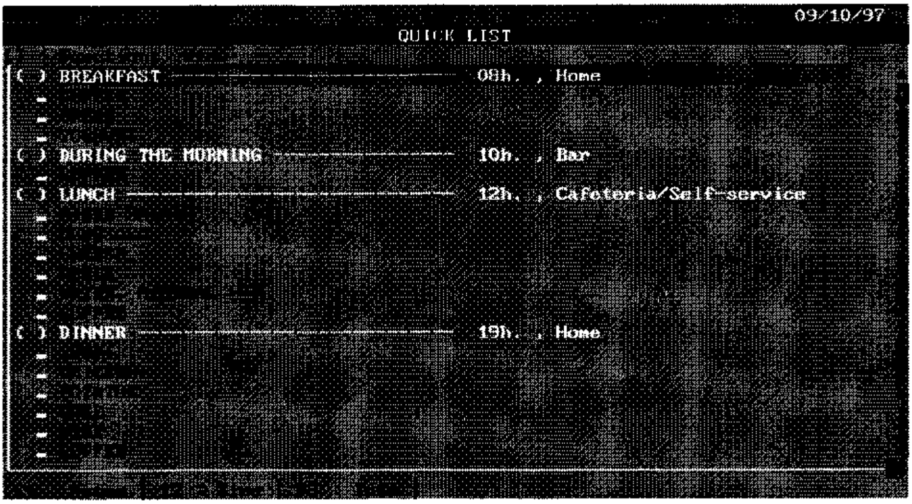
In this first part of the interview, the questions to be answered are indicated as empty fields on the screen. The interviewer can either write down or select from a pre-defined list the answers given by the subject. The system then checks for incomplete fields or extreme values entered manually. For example, if the weight or height is outside the pre-defined country-specific range, a warning comes up on the screen and the entry has to be confirmed. In addition, to speed up the interview and minimize typing errors, a file containing all the information on the subject to be interviewed can be loaded in EPIC-SOFT. When, during the interview, the subject is selected from this pre-defined list, all the individual characteristics (name, surname, age, sex) are automatically displayed in the corresponding fields.

Once the two general information screens are completed, the dietary recall interview can start.

Quick list

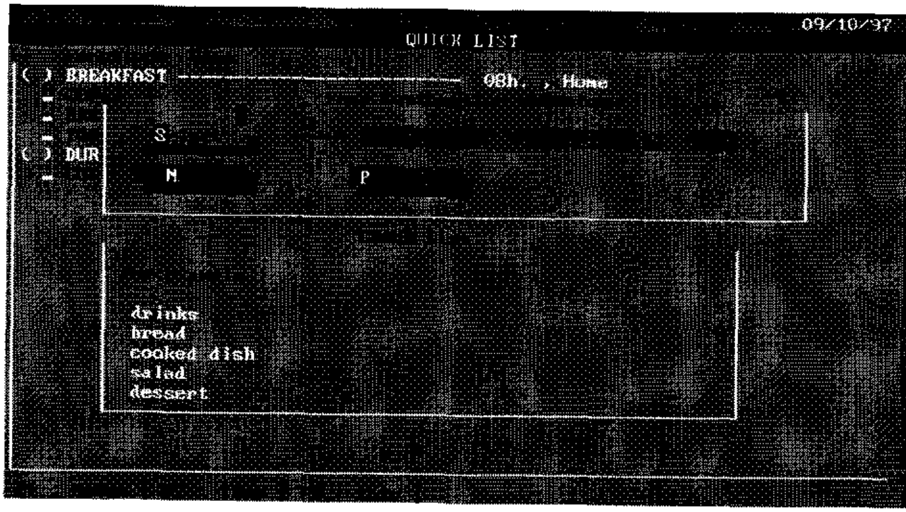
The so-called "quick list" is an open field where the interviewer manually enters all the foods and recipes consumed the day preceding the interview in chronological order (QLI : quick list items). This computerized approach is adapted from the traditional method of conducting a 24-hour diet recall, where all the foods consumed the day before are reported on a sheet of paper.

The main objective of this first phase is for the subject to remember all the foods and beverages consumed the previous day with no omissions. To focus the attention of both the interviewer and interviewee, the information should be reported in generic terms (e.g. milk, yogurt, bread, moussaka) with no attempt at this stage to describe or quantify the foods in detail (**Screen 3**)



***Screen 3** : Example of quick list (QL) already filled in*

The foods are entered per meal (or "food consumption occasions": FCO) common to all EPIC countries. Eleven FCOs cover the entire period of the day starting from before breakfast to during the night. This chronological structure has two purposes. First of all, as a cognitive approach to help the subjects remember what they did the previous day and then associate what they consumed. Second, as a way to standardize the interview procedure by covering the entire day and asking the same sequence of questions on the foods consumed at the different meal occasions. Each FCO comes up on the screen in chronological order (**Screen 4**).



Screen 4 *: Example of Food consumption occasion (e.g Lunch) and related check list*

The subject then has to answer whether or not anything was consumed at that particular FCO. If so, the subject is asked to indicate the place and approximate time as a further cognitive aid. The interviewer lists all the foods reported spontaneously by the subject, and checks the completeness of the answers using a "check list" automatically displayed on the screen (e.g. for lunch: drinks, bread, cooked dish, salad, dessert) (Screen 4). The interviewer is not allowed to go on to the next stage of the interview until all the FCOs have been covered, and a check has been made on the empty FCOs. The interviewer has therefore still the possibility to make any changes or additions before ending the quick list.

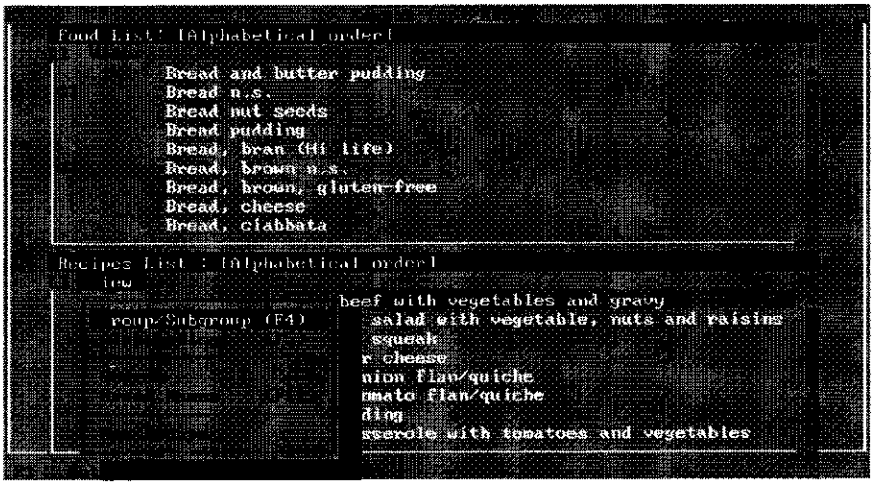
Description and quantification of foods and recipes

Once the quick list is completed, the software moves to the next phase in which, each quick list item (QLI) is described in detail and quantified following four systematic steps displayed automatically on the screen: *a. Search/Identification, b. Description, c. Quantification, d. Probing questions.*

(a) Search/Identification

The search/identification stage involves finding out from the pre-defined food or recipe lists the items corresponding to those actually consumed (**Screen 5**). The QLI is often generic (e.g. cake), and at this stage of the interview the subject is asked additional questions to describe or name precisely the food actually consumed (e.g. fruit pie, wholemeal, one crust).

Different methods of searching in the food or recipe lists are available. The QLI is first automatically searched alphabetically in both the food and recipe lists. In addition, two alternative searches are available by main food or recipe (sub-) groups or by searching a given string whatever its position in the name (e.g. all the foods containing the term "chocolate").



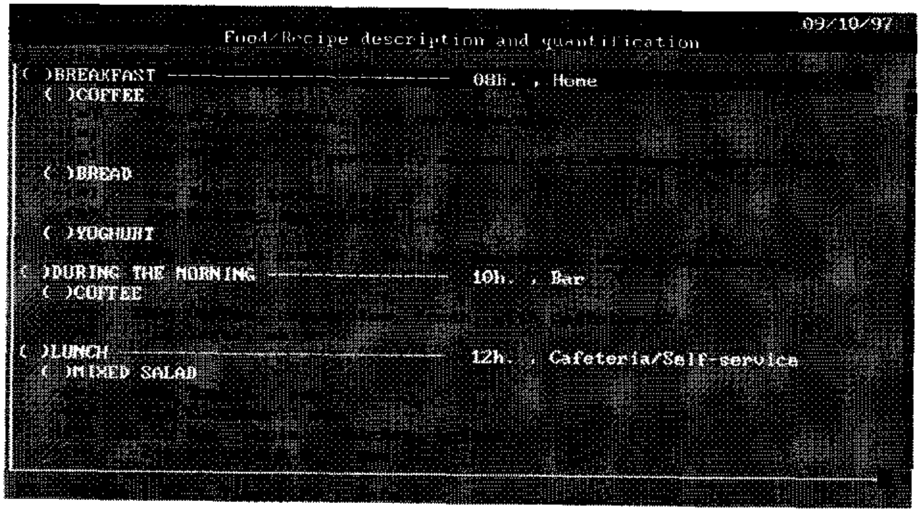
***Screen 5** : Search functions in the pre-defined food and recipe lists*

If the food or recipe consumed is not found in the lists, the interviewer has the possibility of recording a temporarily "new" food or recipe. The interviewer classes the "new" food in the most appropriate (sub-) group and the food is described and quantified according to default options available for the given (sub-) group. These new items are, however, flagged so that, after controls, they can be added to the EPIC-SOFT databases.

(b) Description of foods/ingredients and recipes

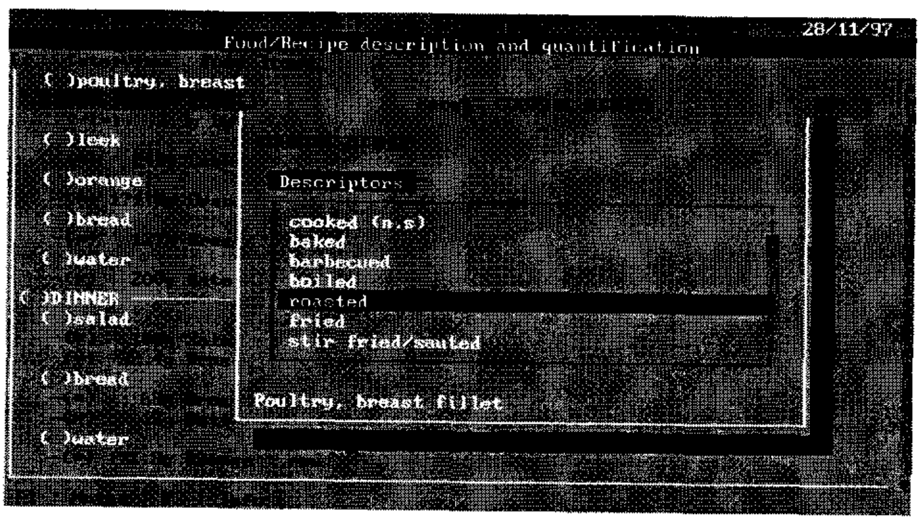
When the item actually consumed is identified, the food or mixed recipe pathways are followed to describe and subsequently quantify the item. Mixed recipes are broken down into ingredients that are described and quantified as foods, following the same pathway as (simple) foods.

All foods or ingredients identified from a given quick list item are described according to a pre-defined, standardized level of detail depending on the food (sub-) group (**Screen 6**). A sequence of windows corresponding to questions to be asked on different characteristics of foods called "facets" (e.g. cooking method, preservation method, etc.) are automatically displayed on the screen.



***Screen 6** : Example of complete description and quantification of quick list items (QLIs)*

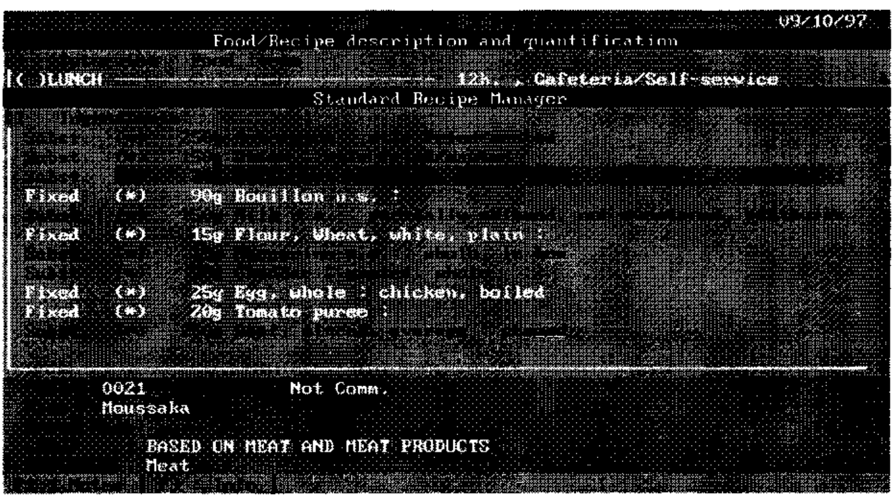
Each of these windows contains a series of "descriptors" corresponding to the pre-defined possible answers (e.g. "grilled", "roasted", "fried" for the facet "cooking method") (**Screen 7**). If the subject is not able to answer the question precisely, generic description names "non specified" are available in each facet (e.g. cooked n.s for the facet "cooking method").



***Screen 7** : Example of « cooking method » facet/descriptors pair*

For some facets such as the facet "brandname" it is possible to add manually new brand names not available in the pre-defined list. This facet/descriptor approach has two advantages: first, to standardize the interview procedure between interviewers by forcing them to ask the same questions systematically and in the same order; second, to standardize the level of detail to be asked to describe given types of foods within and between countries.

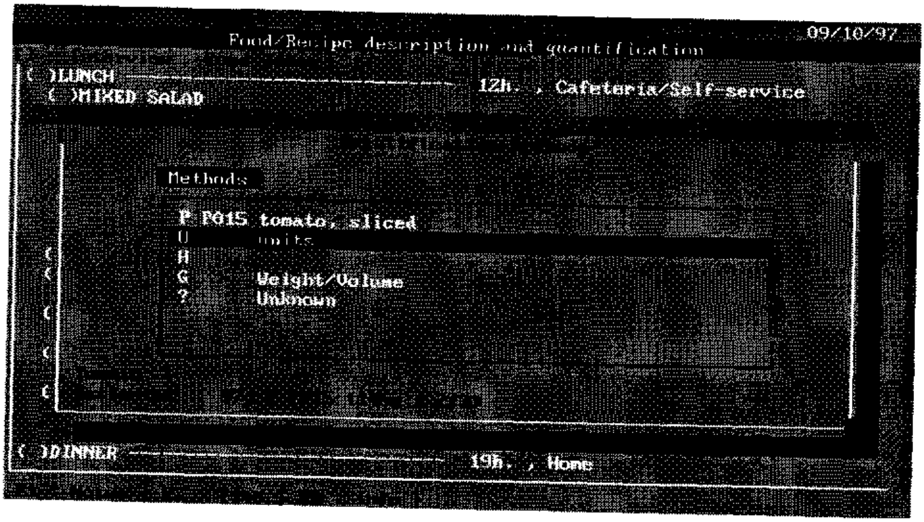
The mixed recipes are treated in a so-called "recipe manager" screen, where the information on the recipe (name, classification into main (sub-) groups, identification code) and its ingredients can be entered when the subject knows his/her recipe (*"Individual recipes")*. If the subject is not able to report the ingredients of the recipe consumed, a *"standard recipe"* can be selected which should be as close as possible to what the subject consumed. However, some ingredients in these standard recipes, which may vary between subjects (e.g. fat, milk), can be substituted if the subject provides more specific information on the actual food consumed (e.g. olive oil, whole cow's milk) (**Screen 8**). If the subject does not know the recipe or it is not available in the recipe list, all the information he/she actually provides can be entered in the so-called *"note file"* which comes up automatically on the screen. This information can then be used after the interview to update the recipe file.



***Screen 8** : Example of standard mixed recipe with possible substitution of ingredients (e.g. oil, vegetable n.s.)*

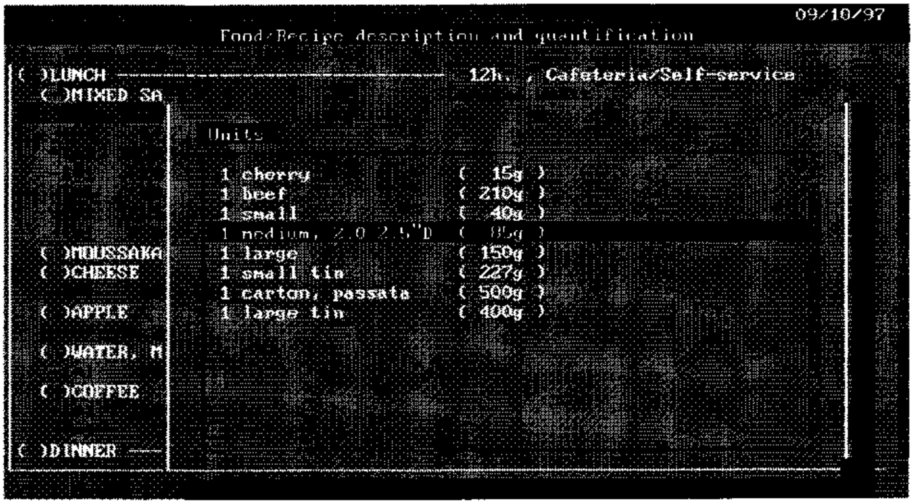
(c) Quantification

Six quantification methods are available to estimate portion sizes: Photos/shapes, household measurements, standard units, standard portions, gram/volume method, and "unknown" method. At the quantification step, the pre-selected and specific methods for estimating each food/ingredient or recipe are proposed on the screen (**Screen 9a**).

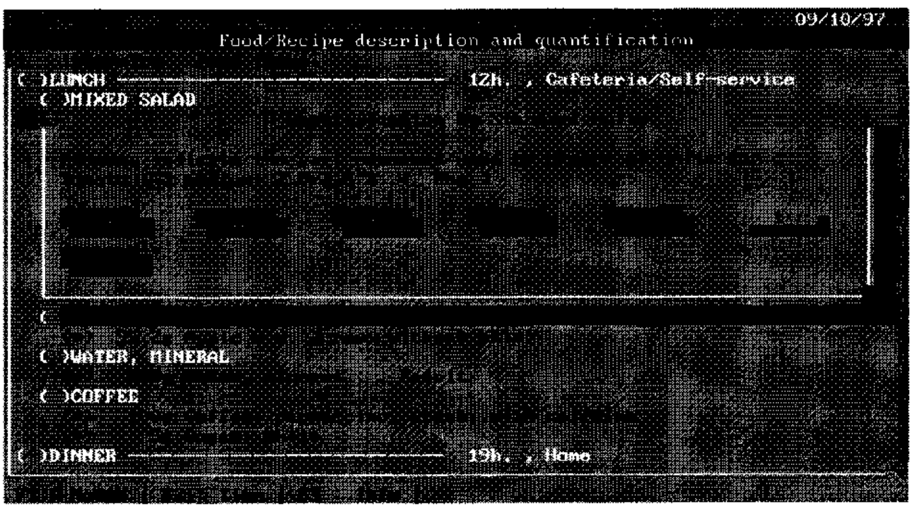


***Screen 9a** : Example of pre-selected quantification methods for « raw tomato »*

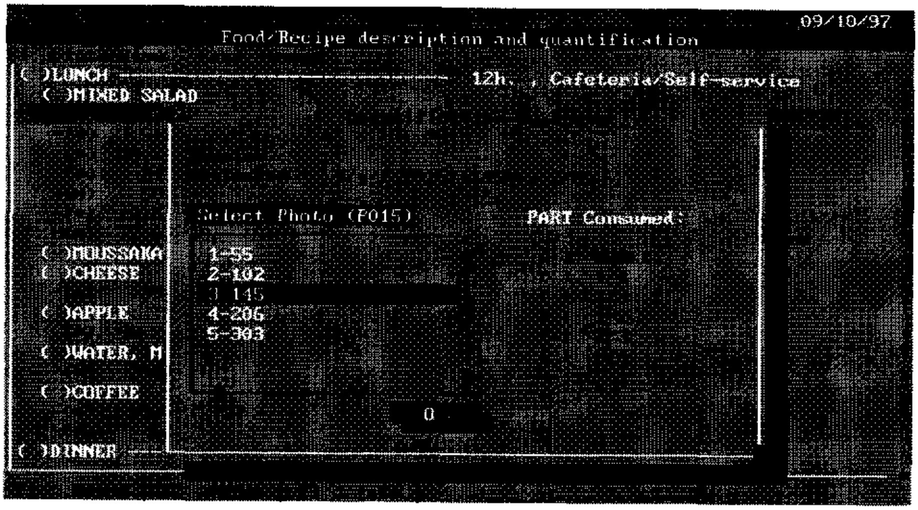
Depending on the quantification method selected, it is possible to indicate the number of units, or whether the whole or only a fraction of the portion was consumed (**Screens 9a, 9b, 9c** for standard units, and **Screens 9a, 9d** for photo). These pre-defined fractions are common to similar foods belonging to the same (sub-) group.



***Screen 9b** : Example of pre-selected standard units for « raw tomato ».*
The quantity in brackets is the unit weight



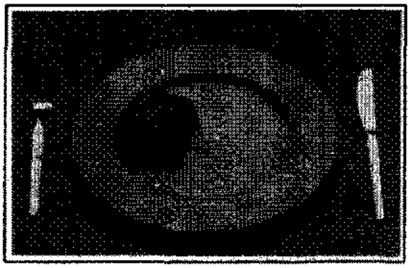
***Screen 9c** : Example of pre-selected standard unit fractions for « raw tomato »*



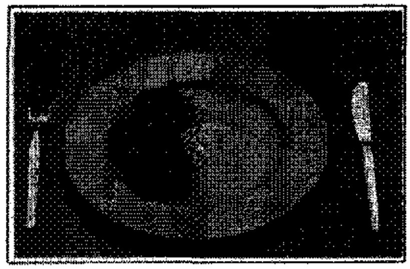
***Screen 9d** : Example of pre-selected photo for « raw tomato »*

For photos, an additional question is asked on whether or not a second helping was consumed. EPIC-SOFT is accompanied by a book of coloured photos to help the subjects estimate the portion sizes they actually consumed. This contains sets of photos of 94 foods and 46 recipes (20) (see example of a photo set: tomato) with 4 to 6 portion sizes in increasing size.

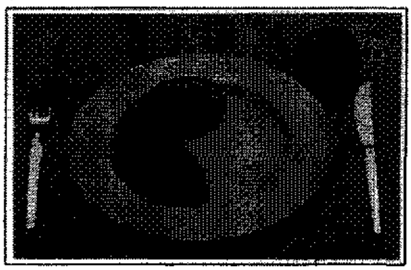
Tomatoes 15



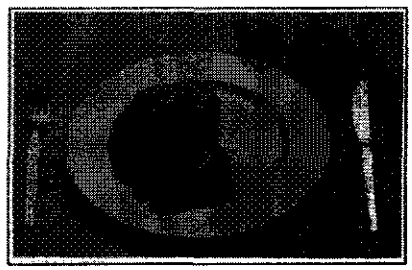
15 - 1



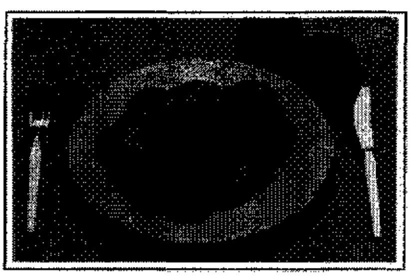
15 - 2



15 - 3



15 - 4



15 - 5

Example of the photo series of the EPIC-SOFT picture book: tomatoes

The difference in weight between two adjacent portions is at least 25% and their range was based on consumption data collected during the EPIC pilot phase or on national data. The same photo may be used to quantify several

foods with similar form and density. For practical reasons, some household measurements (HHMs) such as glasses, bowls, are available on pictures.
A ruler is used during the interview to check that the actual height and/or circumference correspond to the HHMs the subject actually used. The other HHMs are shown by the interviewer so that the subject can check the dimension of the selected measure. The ruler is also used when three standard units of different sizes (small, medium and large) are proposed to quantify certain foods (e.g. for apple, potato). Using the ruler, the interviewer indicates to the subject the circumference/length of the *medium* standard portion. The subject has then to say whether the size of the consumed portion was smaller, larger or about the same as the medium one. This approach, using an absolute scale as common reference, was adopted to standardize better the notion of small, medium, or large portions between subjects.

For bread it was decided to use two-dimensional models of bread slices because there are so many types of loaves, both regionally and internationally, that they could not all be photographed. In addition, it is easier to identify the real thickness and size of bread slices from shapes with the real form of bread than from pictures.

The "standard portions" method is relatively rarely used. They are available for foods in small quantities on the dish and not known by the subject (e.g. grated cheese on pasta; or onion, nut or ham in a mixed salad). The "gram/volume" method is systematically proposed for all foods. This method is used particularly for the ingredients of mixed recipes which are frequently reported in their raw form before any preparation and/or cooking (e.g. 500 g of potatoes), or when the subject knows the precise weight of the portion actually consumed (e.g. a 250 g packet of commercial French fries).

The "Unknown quantity" method is used either when the subject cannot estimate how much was consumed or when a quantification method is lacking. For example, when a standard unit weight of a commercial product (e.g. yogurt, biscuit) is missing in the EPIC-SOFT database, the type and number of standard units and any further information can be entered in the "note file" automatically prompted on the screen. The missing quantification is updated afterwards in the EPIC-SOFT files. The fat absorbed by a food during cooking is automatically calculated based on the food weight and according to the estimated amount of fat used for a given cooking method. If a subject is unable to estimate the quantity of fat, sauce, sweetener added on food or recipe, the system also automatically calculates a standard quantity adjusted for the weight of the food (or recipe) actually consumed.

Note that foods are always estimated "as consumed", whereas in mixed recipes the consumed quantity is calculated in three steps:

(1) Estimation of the individual "as consumed" recipe portion with quantification methods proposed by the system (e.g. photos of ready-to-eat moussaka);

(2) quantification of each ingredient of the mixed recipe. As for foods, common algorithms are used to adjust for raw-to-cooked, edible part and density, if the ingredients are not estimated in their "as consumed" form (i.e. quantification of the ingredients of moussaka: aubergines, tomatoes, meat, etc...); and

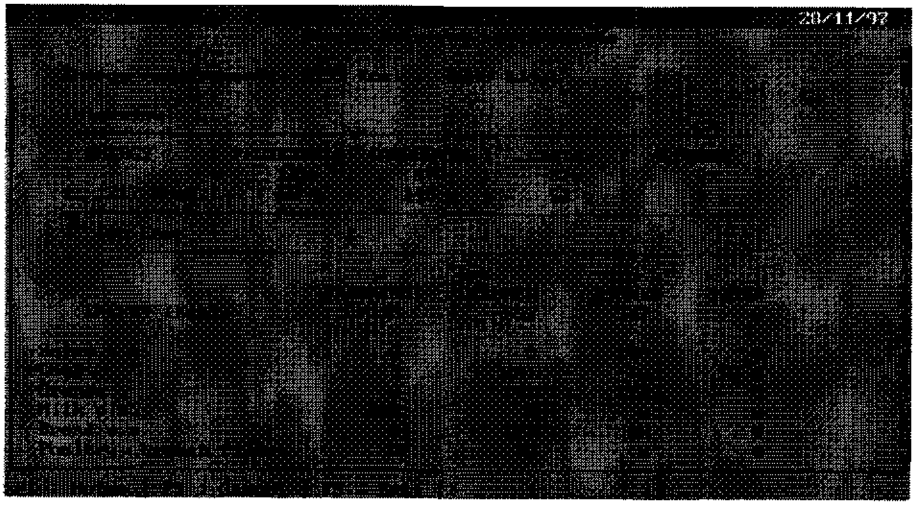
(3) automatic calculation of the "as consumed" weight of each ingredient based on the recipe portion actually consumed (i.e. relative weight of each moussaka ingredients actually consumed).

d) Check list

After the description and quantification of each food or recipe, a checklist containing foods which can easily be forgotten in a given QLI is automatically displayed on the screen to ensure that the subject does not forget any food. Particular attention is given to fat, sauce and sugar added to foods, as they are not reported as separate QLIs in the quick list.

Quality controls

Various quality controls are made at different stages of the interview procedure. Checks are made on outlier values either entered by the interviewer or calculated by the system (e.g. food quantity above a pre-defined maximum value). Before ending the interview, a screen summarizing the foods/recipes with missing quantities is displayed automatically and any incomplete information can be identified and corrected. Thereafter, a final control at the nutrient level is made by comparing the energy and macronutrient intakes estimated during the interview with standard energy requirements based on the subject's age, sex, weight and height. If the nutrient intake seems too high or too low, the interviewer has the possibility to correct for possible errors while the subject is still present (**Screen 10**).



***Screen 10** : Comparison between estimated and required energy and macronutrient intakes. Example of «too low energy intakes » to be checked by the interviewer*

At the end of the interview, questions are asked about the consumption of supplements or medication which may contain vitamins or minerals. A pre-defined list of locally available products is included in each country-specific EPIC-SOFT version. However, if a product has not been included, it can be added during the interview and described and quantified in terms of daily dose. The actual inclusion of new products in the database is done afterwards.

Maintenance/update of the EPIC-SOFT databases

Like any open-ended method, the computerized 24-hour diet recall method needs regular update of its databases in order to add new foods, recipes, or other information reported by the study subjects. To maintain a high level of control and standardization of EPIC-SOFT databases and to facilitate updating, it was decided that only one version will be available per country and that any modifications to the EPIC-SOFT files will be centralized at IARC. In addition, a four-step hierarchical structure involving the *interviewers, centre coordinators, country coordinator and international coordinator* (IARC) was set up to share the work involved in standardizing the collection, update, and export of the EPIC-SOFT 24-hour diet recalls and ensure controlled transfer of modifications required to the common or country-specific EPIC-SOFT files from the bottom level (interviewers) to the top (IARC).

Various tools and guidelines were developed at IARC to ensure that the above tasks are highly standardized. Common food or recipe update forms are used to fill in all the information needed to update a new food or recipe not available in the EPIC-SOFT databases (or any other missing information). The country coordinator sends these forms to IARC at regular intervals. Additional software (RECIPE manager) was developed to facilitate the update of the standard recipe files of the EPIC-SOFT versions.

Other software (EPIC-SOFT Recompute and Export programmes) was developed to recalculate automatically the individual food (or nutrient) intakes collected with the 24-hour diet recalls according to the latest modifications made to the EPIC-SOFT standard files (e.g. food standard unit; edible, density, raw-to-cooked standard coefficient files). Before performing statistical analyses, it is particularly important to ensure that the final food and nutrient intakes were estimated from the most up-to-date EPIC-SOFT version. This programme also checks for incompleteness in the 24-hour diet recalls by indicating precisely the missing items (e.g. missing food or recipe, missing coefficient, missing food portion). In addition, this software allows the 24-hour diet recall interview data to be exported in a common export file. This facilitates the storage, exchange and combined analyses of the EPIC calibration data.

EPIC-SOFT software and hardware requirements

EPIC-SOFT is written in Clipper. The system is PC IBM-compatible and needs about 2 Mbytes of RAM memory and at least 5 Mbytes of hard disk space. MS-DOS version 3.00 or later is required.

Conclusion

In this paper we presented the general structure and principal functions of a new software (EPIC-SOFT) developed to obtain standardized 24-hour diet recall interviews between the nine European countries involved in the EPIC project. It was decided to develop a computerized 24-hour diet recall interview as it was felt that this was the most appropriate methodological approach to standardize the interview procedure (21-22), and to minimize the potential sources of error associated with the 24-hour diet recall method in a project which was going to mobilize about 90 interviewers in 22 participating centres. Common rules for describing, quantifying and checking the 1500 to 2200 foods and 150 to 350 recipes per country were pre-entered in the system. In addition, different quality controls on the information reported were included in EPIC-SOFT to reduce the risk of misidentification and misquantification and permit corrections to be made while the subject is still present. Dietary intakes collected with EPIC-SOFT can be analyzed at the food group level according to the common EPIC-SOFT food classification system.

EPIC-SOFT standardizes the collection of food consumption intakes between countries. However, the nutrient databases implemented in the system are not yet standardized. These temporary food composition tables, derived from national values, are used during the interview to calculate energy and macronutrients and identify and correct gross over- or under-estimation while the subject is still present. A study of the comparability of food composition tables available in the EPIC countries shows that they are extremely heterogeneous (23). Further work is needed to reduce potential sources of systematic bias due to different definitions, sampling and analytical methods, number and precision of food items used in the different national food composition tables. It was therefore decided to develop standardized nutrient databases, which can be used for the combined analyses of the EPIC 24-hour diet recall data and the baseline questionnaires.

The standardized EPIC-SOFT programme exists in 9 European versions and it takes about 30-35 minutes to conduct a 24-hour diet recall interview, whatever the country. Additional time, depending on the initial completeness of the country EPIC-SOFT versions, is requested to add any new foods or recipes (or any other missing information) reported during the interviews to the EPIC-SOFT databases. It is planned to conduct a series of analyses on about 32,000 interviews to determine the quality and level of standardization of the 24-hour diet recalls collected with the nine EPIC-SOFT versions. In addition, the EPIC-SOFT estimates will be validated (correlated) against some biochemical markers such as total urine nitrogen (24), potassium and certain serum components (e.g. carotenoids, fatty acid profile) using biological samples collected and stored within EPIC and newly developed HPLC and GC chemical analysis techniques for single carotenoids and fatty acids (25-26). The results of these statistical analyses, will furnish an independent appraisal of the ability of EPIC-SOFT to provide accurate and standardized 24-hour diet recall estimates. So far, EPIC-SOFT is the only known available software which was designed to provide comparable 24-hour diet recall data across several European countries. This programme is likely to be of interest for other nutritional (multi-centre) studies within or outside EPIC. In the near future, it is planned to develop additional software programmes to make EPIC-SOFT more independent of the EPIC study and its logistics so that it can be used by other researchers.

Acknowledgements

The authors wish to thank everyone who participated actively in the development of the local EPIC-SOFT versions and in the field work of the calibration studies as well as those in charge of each centre's technical assistance: Johanne Amoyel and Maryvonne Niravong (France), Stefanie Bohlscheid-Thomas, Susanne Voss and Heiner Boeing (Germany), Georga Kornilia (Greece), Manuela Bellegotti and Sabina Sieri (Italy), Janneke Verloop and Ester Goddijn (The Netherlands), Ulrika Pettersson (Sweden), Alison McTaggart (UK), and Corinne Casagrande and Bertrand Hémon (IARC). The authors would like also to thank Dr. Christine Friedenreich of the Division of Epidemiology, Prevention and Screening, Calgary (Canada) for her helpful comments. This work was part of the EPIC project supported by the Europe Against Cancer Programme of the European Union.

References

1. E. Riboli and R. Kaaks, The EPIC project : Rationale and study design. Int. J. Epidemiol. 26 (suppl. 1) (1997) S6-S14.
2. K. Overvad, A. Tjønneland, J. Haraldsdóttir, S. Bang, M. Ewertz and O. Møller-Jensen, Development of a semi-quantitative food frequency questionnaire to assess food, energy and nutrient intake in Denmark. Int. J. Epidemiol. 20 (1991) 906-912.
3. S.A. Bingham, C. Gill, A. Welch, K. Day, A. Cassidy, K.T. Khaw, M.I. Sneyd, K.T.J. Key, I. Roe and N.E. Day, Comparison of dietary assessment methods in nutritional epidemiology : weighted records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. Br. J. Nutr. 72 (1994) 619-643.
4. M.C. Ocké, H.B. Bueno de Mesquita, H.E. Goddijn, A. Jansen, M.A. Pols, W.A. van Staveren and D. Kromhout, The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire and validity and reproducibility for food groups. Int. J. Epidemiol. 26 (suppl. 1) (1997) S37-S48.
5. S. Bohlscheid-Thomas, I. Hoting, H. Boeing and J. Wahrendorf, The reproducibility and relative validity of a food frequency questionnaire newly developed for the German part of the EPIC project. I. Food group intake. Int. J. Epidemiol. 26 (suppl. 1) (1997) S59-S70.
6. The EPIC group of Spain, Validity and reproducibility of a diet history questionnaire in Spain. I. Foods. Int. J. Epidemiol. 26 (suppl. 1) (1997) S91-S99.
7. K. Katsouyanni, E.B. Rimm, C. Gnardellis, D. Trichopoulos, E. Polychronopoulos and A. Trichopoulou, Reproducibility and validity of an extensive semi-quantitative food frequency questionnaire using dietary records and biochemical markers among Greek school teachers. Int. J. Epidemiol. 26 (suppl. 1) (1997) S118-S127
8. M.J. van Liere, F. Lucas, F. Clavel, N. Slimani and S. Villeminot, The validity and reproducibility of a French dietary history questionnaire. Int. J. Epidemiol. 26 (suppl. 1) (1997) S128-S136
9. P. Pisani, P. Vineis, V. Krogh and F. Berrino, Validity and reproducibility of the dietary questionnaire for the Italian component of EPIC. Int. J. Epidemiol. 26 (suppl. 1) (1997) S152-S160.

10. E. Riboli, S. Elmståhl, R. Saracci, B. Gullberg and F. Lindgärde, The Malmö Food Study : Validity of two dietary assessment methods for measuring Nutrient intakes. Int. J. Epidemiol. 26 (suppl. 1) (1997) S161-S171.
11. M. Plummer, D. Clayton and R. Kaaks, Calibration in multi-centre cohort studies. . Int. J. Epidemiol. 23 (1994) 419-426.
12. R. Kaaks, E. Riboli and W.A. van Staveren, Calibration of dietary intake measurements in prospective cohort studies. Am. J. Epidemiol. 142 (1995) 548-556.
13. R. Kaaks and E. Riboli, Validation and calibration of dietary intake measurements in the EPIC project. Int. J. Epidemiol., 1997, Vol 26. Suppl. 1: S15-S25.
14. M. Buzzard and Y.A. Sievert, Research priorities and recommendations for dietary assessment methodology. Am. J. Clin. Nutr. 59 (Suppl.) (1995) 275S-280S.
15. E. Callmer, U. Hagman, J. Haraldsdóttir, E. Bjørge Løken, R. Seppänen and Trygg K, Standardization of 24-hour recall. Vår Föda 38 (Suppl. 4) (1986) 201-298.
16. S.A. Bingham, The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. Nutrition Abstracts and Reviews (series A) 571 (1987) 705-742.
17. M.E. Cameron and W.A. van Staveren, Manual on Methodology for Food Consumption Studies. Oxford University Press, Oxford (1988).
18. J.C. Witschi. Short-term dietary recall and recording methods. In: W. Willett (ed.) Nutritional Epidemiology, pp. 52-68. Oxford University Press, New York (1990).
19. N. Slimani, W.A. van Staveren and E. Riboli. Approach towards the standardization of the 24-hour diet recall method used for calibration of dietary intake measurements in EPIC (in preparation).
20. A.L. van Kappel, J. Amoyel, N. Slimani, B. Vozar and E. Riboli, EPIC-SOFT picture book for estimation of food portion sizes. International Agency For Reasearch On Cancer, Lyon, France (Internal Report, 1995).
21. B. Dennis, N. Ernst, M. Hjortland, J. Tillotson and V. Grambsch, The NHLBI nutrition data system. J. Am. Diet. Assoc. 77 (1980) 641-646.
22. D. Feskanich, B.H. Sielaff, K. Chong and M. Buzzard, Computerized collection and analysis of dietary intake information. Comput. Methods Programs Biomed. 31 (1989) 47-57.
23. G. Deharveng, R. Charrondière, N. Slimani, E. Riboli and D.A.T. Southgate, Comparison of nutirents in the food composition tables available in the nine European countries participating in EPIC. Eur. J. Clin. Nutr. (in press).
24. S.A. Bingham and J.H. Cumming, Urine nitrogen as an independent validatory measure of dietary intake. Am. J. Clin. Nutr. 42 (1985) 1276-1289.
25. J.P. Steghens, A.L. Van Kappel, E. Riboli E and C. Collombel, Simultaneous measurement of seven carotenoids, retinol and alpha-tocopherol in serum by High-performance liquid chromatography. J. Chromatograph. B 6941(997) 71-81.

26. V. Chajès, P. Toniolo, A.L. Van Kappel, A. Zeleniuch-Jacquotte and E. Riboli, Intake of specific foods is associated with serum phospholipid fatty acid composition among women in New York City (submitted).

Chapter 4

Standardization of food composition databases for the European Prospective Investigation into Cancer and Nutrition (EPIC): general theoretical concept

Nadia Slimani, Ute Ruth Charrondière, Wija van Staveren, Elio Riboli

J Food Comp Anal 2000; 13: 567-584

Abstract

EPIC is a prospective cohort study on diet and cancer involving 480,000 subjects from 9 European countries. In order to establish the relationship between nutrient intakes and disease, standardized food composition databases are needed. In the absence of already existing comparable European nutrient database(s), an *ad hoc* approach was developed to standardize the EPIC databases.

New matrixes were built using information as collected from the EPIC study subjects in order to overcome the difficulty of reducing the systematic differences between food lists in national databases. In the EPIC databases, the food lists, *vertical axes*, are based on information derived from standardized computerized 24-hour diet recalls collected from 35,000 subjects. The criteria of selection and level of detail on foods reported in the EPIC databases are therefore highly standardized between countries. In addition, reported recipes are systematically broken down into ingredients to optimize the comparison between countries. For the nutrient list, *horizontal axes*, the number, mode of expression, definition, unit and methods of analysis are fixed. The compilation of nutrients, *nutrient value cells*, is carried out using standardized sources of nutrient data or algorithms. Depending on whether the food is *common* or *country-specific*, the same or country-specific source of values will be used. This approach addresses some methodological issues, which may have implications for future priorities and development of food composition tables.

Introduction

There is growing interest in large (prospective) multi-centre studies on diet and chronic diseases. These studies offer an optimal setting for substantially improving our knowledge of the complex relationship existing between diet and diseases by increasing both the scale and the geographical heterogeneity of the exposure (*diet*) and the incidence of the diseases (e.g. *cancer, cardiovascular diseases*) (1). These new study designs, however, raised a major new methodological issue related to the comparability, at the food and nutrient levels, of dietary measurements collected from populations differing in eating habits and geographical and ethnic origin. An increasing amount of literature is being published on the methodological and statistical approaches to improve the standardization and pooled analysis of dietary measurements collected in large multi-centre studies (2-5). A number of different international and regional initiatives to improve the harmonization of analytical laboratory methods, definition and mode of expression of foods or nutrients have been promoted during the last fifteen years (6-11). The INFOODS organisation initiated this work by proposing systems to describe food and food components to facilitate the international food data interchange (7,8,12), and this was followed by other regional initiatives and proposals, particularly from the EUROFOODS COST99 and NORFOODS groups, to improve data interchange (13-15). However despite these major contributions, mainly oriented toward the harmonization of data management and interchange, there are still no standardized European food composition databases for international use available.

Food composition databases are known to be a major source of error in the estimation of nutrient intakes (16-18). Several studies that compared different computerized nutrient databases, have shown significant differences between the softwares for a number of tested nutrients (19-23). These were methodological exercises which compared the nutrient intakes calculated by several computerized nutrient databases, when the potential sources of error attributable to food consumption data were controlled by using the same food data set. These differences were observed between softwares from the same country and may be much larger in international multi-centre studies where food composition databases available in the different participating countries differ substantially in form, content and relative quality (24). By introducing systematic artefactual differences in the estimates, systematic (or random) errors in food composition databases used in international contexts can then bias the comparison of nutrient intakes across countries and the subsequent interpretation of the diet-disease relationships. In this paper we will use the term *food composition database* (FCDB) for nutrient data available on paper or on hard disk support.

In the absence of an already existing standardized European FCDB for nutritional epidemiology, the European Prospective Investigation into Cancer and Nutrition (EPIC) developed an *ad hoc* concept to improve the comparability of the FCDBs to be used for pooled analyses of dietary data collected from the nine Northern, Central and Southern European countries involved in the project (Denmark, France, Germany, Greece, Italy, Netherlands, Spain, Sweden and United Kingdom). This paper will present the overall theoretical concept developed to standardize FCDBs between the EPIC countries, using the unique setting and resources offered by this project. The actual practical development of these standardized FCDBs is ongoing and will be reported in further papers.

Presentation of the European Prospective Investigation into Cancer and Nutrition (EPIC)

EPIC is a multi-centre prospective study aimed at investigating the relationship between nutrition and various lifestyle factors and the etiology of cancer and other chronic diseases (25). The project involves 22 regional centres located in 9 European countries representing a total cohort of about 480,000 subjects. Dietary information was collected using validated country-specific dietary history questionnaires containing between 150 and 300 food items to estimate the individual usual food intakes (26-28). In addition, a second dietary method based on a single open-ended 24-hour diet recall interview, collected from a representative sample of each of the nine cohorts (~ 35,000 subjects), was used as common reference calibration method to correct both for systematic errors associated with the baseline questionnaires and for attenuation bias in the relative risk when measuring the association between diet and disease (2-3). The 24-hour diet recalls were collected using a highly standardized user-friendly computerized system (EPIC-SOFT) containing between 1500 to 2000 foods and 200-350 mixed recipes depending on the countries (5). Standardized rules were implemented in the software to *identify*, *describe*, *quantify*, and *check* reported foods and mixed recipes similarly across the countries involved in EPIC. In addition, a blood sample was collected from 360,000 study subjects using standard protocols, and stored in liquid nitrogen (-196°C). These specimens will be used within EPIC to investigate a series of markers of nutritional, metabolic and hormonal status and genetic susceptibility and exposure to genotoxic and non-genotoxic components. Urine samples were also collected from about 2000 individuals in order to validate the 24-hour diet recalls against independent biological measurements such as urine nitrogen and potassium (29-30). Specific questions on consumption of supplements of vitamins/minerals or other compounds were asked systematically at baseline before blood collection and during the 24-hour diet recall interviews.

FCDBs are essential for EPIC in order to estimate nutrient or non-nutrient intakes collected from three different sources: *Individual dietary questionnaire*, *24-hour diet recalls* and *supplement intakes*. FCDBs are primarily needed for nutritional studies investigating the relation between dietary and supplement intake and biochemical measurements. In addition, biomarkers of diet and dietary measurements can be combined for the etiological investigation of cancer and cardiovascular disease. New statistical approaches based on latent variable models have been developed (31), but accurate measurement of nutrient intake is required. Finally, FCDBs will be essential to translate research findings in terms of public health recommendations on foods or food patterns, for disease prevention.

Needs for Food composition databases in large multi-centre epidemiological studies such as EPIC

The requirements of FCDBs in nutritional epidemiology depend directly on the methodological and etiological specifications: *nutritional objectives*, *study design*, *study population* and *dietary measurements*. These needs are similar to those of other users, but are amplified in large international epidemiological studies such as EPIC (32-33).

Needs for tailored and complete databases

Nutritional epidemiology requires the FCDBs to be specifically tailored to the variety and the detail of reported food consumption in different populations according to the dietary method(s) used. In EPIC two dietary methods are used: 1) The *baseline questionnaire*s contain 150-300 generic items (e.g. meat, citrus fruits, cakes, etc.) and it is used to estimate usual long-term individual intakes. 2) The *24-hour diet recalls*, which measure all foods and beverages consumed the previous day, provide detailed information on several thousand specific foods including characteristics such as method of cooking and preservation, meat cut and food product brand name.

It is practically impossible to have FCDBs with all the foods reported by the study subjects. Ideally, reference guidelines should be available for deciding alternative foods of nutritional similarity and on algorithms to be applied for defining standard mixed recipes to estimate mean nutrient values of mixed dishes or average weighted food items. A high proportion of missing values in FCDBs is another source of incompleteness in the databases which may introduce random or systematic under-estimation. In this case, it is better to replace a missing value by an estimated or (weighted) mean value, than to have no value at all (34).

The nutrients reported in FCDBs should cover the nutritional objectives of the study. EPIC has been set up to investigate the relationship between diet, cancer and other chronic diseases prevalent in western countries. For each disease there are specific etiological hypotheses on one or more nutrients or other components vectored by foods which need to be investigated, e.g. the role of some vitamins in carcinogenesis as antioxidants (carotenoids, vitamins C and E) and methyl donors (folates, B6), or the relationship between total fat and specific fatty acids or fractions and cancers of the breast, prostate and colorectum. Ideally, FCDBs should contain all the components of etiologic, descriptive or methodological interest for researchers.

Need for valid databases

The nutrient values reported in FCDBs should be valid and representative of local consumptions. The validity will depend upon the method of sampling and chemical analysis, the representativeness of values, and the accuracy and standardization of laboratory quality controls (9). With the increasing use of biological markers in epidemiological studies and the increasing consumption of vitamin/mineral supplements, further information is needed on the bioavailability of these nutrients to improve the estimation of their absolute intakes. This area of research on nutrient bioavailability is being developed particularly on vitamins and minerals such as calcium, iron, B vitamins, but needs to be extended to other nutrients in the future (35-36).

Need for consistent databases

FCDBs used in multi-centre studies should also be consistent within and between countries. At the food level, there should be a certain homogeneity in the mode of expression, level of detail of foods belonging to the same groups (e.g. meat, vegetables). At the nutrient level, the same units, definition and methods of analysis providing comparable results should be used consistently in the tables.

Need for comparable and standardized databases

This is required particularly for international multi-centre studies where different FCDBs are used. The standardisation of FCDBs should prevent or minimize bias (systematic errors) which could affect the pooling of data collected in different centres or countries. Two independent FCDBs might provide equally good nutrient estimates on a relative scale within each population, but may introduce systematic differences at an absolute level if they are used together in pooled data analyses. For example, two tables could both have analytical values obtained from the same chemical method, but provide systematically different results if one expressed the results for raw foods and the other for foods as consumed. The values expressed for raw foods will systematically over-estimate intake, particularly of thermolabile vitamins (e.g, vitamins B and C), if no correction or adjustment is performed.

Need to maintain databases over time

The maintenance of FCDBs over time is particularly important in prospective studies where study populations are followed for a relatively long period of time. FCDBs should be updated regularly to take into account changes, i.e. availability of more reliable data or addition of nutrients or components of new research interest.

Need for documented databases

FCDBs should provide full information on the general and specific conventions applied in the compilation. At the food level, adequate nomenclature, identification and classification of foods and description of food samples is needed. At the nutrient level, the food components reported should be identified, and information should be provided to enable users to judge the origin, precision and representativeness of the nutrient values.

Need for a database management system (DBMS)

It is essential for food composition data to be available in a computerized form, to facilitate the management of large datasets. A database management system (DBMS) is necessary as a support for the overall data processing, i.e. retrieval, changes, calculations, etc.

Comparability of food composition databases available in the countries participating in EPIC

Before deciding on a general approach for standardizing FCDBs between the countries involved in EPIC, an extensive review was made to compare FCDBs available in the nine European countries participating in the project (24). The main objective was to evaluate the nature and magnitude of systematic differences between tables from the point of view of *availability, definition, mode of expression and analytical methods* for nutrients of interests for EPIC. Despite improvements in comparability between European tables promoted by international programmes such as Eurofoods-Enfant or Infoods, there are still large differences between national food composition and nutrient databases at different levels:

At the food level, the number of food items reported varies enormously between national tables from several hundreds to several thousands. In addition, the level of detail reported for these foods (e.g. information on cooking method, method of preservation or physical state, or different meat cuts) varies substantially within and between tables (unpublished data).
Table 1 summarizes the magnitude of differences in the number and level of detail of descriptions reported for beef and veal. It shows that the information on cooked foods is not reported systematically and that the numbers of specific meat cuts and cooking method combinations vary enormously across tables. These differences may introduce systematic errors in the estimation of, in particular, total fat, fatty acids, cholesterol and thermo-labile vitamin intakes, when food consumption data are matched with FCDBs.

Table 1 : Example of information reported on foods in the principal FCDBs available in the nine countries involved in EPIC: *beef, veal* (*)

	Beef		Veal	
	Number of items	Description	Number of items	Description
Denmark Levnedmiddeltabeller (1996)	24	*21 meat cuts with 24 raw only, sometimes different fat contents*	0	-
France CIQUAL (1995)	17	*9 meat cuts with raw (5), and various cooking methods (12)*	7	*6 meat cuts with raw (3), and various cooking methods (3)*
Germany BLS (1996)	134	*29 meat cuts with raw or deep frozen (95), cooked n.s. (39)*	147	*24 meat cuts with raw or deep frozen (107), cooked n.s. (40)*
Greece Greek food composition table (1992)	27	*11 meat cuts with raw (11), and various cooking methods (16). Based on McCance data*	4	*3 meat cuts with raw (1), various cooking methods (3). Based on McCance data*
Italy Banca dati italiana di composizione degli alimenti per studi epidemiologici (1998)	6	*1 (young beef) raw with 6 different fat contents + 1 canned*	2	*1 raw meat cut with 2 different fat contents*
Netherlands NEVO (1996)	36	*18 meat cuts with raw (18) and cooked n.s. (18)*	7	*4 meat cuts with raw (4), and cooked n.s. (3)*
Spain Tablas de composicion de alimentos espanoles (1997)	0	-	2	*2 raw meat cuts:*
Sweden Livsmedelstabel (1996)	12	*7 meat cuts with raw (9) – with or without bone; and canned (1), boiled (1), roasted (1)*	2	*2 raw meat cuts:*
United Kingdom McCance and Widdowson's - meat, poultry and game supplement (1995)	91	*15 meat cuts with raw (31)– also different fat contents – or with bone; and various cooking methods (60)*	4	*2 cuts with raw (2), fried (1), stewed (1)*

(*) : Excluding organs and offal

At the nutrient level, nutrients are comparable to varying degrees across national FCDBs. Deharveng et al. (24) distinguished three main groups of nutrients. The first set contains nutrients with comparable values across tables but which may slightly differ in the definitions or analytical methods used (nitrogen, lactose, alcohol, fat, fatty acids, retinol, Vitamin D, tocopherols, thiamin, riboflavin, vitamin B6 and B12, calcium, iron and potassium). The second set are nutrients which are not readily comparable due to different modes of calculation or of expression but which could easily be converted to comparable values (energy, protein, carbohydrates, starch, sugar, carotene, vitamin A and vitamin E). The last group are nutrients showing different values due to method or definition used (e.g. folates, fibre). Data on these nutrients should be used with particular caution in any multi-centre comparisons.

A further problem is the lack of documentation on the sources of nutrient values and related chemical methods, particularly when the values are derived both from original analyses and from borrowed data and there is no way to retrieve the information and check the consistency between the different sources. To improve comparability between tables it would also be useful to have information on food sampling, number of samples, means, and coefficients of variation or minima/maxima in order to assess the representativeness and the range of (natural) variation of selected values. Unfortunately this is so far available only in a restricted number of FCDBs.

In conclusion, a critical evaluation of the existing FCDBs and nutrient databases available in the nine European countries involved in EPIC indicated that further work is needed to standardize FCDBs in order to minimize the systematic and random errors which may affect nutrient intake measurements, particularly in pooled data analyses on diet and cancer risk (24).

Overall theoretical strategy developed to standardize food composition databases within EPIC

FCDBs are bi-dimensional matrixes, where the *vertical axis* corresponds to the list of food items and the *horizontal axis* to the list of nutrients or other components. The *cells of the matrix* correspond to the nutrient values of specific food-nutrient pairs. A third *virtual* dimension of the standardization of food composition databases is an adequate documentation of food, nutrient and component value, but which, however, will not be discussed in the paper. As shown in the previous section (Table 1), one of the prime differences existing between FCDBs is the number, level of detail, and type of food items reported.

These differences in the vertical axes of FCDBs explain to a large extent the difficulty of comparing further nutrient values across tables, and the failure of attempts to standardize them on the basis of the existing FCDBs. In addition, the discrepancy between the information on foods available in FCDBs and information on diet spontaneously reported by subjects constitutes the principal source of coding errors when food consumption data and FCDBs are matched together, and subsequent mis-estimation of nutrient intakes (20, 37).

The approach proposed to standardize FCDBs within EPIC, and overcome these major problems, consisted first in building up new matrixes, starting with the preparation of food lists (vertical axis), from the 35,000 detailed 24-hour diet recalls collected by means of a standardized computerized programme (EPIC-SOFT), rather than using the food lists available in the national FCDBs themselves. This was achieved by using indirectly the approach for standardizing food descriptions implemented in EPIC-SOFT to ensure that the same (or similar) foods across countries are described at equal levels of detail (38).

Standardization of the vertical axis of the FCDB matrixes (food lists)

In EPIC-SOFT, foods are described using two complementary *implicit and explicit food description approaches*. The *implicit descriptive approach* involved including in the nine EPIC-SOFT versions a food list containing 1500-2200 items classified according to a common classification system (38). Depending on the food (sub) groups, common criteria were defined to name the foods similarly, using both the local language, English and taxonomic names added afterwards, when relevant. For foods common or similar between countries, such as fruits, vegetables, fish and meat, only the generic names were reported in the food lists (e.g. orange, tomato, trout, etc). The *explicit descriptive approach* consisted in attaching to these foods an additional series of pre-defined questions asked during the interviews to obtain more details on how the foods were prepared or consumed specifically by the subjects. This explicit descriptive approach used the concept of *"facets"* and *"descriptors"* derived from the LANGUAL multi-factorial coding system initially developed to describe technological and toxicological food characteristics (39, 40). The *"facets"* (e.g. cooking method and preservation method) were used in EPIC-SOFT as systematic questions to describe similar foods. The "descriptors", which are the terms associated with each facet, were used as pre-defined potential answers to be given by the subjects (e.g. boiled, fried for the "cooking method" facet).

This system implemented in EPIC-SOFT contains 16 "facets" each having between 3 to 33 descriptors with common codes and definitions between countries. Depending on the food (sub-) group and the specific food item, different facet strings were used to describe foods. For salmon, for example, which is included in the nine food lists under the same food (sub-) group, the same series of questions (facets) about preservation method, physical state and cooking method containing the same descriptors were automatically prompted by the software and asked to describe it as finally consumed by the subjects (**Figure 1**). Complex multi-ingredient foods (e.g. cakes, breads, and sauces), which are more heterogeneous between countries, were treated with the same global approach as foods similar between countries. However, for these foods the implicit description was used more intensively, by introducing directly the product or brandname in the food names. The facet/descriptor approach was mainly used to discriminate between the same generic names by asking further questions about, for example, brandname for margarine (**Figure 2**). To improve the comparability of dietary data between countries and minimize mis-classification errors, mixed recipes were broken down into ingredients (41, 38). Each of the ingredients was then described by the same facet/descriptor approach presented above for single foods.

Figure 1 : Example of the food description in EPIC-SOFT: Salmon.

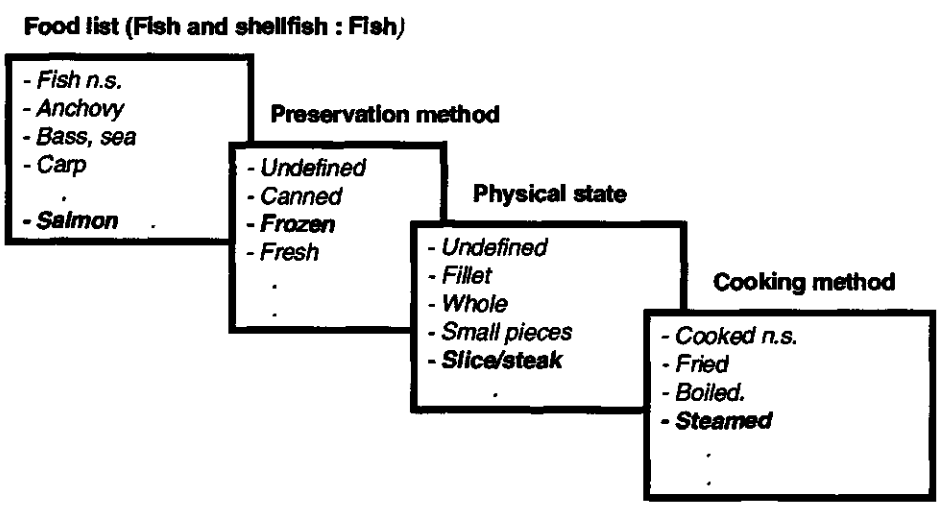
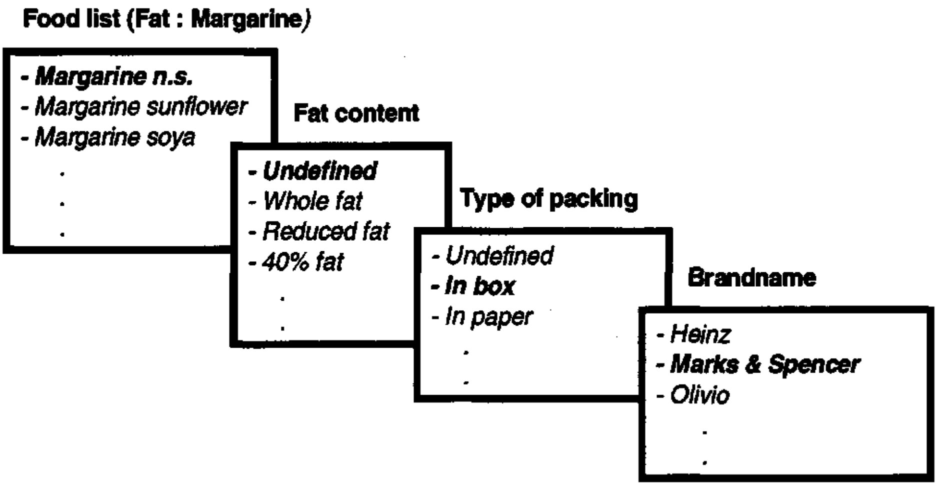
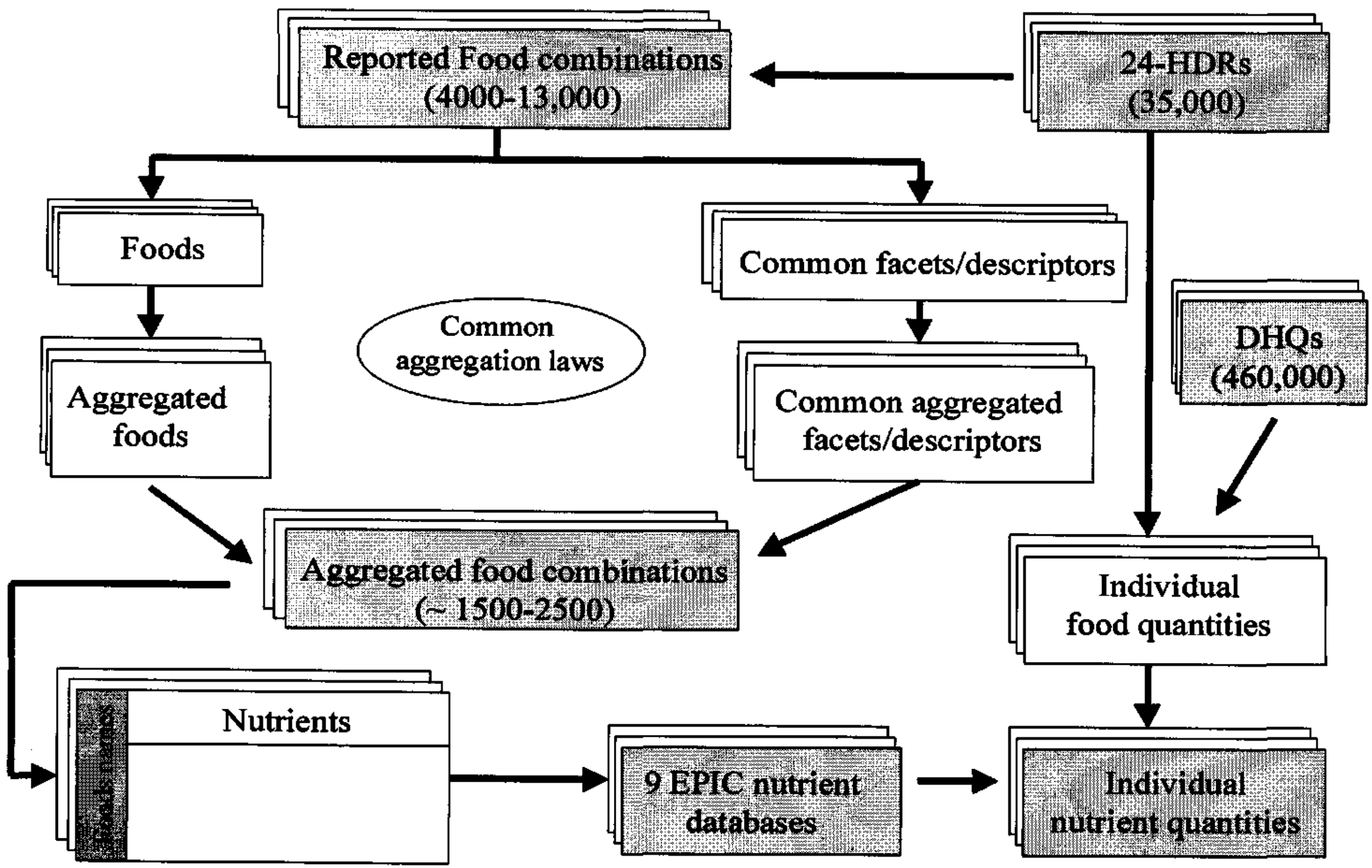


Figure 2 : Example of the food description in EPIC-SOFT: Margarine.



This combination of the food list (*implicit description*) and the facet/descriptor system (*explicit description*) produced a series of several thousand possible combinations corresponding to the complete identification and description of foods as consumed by the subjects, and highly comparable between countries for the naming, level of detail, and facets/descriptor codes. However, as it is impossible in practice to find specific nutrient values for each of the thousands of foods reported, a further major methodological step in the definition of the vertical axis was to aggregate common food combinations according to both the nutritional objectives and the information currently available in national food or nutrient databases or literature. Between 4000 and 13,000 different food combinations (raw data) were reported per country. These will be aggregated to obtain about 1500 to 2500 foods per country as entries in the vertical axis of the EPIC FCDBs. This procedure will be done semi-automatically by specifically developed software programmes. These aggregation rules programmes will apply to food lists and facets/descriptors common between countries to obtain restricted lists which will be linked together again, and used as entry points in the vertical axes of FCDBs (**Figure 3**). The standardised EPIC FCDBs will then be used to estimate the nutrient intakes derived both from the 24-hour diet recalls and from the individual baseline questionnaires (DHQ) collected from the entire cohorts. Only the latter will be used for etiological analyses, after being calibrated by means of 24-HDR mean estimates.

Figure 3 : Overall Flow of the standardization of the vertical axis (food lists) of the food composition matrixes



This (semi-) automatic aggregation approach provides the flexibility to reconsider the aggregation rules according to the food components of interest using the same original set of dietary data. It leaves the possibility to develop subsequently a series of databases on, for example, phytohormones, polyphenols, contaminants or pesticides which require different criteria of aggregation of the same original food combination lists.

Standardization of the horizontal axis of the FCDB matrixes (nutrient list)

In order to standardize the horizontal axis of the FCDB the nutrients to be included in the FCDBs must be defined. Forty nutrients and up to 20-40 fatty acids were selected for inclusion in the EPIC FCDBs (**Table 2**). Additional non-nutrients of recent major etiological interest such as *phyto-estrogenes* were also tentatively included. This list of nutrients, which could be extended in the future, was drawn up and will be compiled first in practice, according to a number of criteria such as:

a. Scientific interest for cancer and other chronic diseases to be investigated in EPIC.
b. Nutrients of interest for cross-sectional descriptive studies or studies comparing dietary and biochemical measurements.
c. Accuracy and comparability of nutritional data available in current national FCDBs or literature.
d. Confidence regarding nutrient intakes estimated from dietary assessments (reason to exclude sodium) and/or for which the status estimated using biological markers will be more reliable for ranking subjects according to blood concentration levels (selenium).

Table 2 : Non exhaustive list of food components proposed for inclusion in the first EPIC FCDBs

[Energy]	Cholesterol	Free Retinol	Vitamin C	Potassium
Water	Total	4 carotenes	Thiamin	Calcium
Total N	Carbohydrates	Vitamin D	Riboflavin	Iron
Protein	Oligosaccharides	Total vit. E	Vitamin B6	
Total fat	Starch	Individual	Folate	Phyto-estrogens
Fatty acids:	Lactose	tocopherols&	Vitamin B12	(Lignans,
fractions &	Alcohol	tocotrienols		Isoflavonoids)
individual	Dietary fibre			Polyphenols

For each of these nutrients, the *definition*, *mode of expression*, *units* will be harmonized between the potential sources to be used for the compilation. In addition, the same *chemical method of analysis* or those providing comparable results will be retained as reference methods for the compilation. Based on a critical evaluation of the information available in the national FCDBs carried out in collaboration with an expert food chemist (Professor D.A.T. Southgate), the following rules of standardization were defined for each specific nutrient (**Table 3**): For example, all energy values will be recalculated using factors 4, 9, 3.75, 7 respectively for protein, fat, carbohydrates, and alcohol (Kcal/g). The values given for total available carbohydrates will be recalculated to "as monosaccharides" using conversion factors. For dietary fibre values, there is a global problem, as the UK tables are the only ones which indicate non-starch polysaccharide (NSP) values obtained by the Englyst method, whereas all the others indicate total dietary fibre (TDF) values obtained through AOAC which are incompatible with Englyst NSP values. Englyst NSP has been chosen as the EPIC reference method because the Englyst NSP values are more consistent than AOAC values which are obtained by an enzymatic-gravimetric method and need to be adjusted for non dietary fibre components left in the residue. On the other hand, the Englyst NSP method measures precisely the sum of constituent sugars. Dietary fibre values can be regarded as compatible when necessary for fruits and vegetables (which have low lignin content and no resistant starch) as the two methods give very similar results. The main difficulties arise where foods are country-specific, such as many types of bread, rye flours, etc., but for which no NSP data are available. In these cases, the regression equation of Mongeau and Brassard (42) will be used to compute the predicted NSP values. Both TDF and NSP measurements will be reported in the EPIC tables.

Table 3 : Standardized reference definition, mode of expression, unit and chemical method of analysis of each selected nutrient

Nutrients (*)	Comparability of nutrient values	Appropriate reference analytical methods (s)	Reference definition
[Energy (kcal)]	To be calculated	-	4 protein,7 alcohol,9 fat, 3.75 available CHO
[Energy (kJ)]		-	17 protein,29 alcohol,37 fat, 16 available CHO
Water (g)	Comparable	Various drying methods	H2O
Total N (g)	To be calculated by FAO/WHO factors of 1973	Kjeldahl	Nitrogen
Protein (g)		-	Total N – non-protein N times FAO/WHO factors (1973)
Fat (g)	Comparability depends on methods	Acid hydrolysis+ mixed solvent extraction. Rose-Gottlieb values acceptable for dairy products. Soxhlet not desirable for cereals.	Triglycerides, phospholipids, sterols and related compounds
Fatty acids: fractions and individual (g)	All Methods compatible	Folch-type or Bligh and Dyer method + GLC	Latin names; chemical names or formulas
Cholesterol (g)	Globally comparable	GC, GLC or calculated	Cholesterol
Alcohol (g)	Conversion if %vol. or g/ml.	Distillation	Ethanol
Total Carbohydrates CHO (as such) (g)	Compatibility depends on modes of expression, methods (sum of analysed fraction vs. "by difference"), and if "dietary fibre" included or not.	"by difference" [100 - (fat + protein + ash + water)]	"by difference" [100 - (fat + protein + ash + water)]
Total CHO (as monosaccharides.) (g)		Sum of the analyzed CHO fractions	sum of the analyzed CHO fractions

Table 3 (continued)

Nutrients (*)	Comparability of nutrient values	Appropriate reference analytical methods (s)	Reference definition
Starch (g)	Comparable	Enzymatic hydrolysis + glucose-specific method/polarimetry	Starch including dextrins and glycogen
Sugar (g)	Comparable	Colorimetry, reductiometry, polarimetry, or HPLC	mono-, di- and oligosaccharides
Oligosaccha-rides (g)	Comparable		Raffinose, maltotriose, higher homologues
Lactose (g)	Comparable		Lactose
Dietary fibre (AOAC gives TDF) (g)	Similar values among TDF but higher than NSP (Englyst method) because AOAC-type methods include lignin and resistant starch. TDF and NSP values are globally compatible for fruits and vegetables (which have low lignin content and no resistant starch). NSP can be calculated by regression equation of Mongeau and Brassard (1989): (TDF + 0.02)/1.28 = NSP.	**AOAC methods** (based on the indigestibility of the components):	Total dietary fibre (**TDF**) = non-starch polysaccharides + lignin + resistant starch
Dietary fibre (NSP) (g)		**Englyst-type** methods	non-starch polysaccharides (**NSP**) without lignin, waxes, cutins, or resistant starch
Free Retinol (μg)	Comparability depend on methods	HPLC	all-*trans* retinol equivalents = all-*trans* retinol + 0.75 13-*cis* retinol + 0.90 retinaldehyde
Carotenes (μg)	Comparability depends on methods and definition. Caution because of high natural variability.	HPLC	β-carotene equivalent = 1 β-carotene+ 2 α-carotene+ 2 α-cryptoxanthin+ 2 β-cryptoxanthin

Table 3 (continued)

Nutrients (*)	Comparability of nutrient values	Appropriate reference analytical methods (s)	Reference definition
Vitamin D (total) (µg).	Comparability depends on methods	HPLC	Natural: cholecalciferol + 5 25-hydroxy-vitamin D. Manufactured: ergocalciferol
Vitamin E (total) (α-TE) (mg)	Comparability depends on method and definition	HPLC	α-TE = α-tocopherol tocopherols (0.4 β + 0.1 γ + 0.01 δ) + tocotrienols (0.3 α + 0.05 β + 0.01 γ)
Tocopherols (mg)	Comparability depends on methods	HPLC	β-, γ- ,δ-tocopherol
Tocotrienols (mg)		HPLC	α-, β-,γ-tocotrienol
Vitamin C (total) (mg)	Comparable, except titrimetry only AA). Caution due to handling of food samples, high natural variation.	AA+DHAA: HPLC, fluorimetry, colorimetry.	L-ascorbic (AA) + L-dehydroascorbic acids (DHAA),
Thiamin (mg)	Globally comparable	Microbiological, fluorimetry,	Thiamin
Riboflavin (mg)	Globally comparable	HPLC	Riboflavin
Folate (µg)	Incomparable values due to method and definition.	Microbiological, HPLC	Manufactured: pteroyl glutamic acid (folic acid). Natural: derivates of 5,6,7,8-tetrahydrofolic acid in mono- and polyglutamate (="bound").
Vitamin B6 (mg)	Globally comparable	HPLC (problem with extraction)	Pyridoxal + pyridoxamine + their phosphates and pyridoxine

Table 3 (continued)

Nutrients (*)	Comparability of nutrient values	Appropriate reference analytical methods (s)	Reference definition
Vitamin B12 (µg)	Compatible, but high potential measurement error because of low concentration in food	Microbiological assay	Vitamin B12
Potassium (mg)	Compatible if good laboratory practice is assured. Caution due to fortification	Atomic absorption spectrometry, flame or emission spectrophotometry. Titrimetry for Ca. Colorimetry for Fe.	K
Calcium (mg)			Ca
Iron (mg)			Fe

(*) : Unit and mode of expression

Standardization of the cells of the FCDB matrixes (nutrient values)
When the empty matrixes, defined by the final food and nutrient (or non-nutrient) axes, are completed the compilation of nutrient cells can actually start using a semi-automatic procedure. It is planned to develop a database management system (DBMS) to support the overall procedure of data compilation, which includes data entry, calculation, checking, storing, automatic comparison of nutrient values from different sources, and data retrieval and exchange. This compilation will be done in several steps:

Critical evaluation and standardization of the potential sources of nutrient values.
It is proposed to create different working groups involving experts on national food composition data and/or specific nutrients (carbohydrates, folates, carotenoids) who will evaluate the scientific value and comparability of the selected published and unpublished sources of nutrient values according to common criteria. For example, where only protein values are available in national tables, total nitrogen will be recalculated using the factors cited in the corresponding tables, or using the default 6.25 factor. Most of these data (national food or nutrient databases) will be available on hard disk support. This will facilitate the automatic transformation of the original data according to a reference standardized mode of expression and unit for each specific nutrient, using common algorithms implemented in the DBMS.

Individual evaluation and selection of nutrient values
For each specific food-nutrient pair stated in the FCDBs, the expert working groups will make a critical evaluation in order to decide on the final nutrient values to be applied after considering two types of foods: *common foods* and *country-specific foods*. For *common foods* it is assumed that the natural variability in their nutrient components within and between countries is much higher than the differences which may be attributed to differences between tables in sampling or chemical analysis. In addition, the information provided by study subjects is often relatively imprecise and it is not always possible to retrieve, for example, the specific botanic species of the apple or tomato actually consumed. For these common foods, the same source of nutrient values will be applied for all countries. The *country-specific foods,* for which the composition varies from country to country, will represent the majority of reported foods. For these foods, country-specific values will be used, after standardizing the source according to the reference criteria. This critical evaluation (compilation) will be done with particular emphasis on the foods most frequently reported in the 24-hour diet recalls and/or which are nutritionally important.

Actual compilation of food composition values

The nutrient values will be compiled automatically with the aim of reducing substantially the processing time and possible typing errors. This will be performed by selecting standardized nutrient value(s) from the satellite files (i.e. standardized national FCDBs stored on the hard disk) and copied automatically to a nutrient cell(s) or food line. The codes identifying the source of these nutrient values will also be systematically stored in a separate reference documentation file. When the information is derived from paper sources (e.g. scientific manuscripts, unpublished data) the standardization and data entry will be done manually.

A series of quality controls on entered or generated values will be implemented in the DBMS to identify missing, erroneous or inconsistent values.

Calculation of missing values

When no (or no alternate) values are available, missing values will be calculated automatically using harmonized procedures. Different cases exist. For foods consumed cooked but for which no corresponding nutrient values are available, adjustment for water, fat, vitamin and mineral losses will be done using reference standard coefficients. If a nutrient value is missing for a specific nutrient, alternate values will be borrowed from a comparable food or weighted values will be calculated. If the nutrient values of entire foods are missing, either standard recipes (e.g. cakes) or weighted mean values (e.g. vegetables n.s) will be automatically calculated.

Documentation

Foods reported in the EPIC FCDBs will be identified by their local English and, if relevant, taxonomic names. This will improve the subsequent identification, comparison and exchange of data within and between countries. Information on the sources of each nutrient value will be stored systematically in the DBMS to facilitate the retrieval of the original data. Specific codes will be used to identify the values derived from calculation, and a link will be kept with the specific calculation algorithm used. Recalculation of these values will then be possible at any time.

Database maintenance

The DBMS will be structured in such a way that recalculation and up-dating using more recent or reliable data will always be possible. The system will also be flexible enough to add new nutrients or food components in the FCDBs.

Discussion/Conclusion

The lack of comparable FCDBs is a major obstacle to investigating the wide range of food components of etiological interest in international multi-centre epidemiological studies. This reveals the paradoxical gap between the increasing scientific interest in large international multi-centre studies on diet and diseases and the lack of standardized tools to obtain comparable food and particularly nutrient measurements. To a large extent, this situation can be explained by the difficulty and cost of compiling comparable nutrient values. Another major reason is that so far attempts at standardization between national FCDBs have always been restricted to the harmonization of the mode of expression, unit and chemical analytical methods of *nutrients* (horizontal axis of the matrixes). Complete standardization of existing FCDBs has, however, always come up against the difficulty of standardizing the *food lists* (vertical axis of the matrixes) and subsequently their nutrient values. In this paper, we propose an approach to overcome these obstacles using the unique setting offered by the EPIC project which consists in using detailed food lists derived from already standardized 24-HDRs obtained from 35,000 subjects to build the vertical axes of the matrixes. This novel approach has the advantage of defining nine new matrixes, where the vertical axes are developed independently of the food lists reported in the national FCDBs. In addition, the level of detail indicated in these tables corresponds closely to that spontaneously reported by subjects. At this stage of the project, the matrixes of FCDBs are being defined using the infrastructure and data available in EPIC. International multi-disciplinary collaboration involving epidemiologists, nutritionists, food chemists and computer experts in DBMS is crucial for the success of such a project and is being developed. In addition, it is planned to take advantage of the INFOODS, EUROFOODS COST99 and NORFOODS experiences and recommendations on food, food component description systems and data interchange to enable the EPIC FCDBs to be easily interchangeable and compatible with other data management systems, at national or international level. A number of major methodological problems, particularly related to the actual compilation of FCDBs, still need to be solved due to the limitations of original national analytical data, lack of information on the sources, definition, food sampling, and on reference guidelines, standard coefficients and algorithms for calculating standard recipes or missing values. The harmonization of laboratory analyses, definition, mode of expression and units, already undertaken by international authorities should be promoted.

In addition, the decision to break down mixed recipes systematically into ingredients at the stage of data collection, in order to improve the comparability of food consumption data between countries and minimize food misclassification, has two major implications for the design and requirements of the FCDBs to be used. The FCDBs must list a large number of basic foods in the different ways they are usually consumed (i.e. raw/cooked, different cooking methods, physical states, meat cuts, part of plants, etc.). Ideally these food items should all be based on local analytical data using standard references. In addition, sophisticated algorithms should be developed to break down or calculate the mean nutrient content of multi-ingredient foods or standard recipes (e.g. cakes, soups). This implies investigating subsequently the algorithms for calculating nutrient values of multi-ingredient items (and breaking them down into ingredients) and having the reference information on the recipes, standard or derived from study subjects. The number of new foods and recipes increases continually, reflecting both the dynamic food industry and the large within- and between-individual variations in the way foods are prepared and consumed. We think that increasing the number of basic foods in different forms in FCDBs, and defining a systematic approach to deal with mixed recipes and multi-ingredient foods is the only way to reduce in the future the gaps between the needs expressed by the users and the information actually available in FCDBs.

As the use of biological markers in epidemiological research increases, it is also necessary to improve the estimation of absolute nutrient intake derived from dietary assessment methods, and particularly to express the foods in their "final as consumed" forms (i.e. including all nutrient losses due to food processing and cooking). In addition, the bioavailability of food components and its main determinants should be investigated further.

In the future, it is expected to have available a first series of nine comparable European FCDBs for international nutritional studies. These tables should each contain about 1500-2000 food items per country and nutrients covering most of the basic needs of nutritional studies. Although the subjects involved in EPIC are not strictly representative of the general local populations, the foods reported throughout the whole year in about 35,000 24-HDRs cover common foods consumed in the European countries participating in EPIC. Interest in the food composition databases to be developed therefore goes beyond the EPIC study to a wider spectrum of potential users in other nutritional or epidemiological studies, macro-economy, food industries and international institutions in responsible for establishing nutritional guidelines or nutrition policies.

In addition, this particular setting will be propitious for testing different methodologies such as various existing food or food component coding, description and classification systems using comparable empirical datasets. It is also planned to measure the impact of standardized food composition databases in the estimates of nutrient intake and relative risks. This will be obtained by comparing nutrient intakes estimated from national and EPIC FCDB, using the same set of dietary data. This will give an appraisal on the actual importance of standardizing food composition tables, particularly for pooled data analyses.

Acknowledgements

We would like to thank Professor. DAT Southgate for his precious contribution to this work, and his major imput in the selection and criteria of standardization of nutrients to be included in the food composition database . The authors would like to thank their former colleague, Geneviève Deharveng, whose review of the food composition tables available in the countries participating in EPIC was very valuable for the work described in this paper. The work was part of the EPIC project supported by the Europe Against Cancer Programme of the European Union.

References

1. Willett, W. (1998). Overview of nutritional epidemiology, and future research directions. In *Nutritional Epidemiology*, 2nd Edition (W Willett, Ed.). *Monogr. Epidemiol. Biostat.* 30, 1-17 and 484-496. Oxford University Press, New York, Oxford
2. Plummer, M., Clayton, D. And Kaaks, R. (1994). Calibration in multi-centre cohort studies. *Int. J.·Epidemiol.* 23, 419-426.
3. Kaaks, R., Plummer, M., Riboli, E., Estève, J. And Van Staveren W.A. (1994). Adjustment for bias due to errors in exposure assessments in multicenter cohort studies on diet and cancer : a calibration approach. *Am. J. Clin. Nutr.* 59 (Suppl.), S245S-S250.
4. Freedman, L. (1998). Challenges for statistical approaches to dietary assessment. *Eur. J. Clin. Nutr.* 52 (Suppl. 2), S6.
5. Slimani, N., Deharveng, G., Charrondière, R.U., Van Kappel, A.L., Ocké, M.C., Welch, A., Lagiou, A., Van Liere, M., Agudo, A., Pala, V., Brandstetter, B., Andrén, C., Stripp, C., Van Staveren, W.A. And Riboli, E. (1999). Structure of the standardized computerized 24-hour diet recall interview used as reference method in the 22 centers participating in the EPIC project. *Comput. Methods Programs Biomed.* 58, 251-266.
6. Polacchi, W. (1986). Standardized food terminology : an essential element for preparing and using food consumption data on an international basis. *Food Nutr. Bull.* 8, 66-68.
7. Klensin, J.C., Feskanich, D., Lin, V., Truswell, A.S. And Southgate, D.A.T. (1989). *Identification of Food Components for INFOODS Data Interchange.* UNU Press, Tokyo.

8. Truswell, A.S., Bateson, D., Madafiglio, K.C., Pennington, J.A.T., Rand, W.M. And Klensin, J.C. (1991). Committee report : INFOODS guidelines for describing foods: A systematic approach to describing foods to facilitate international exchange of food composition data. *J. Food. Comp. Anal.* 4, 18-38.
9. Greenfield, H. And Southgate, D.A.T. (1992). *Food Composition Data: Production, Management and Use.* Elsevier, London and New York.
10. Greenfield, H., Ed (1995). *Quality and Accessibility of Food-related Data.* AOAC International, Arlington, USA:
11. Scrimshaw, N. (1997). INFOODS: The international network of food data systems. *Am. J. Clin. Nutr.* 66 (Suppl.), 1190S-1193S.
12. Klensin, J.C. (1992) INFOODS Food Composition Data Interchange Handbook. The United Nations University, Tokyo, Japan.
13. Moller, A (1992). NORFOODS Computer Group – Food Composition Data Interchange among the Nordic Countries : A report In International Food Databases and Information Exchange. World Rev. Nutr. Diet.eds A.P. Simopoulos, R.R. Butrum. Karger, Basel, Switzerland, 94-103.
14. Schlotke, F. (1996). Using Internet services to improve international food data exchange. Food Chemistry - The Second International Food Database Conference, 57, 1, 137-143
15. Unwin, I. Becker, W (1996). The Component Aspect Identifier for compositional values. Food Chemistry – The Second International Food Database Conference. 57, 1, 149-154.
16. Bingham, S.A. (1987). The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutrition Abstracts and Reviews (series A)* 57, 705-742.
17. Cameron, M.E. And Van Staveren, W.A. (1988). *Manual on Methodology for Food Consumption Studies.* Oxford Medical Publications, Oxford.
18. Haraldsdóttir, J. (1993). Minimising error in the field: Quality controls in dietary surveys. *Eur. J. Clin. Nutr.* 47, (Suppl. 2), S19-S24.
19. Adelman, M., Dwyer, J.T., Woods, M., Bohn, E. and Otradovec, C.L. (1983). Computerized dietary analysis systems: A comparative view. *J. Am. Diet. Assoc.* 83, 421-429.
20. Hoover, L.W. (1983). Computerized nutrient databases: I. Comparison of nutrient analysis systems. *J. Am. Diet. Assoc.* 82, 501-505.
21. Nieman, D.C. And Nieman, C.N. (1987). A comparative study of two microcomputer nutrient data bases with the USDA nutrient data base for standard reference. *J. Am. Diet Assoc.* 87, 930- 932.
22. Eck, L.H., Kledges, R.C., Hanson, C.L., Baranowski, T. And Henske, J. (1988). A comparison of four commonly used nutrient database programs. *J. Am. Diet. Assoc.* 88, 602-604.
23. Lee, R.D., Nieman, D.C. And Rainwater, M. (1995). Comparison of eight microcomputer dietary analysis programs with the USDA nutrient data base for standard reference. *J. Am. Diet. Assoc.* 95, 858- 867.

24. Deharveng, G., Charrondière, R., Slimani, N., Riboli, E. And Southgate, D.A.T. (1999). Comparison of food composition tables available in the nine European countries participating in EPIC. *Eur. J. Clin. Nutr.* 53, 60-79.
25. Riboli, E. And Kaaks, R. (1997). The EPIC project: Rationale and study design. *Int. J. Epidemiol.* 26 (Suppl. 1), S6-S14.
26. Overvad, K., Tjønneland, A., Haraldsdóttir, J., Bang, S., Ewertz, M. And Møller-Jensen, O. (1991). Development of a semi-quantitative food frequency questionnaire to assess food, energy and nutrient intake in Denmark. *Int. J. Epidemiol.* 20, 906-912.
27. Bingham, S.A., Gill, C., Welch, A., Day, K., Cassidy, A., Khaw, K.T., Sneyd, M.I., Key, T.J.A., Roe, I. and Day, N.E. (1994). Comparison of dietary assessment methods in nutritional epidemiology: weighed records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. *Br. J. Nutr.* 72, 619-643.
28. Margetts, B.M., Pietinen, P., Riboli, E. (1997). EPIC European Prospective Investigation into Cancer and Nutrition. Validation studies on dietary assessment methods. *Int. J. of Epidemiol.* 26 (Suppl. 1), 1-189.
29. Bingham, S.A. and Cummings, J.H. (1985). Urine nitrogen as an independent validatory measure of dietary intake: a study of nitrogen balance in individuals consuming their normal diet. *Am. J. Clin. Nutr.* 42, 1276-1289.
30. Bingham, S.A., Cassidy, A., Welch, A., Runswick, S.A., Black, A.E., Thurnam, D., Bates, C., Khaw, K.T., Key, T.J.A. and Day, N.E. (1995). Validation of weighed records and other methods of dietary assessment using the 24 h urine nitrogen technique and other biological markers. *Br. J. Nutr.* 73, 531-550.
31. Ocké, M.C. And Kaaks R.J. (1997). Biochemical markers as additional measurements in dietary validity studies: application of the method of triads with examples from the European Prospective Investigation into Cancer and Nutrition. *Am. J. Clin. Nutr.* 65 (Suppl), 1240S-1245S.
32. Baghurst, K.I. and Baghurst, P.A. (1990). The uses of food compositional data in nutritional epidemiology. *Food Australia* 42, S11-S13
33. Slimani, N., Riboli, E. And Greenfield, H. (1995). Food composition data requirements for nutritional epidemiology of cancer and chronic diseases. In *Quality and Accessibility of Food-related Data* (H. Greenfield, Ed.), pp 209-215. AOAC International, Arlington, USA.
34. Kohlmeier, L. (1991). Problems and pitfalls of food-to-nutrient conversion. In *Food and Health Data. Their Use in Nutrition Policy-making* (W. Becker and E. Helsing, Eds.), pp 73-84. WHO Regional Publications, European series No. 34, WHO, Copenhagen.
35. West, C. The future information needs for research at the interface between food science and nutrition. In *Quality and Accessibility of Food-related Data* (H. Greenfield, Ed.) pp 141 –151. AOAC International, Arlington, USA:
36. Jackson, M.J., Fairweather-Tait, S.J., Van Den Berg, H. And Cohn, W., Eds. (1997). Assessment of the Bioavailability of Micronutrients (1997). *Eur. J. Clin. Nutr.* 51, (Suppl. 1), S1-S90

37. Franck, G.C., Hollatz, A.T., Webber, L.S. And Berenson, G.S. (1984). Effect of interviewer recording practices on nutrient intake - Bogalusa Heart Study. *J. Am. Diet. Assoc.* 84, 1432-1439.
38. Slimani, N., et al. Standardization of the 24-hour diet recall calibration method used in the the European Prospective Investigation into Cancer and Nutrition (EPIC): general concepts and preliminary results (2000). *Eur. J. Clin. Nutr.* 54, 900-917.
39. Pennington, J.A.T. And Butrum, R.R. (1991). Food descriptions using taxonomy and the "Langual" system. *Trends Food Sci. Technol.*, November, 285-288.
40. USDA (1992). *Langual Users Manual.* Centre for Food Safety and Applied Nutrition, Washington DC:
41. Krebs-Smith, S.M., Cronin, F.J., Haytowitz, D.B. And Cook, A. (1990). Contributions of food groups to intake of energy, nutrients, cholesterol, and fiber in women's diet: effect of method of classifying food mixtures. *J Am. Diet. Assoc.* 90, 1541-1546.
42. Mongeau, R. And Brassard, R. (1989). A comparison of three methods for analyzing dietary fiber in 38 foods. *J. Food Comp. Anal.* 2, 189-199.

Chapter 5

EPIC calibration study: rationale, design and study population characteristics

Nadia Slimani, Rudolf Kaaks, Pietro Ferrari, Corinne Casagrande, Françoise Clavel-Chapelon, Gerhard Lotze, Anja Kroke, Dimitrios Trichopoulos, Antonia Trichopoulou, C Lauria, Manuella Bellegotti, Marga C Ocké, Petra HM Peeters, Dagrun Engeset, Eiliv Lund, Antonio Agudo, Nerea Larrañaga, Irene Mattisson, Caren Andren, Ingegerd Johansson, Gwyneth Davey, Ailsa A Welch, Kim Overvad, Anne Tjønneland, Wija A Van Staveren, Rudolfo Saracci, Elio Riboli

Public Health Nutrition (in press)

Abstract

The European Prospective Investigation into Cancer and Nutrition (EPIC), which covers a large cohort of half a million men and women from 23 European centres in 10 Western European countries, was designed to study the relationship between diet and the risk of chronic diseases, particularly cancer. Information on usual individual dietary intake was assessed using different validated dietary assessment methods across participating countries. In order to adjust for possible systematic over- or underestimation in dietary intake measurements and correct for attenuation bias in relative risk estimates, a calibration approach was developed. This approach involved an additional dietary assessment common across study populations to re-express individual dietary intakes according to the same reference scale. A single 24-hour diet recall was therefore collected, as the EPIC reference calibration method, from a stratified random sample of 36,900 subjects from the entire EPIC cohort, using a software programme (EPIC-SOFT) specifically designed to standardize the dietary measurements across study populations. This paper describes the design and populations of the calibration sub-studies set up in the EPIC centres. In addition, to assess whether the calibration sub-samples were representative of the entire group of EPIC cohorts, a series of subjects' characteristics known to possibly influence dietary intakes were compared in both population groups. This was the first time that calibration sub-studies had been set up in a large multi-centre European study. These studies showed that, despite certain inherent methodological and logistic constraints, a study design such as this one works relatively well in practice. The average response in the calibration study was 78.3% and ranges from 46.5% to 92.5%. The calibration population differed slightly from the overall cohort but the differences were small for most characteristics and centres. The overall results suggest that, after adjustment for age, dietary intakes estimated from calibration samples can reasonably be interpreted as representative of the main cohorts in most of the EPIC centres.

Introduction

Compared to retrospective case–control or ecological studies, large multi-centre prospective studies offer major advantages for investigating the relationship between diet and other lifestyle factors and risk of chronic diseases (1-3). These studies are designed to increase the statistical power to detect an association between diet and disease by including large study populations varying both in the type of dietary patterns and in cancer incidence rates, thus increasing the heterogeneity of both exposure and disease outcomes (4). However, multi-centre studies also raise relatively new statistical and methodological issues for the comparison and pooled analysis of dietary intake data collected from large heterogeneous populations with wide differences in food consumption, language and socio-cultural characteristics. In particular, the inherent difficulty of estimating and comparing individuals' usual dietary intakes is amplified in large multi-centre studies, where dietary questionnaires often differ across study populations in order to capture the specific local diets (5). The magnitude and nature of systematic and random errors in dietary intake measurements may thus vary across study populations and distort the estimation and interpretation of the overall relationship between diet and disease when all cohorts are combined.

Several authors (6-10) proposed the use of a calibration approach in large nutritional studies. The purpose of such calibration studies is twofold: first, at the population level, to adjust for systematic over- and underestimation of the true mean dietary intakes in each centre; second, at the individual level to attempt to correct for attenuation bias in relative risk due to random errors in dietary measurements. For calibration at the population level, where emphasis is on unbiased estimates of mean intake, calibration can be achieved by applying, in addition to the dietary questionnaires, a second highly standardized dietary method in a representative sub-sample from each cohort as a common reference measurement across study populations. In its simplest definition, calibration means re-expressing the individual dietary measurements by means of centre-specific scaling factors. At the individual level, correction for regression dilution can only be fully achieved if the measurement errors of the second dietary assessment instrument are independent from the errors of the main instrument used in the entire cohort.

Although this complex study design initiates a new generation of large nutritional cohorts with nested calibration sub-studies, there is still little experience on how to set up such studies in practice (11,12). This paper describes the design of the calibration sub-studies within the European Prospective Investigation into Cancer and Nutrition (EPIC), a network of prospective cohort studies involving 23 European centres from 10 Western European countries (France, Italy, Spain, United Kingdom, Germany, The Netherlands, Greece, Sweden, Denmark and Norway). Among the most important features of EPIC are its size, geographical distribution and heterogeneity of the dietary patterns and other lifestyle and socio-cultural characteristics of the study populations. Information on usual diet, lifestyle, environmental factors and anthropometry were collected from each individual at baseline, as well as one blood sample. Information on usual individual dietary intakes was assessed using different dietary history questionnaires, food frequency questionnaires or a modified dietary history (13) developed and validated in each participating country (14-16). More details on the EPIC study design, the study cohort populations, the individual information collected, and the EPIC biological bank are given in Riboli et al.(17).

In addition, a single 24-hour dietary recall (24-HDR) was collected from a sub-sample of 36,900 individuals, to be used as the EPIC reference calibration method. Computerized 24-hour dietary recall interview software (EPIC-SOFT) was developed to standardize dietary intakes reported across the EPIC centres and increase the likelihood that measurement errors will be of a similar magnitude and nature in all study centres. The concept of standardization and the structure of the EPIC-SOFT software are described in detail in Slimani et al. (18,19).

This paper describes the design and populations of the calibration sub-studies set up in the centres participating in the European Prospective Investigation into Cancer and Nutrition (EPIC). In addition, in order to assess whether the calibration sub-samples were representative of the overall group of EPIC cohorts, a series of subjects' characteristics known to possibly influence dietary intakes were compared in both population groups.

Study protocol

Sampling procedures

The EPIC study populations were not chosen to provide representative samples. Recruitment was determined by practical and logistic considerations in order to obtain high participation and long-term follow-up from the study participants (17). These study populations represent heterogeneous groups and were population-based (Bilthoven, The Netherlands; Greece; Germany; Sweden; Denmark; Norway; Cambridge and a small part of the Oxford cohort from the UK; Spain; Italy) or participants in breast screening (Utrecht, The Netherlands; Florence, Italy) or teachers and school workers in France. In Oxford, most of the cohort (~87%) was recruited among subjects with interest in health and/or vegetarian eating habits and were self-defined vegans (i.e. consumed no animal products), ovo-lacto vegetarians, fish eaters (i.e consumers of fish but not meat), and meat eaters. Blood donors were also recruited in different proportions in certain Italian and Spanish centres. In France, Norway, Utrecht (The Netherlands) and Naples (Italy) only women were recruited.

The calibration population was defined as a random sample from each of these cohorts, weighted according to the cumulative numbers of cancer cases expected over 10 years of follow-up per gender and 5-year age strata. The sample sizes were chosen to provide calibration at both the individual and population level, even though it was recognized that the 24-hour dietary recall and the main dietary assessment instrument would not have fully independent error structures. A total of about 4000 24-hour dietary recalls, equivalent to a single, large random sample drawn from each *full* country cohort, was recommended per country, according to calculations detailed elsewhere (20). This sample size was achieved in most countries, except in the United Kingdom (1117), Norway (1819) and Greece (2930), and represents, according to the age distribution and size of the cohort, between 5% and 12% of the study population in each national cohort, except in the United Kingdom (~1.5%). In Norway, the calibration sample size requirement was smaller because lower numbers of cancers are expected from the relatively young cohort of women only. In Greece, a 10% representative sample of the entire cohort (28,572) was recruited all over Greece including Athens. In the United Kingdom, the sample size of 1117 was chosen to provide population level calibration, i.e. the sample size was calculated to give a sufficiently accurate estimation of mean intakes. Much of the UK cohort has already completed a second dietary instrument, a 7-day diet diary which included an interviewed 24-hour recall as a component, which could be used for (within-cohort) individual-level calibration should it be required.

In certain countries, the calibration population was sampled strictly at random, particularly when the age distributions were quite narrow, as in Norway (e.g. 49.3 ± 4.3). In France, where the study population was scattered all over the country and it was not possible to interview subjects living far from large urban areas, cluster sampling was used: contiguous, sparsely populated administrative regions were grouped into seven geographical areas in which the subjects could more easily be randomly sampled and approached for home visits or invited to a local centre for the 24-HDR interview. Using cluster sampling, a higher probability of being sampled was given to clusters with a higher number of subjects and vice versa.

In addition, the sampling procedures were defined as having an equal distribution of season and day of interview to control for possible weekday and seasonal variations in dietary consumption. Although in Spain, for example, the response rates obtained during the pilot phase were as high for interviews on Saturdays as for the other days of the week, other countries such as France and the Netherlands experienced high refusal rates for Saturday interviews. In addition, certain examination centres were closed or interviewers did not work during the weekends (e.g. Denmark, The Netherlands). Because of these constraints, alternative methods were considered to approach and interview the subjects during the weekend (i.e. to recall the Saturdays and Fridays). In some instances, interviews on Fridays and Saturdays were collected 48 hours later instead of the following day, and interviews at home were proposed to increase the participation rate. Such methods were rarely used for other weekdays, unless it was impossible to interview on the given sampled day or perform the interview at the examination centre.

Logistics used to set up the EPIC calibration sub-studies

The EPIC calibration fieldwork was conducted over a 5-year period between March 1995 and June 2000. A pilot phase was started initially in France (Rhône-Alpes), Spain (Basque Country) and the Netherlands in order to test the first version of EPIC-SOFT and the overall logistics. It was then extended to the other countries/centres according to the availability of country-specific EPIC-SOFT versions and the date of entry in EPIC. In order to have representative calibration sub-populations, the calibration study lasted until the end of the EPIC baseline recruitment. Depending on the country, it took between 10 and 31 months to collect the interviews sampled to cover both weekday and seasonal variations. Overall it took longer to perform the required dietary interviews in countries where several local, geographically distant centres were involved (e.g. France, Italy or Spain), where total population coverage was attempted (Greece) and where different study populations, study designs and recruitment methods were used. The time required for conducting the calibration field-work was not strictly related to the total number of interviews because at least 1 year was needed to cover all seasons. It is also interesting to note, for example, that the Nordic countries, which joined EPIC later, benefited from more advanced methodology (i.e. the overall logistics and EPIC-SOFT programmes were fully developed and tested), which allowed them to complete the interviews more quickly than other countries.

Table 1 summarizes the methods of recruitment and the localization of the 24-HDR dietary interviews performed in the EPIC centres. Whenever possible, subjects were recruited to the calibration study "by surprise", when they came for their first baseline examination. The dietary interview was then performed 30–40 minutes immediately after their baseline examination. It was anticipated that this recruitment approach would give a higher participation rate, as subjects would not have to return to the examination centres. This method was used for 84–100% of the subjects in Paris and the surrounding area (Ile-de-France), Potsdam, The Netherlands and Denmark, and to a lesser extent in other French centres (i.e. Rhône-Alpes, 36%, and Bretagne/Pays-de-Loire, 28%), Heidelberg (62%) and Ragusa (42%).

In the other centres, the subjects had either been enrolled before joining EPIC or had already been invited for the baseline examination when the calibration study was started. The subjects randomly selected to participate in the calibration study were then re-contacted either by letter (France, United Kingdom, Norway and Sweden) or by telephone (Spain, Italy (except Ragusa) and Greece). In contrast to recruitment at baseline, the subjects re-invited by letter or telephone were asked to return to provide further information but were not informed about the type of dietary method or time period to which it referred (i.e. the previous day). This precaution was taken in order to avoid changes in usual dietary habits and bias during the recalled dietary interview. According to what best suited the subjects and the local facilities available, the face-to-face 24-HDR interview was performed at the local research centre or at home, particularly if the people were living far from the research institute. In France, the overall study was coordinated from Paris and local authorities and cancer leagues made rooms available to conduct the dietary interviews (e.g. schools, town halls, local cancer leagues). In Greece, where it was particularly difficult to recruit local volunteers outside Athens, mobile units were used. In Norway, where the interviews were conducted by telephone, the subjects were all interviewed at home. According to the country and method of recruitment used, one to four reminders were sent when subjects did not reply.

Table 1. Methods of recruitment of the EPIC calibration sub-populations and related issues

	Study period covered	Number of subjects Approached	Recruitment method used (%)			Number of reminders	Type of reminders used (+/-)			Location of interviews (%)			
			Letter	Tel.	Centre		None	Letter	Tel.	Centre	Home	Else where	Unknown /missing
Greece	**05/97–06/99**	**5406**	**-**	**100**	**-**	**2**	**-**	**-**	**+**	**-**	**-**	**100[c]**	**-**
Spain		**3741**								**99.4**	**0.2**	**0.2**	**0.2**
Granada	12/95–10/96	722	-	100	-	0	+	-	-	100	-	-	-
Murcia	10/95–09/96	591	2	98	-	0	+	-	-	98.9	0.2	-	0.9
Pamplona	02/96–01/97	850	-	100	-	0	+	-	-	100	-	-	-
San Sebastian	07/95–06/96	806	100[a]	100[a]	-	0	+	-	-	98.1	0.7	0.8	0.4
Asturias	02/96–01/97	772	-	100	-	0	+	-	-	100	-	-	-
Italy		**4418**								**93.7**	**6.3**	**-**	**-**
Ragusa	04/96–07/98	348	46	12	42	1	-	-	+	92.0	8.0	-	-
Naples	04/96–03/98	482	-	99	1	0	+	-	-	99.0	1.0	-	-
Florence	05/96–08/98	1155	-	99	1	0	+	-	-	92.0	8.0	-	-
Turin	04/96–03/98	1172	-	99	1	0	+	-	-	87.0	13.0	-	-
Varese	04/96–06/98	1261	-	99	1	0	+	-	-	97.0	3.0	-	-

Table 1. (Continued)

	Study period covered	Number of subjects Approached	Recruitment method used (%)			Number of reminders	Type of reminders used (+/-)			Location of interviews (%)			
			Letter	Tel.	Centre		None	Letter	Tel.	Centre	Home	Else where	Unknown /missing
France		**6456**								**29.5**	**19.9**	**45.9**	**4.7**
Languedoc / Roussillon	10/95–04/97	869	100	-	-	2–3	-	+	-	-	22.4	73.1	4.5
Aquitaine	10/95–02/97	803	100	-	-	2–3	-	+	-	26.6	16.5	55.1	1.8
Rhône-Alpes	03/95–03/97	623	64	-	36	2	-	+	-	-	33.8	62.5	3.7
Bretagne/Pays-de-Loire	01/96–04/97	665	72	-	28	1–2	-	+	-	0.2	30.1	69.5	0.2
Ile-de-France	01/96–07/97	1575	16	-	84	0	-	-	-	23.4	24.5	41.5	10.6
Alsace-Lorraine	10/95–02/97	1343	100	-	-	1–2	-	+	-	84.0	8.0	7.5	0.5
Nord-Pas-de-Calais	10/95–07/97	578	100	-	-	2–3	-	+	-	0.2	19.7	67.5	12.6
Germany		**4693**								**81.1**	**18.5**	**0.2**	**0.2**
Heidelberg	06/96–10/98	2349	-	38[b]	62[b]	0	+	-	-	61.7	37.3	0.5	0.5
Potsdam	06/96–04/98	2344	-	-	100	0	+	-	-	98.9	1.1	-	-
The Netherlands		**5642**								**67.0**	**31.9**	**0.1**	**0.9**
Bilthoven	04/95–12/97	3411	-	-	100	0	+	-	-	96.3	2.1	0.0	1.6
Utrecht	04/95–12/97	2231	-	-	100	0	+	-	-	24.8	74.8	0.3	0.1
United Kingdom		**1900**								**99.4**	**0.6**	**-**	**-**
Cambridge	03/97–04/98	905	100	-	-	1–4	-	+	+	98.7	1.3	-	-
Oxford	02/97–07/98	995	100	-	-	0	+	-	-	100	-	-	-

Table 1. (Continued)

	Study period covered	Number of subjects Approached	Recruitment method used (%)			Number of reminders	Type of reminders used (+/-)			Location of interviews (%)			
			Letter	Tel.	Centre		None	Letter	Tel.	Centre	Home	Else where	Unknown /missing
Denmark		**4511**								**96.1**	**3.9**	**-**	**0.1**
Copenhagen	07/96–05/97	3268	-	-	100	0	-	-	-	94.6	5.4	-	0.1
Aarhus	08/96–05/97	1243	-	-	100	0	-	-	+	100	-	-	-
Sweden		**8413**								**100**	**-**	**-**	**-**
Malmö	03/96–07/97	4064	100	-	-	2	-	+	+	100	-	-	-
Umeå	03/97–10/98	4349	100[a]	100	-	0	-	-	+	100	-	-	-
Norway	**05/99–06/00**	**2993**	**100**		**-**	**1**	**-**	**+**	**-**	**-**	**100**	**-**	**-**

[a] The subjects were contacted first by letter then systematically approached by telephone to confirm the date and time of the interview.

[b] The subjects were all approached at the study centre when they came for the baseline examination. If it was not possible to fix an appointment immediately, the subjects were re-contacted by phone.

[c] In Athens, the interviews were conducted in the coordinating and elderly centres. Elsewhere in the country, the interviews were conducted in public places and occasionally at home.

Exclusions and inclusions of subjects from the 24-HDR dataset

The information reported in the following tables was calculated from the final 24-hour dietary recall dataset (*n*=35,955) obtained after further 24-HDR exclusions or inclusions from the original sample. We excluded 358 (~1%) interviews locally, mainly because of technical problems with the software during the interview or because subjects were not properly randomized or excluded from the EPIC cohort for other reasons (e.g. incomplete data). Subjects under 35 and over 74 years of age (who were present only in a few EPIC cohorts) were excluded from the dataset before statistical analyses. This represented a total of 945 subjects, mainly from Bilthoven (583 young people) and Greece (244, mainly elderly people).

In addition, 357 (<1% of the total final sample) subjects not originally sampled were added to the calibration population (29 subjects from Naples, 130 from Potsdam, 46 from Cambridge, 152 from Oxford). These subjects were involved in other EPIC cross-sectional or validation studies on urinary or blood biological markers and diet and most of them were participants sampled from the calibration sub-populations. Apart from Oxford, all have an EPIC-SOFT 24-hour dietary recall measurement collected in previous EPIC pilot studies. In Oxford, the 152 subjects added, essentially vegans and vegetarians, were not part of the representative sample initially selected for the calibration study, had no EPIC-SOFT 24-hour dietary recall measurements and were difficult to re-contact because they were living all over the UK. It was therefore decided to sample randomly 1 day from the 7-day records, collected at baseline in the UK, in addition to the EPIC food frequency questionnaire, and enter them using EPIC-SOFT as a data entry system (i.e. using the same rules as during a classic face-to-face interview).

Redefinition of the EPIC centres

With a view to the final statistical analyses, we decided to redefine the centres and geographical regions used to set up the field calibration studies in France, the United Kingdom and Norway and reported in Tables 1 and 2. In France, the seven geographical regions initially set up to facilitate the calibration field study were reduced, by clustering the 95 French "départements", to four geographical regions more representative of the different dietary patterns existing across the country (i.e. North-east, North-west, South, and South coast). The cohort of subjects recruited from the general population both in Cambridge and Oxford via general practitioners was grouped together ("general population group"). The UK "health conscious" group recruited by post was considered as a separate population group involving heterogeneous sub-populations of vegans, vegetarians, fish eaters and meat eaters.

In Norway, it was decided to subdivide the study populations scattered all over the country to coastal (North and West) and inland (South and East) regions. The coordinating centre "Bilthoven" covers three towns (Amsterdam, Doetinchem, Maastricht), where the subjects were recruited. A total of 27 centres were finally used for the analyses of the EPIC calibration dietary data and for presenting the results reported in Tables 3–9. These include administrative centres and geographical regions in France and Norway, but for convenience the term "centre" is used for both.

Participation rates, general characteristics and representativeness of the EPIC calibration sub-populations

Participation rates in the EPIC calibration sub-studies

The participation rates in the calibration sub-studies obtained in the different EPIC administrative centres are reported in **Table 2.** These calculations were obtained before any of the exclusions or additions discussed previously. At the country level, they ranged from 91.6% (Germany) to 54.2% (Greece), and 7 countries out of 10 had a participation rate of ~75% or more (Sweden, France, the Netherlands, Italy, Spain, Denmark, Germany). The response rate was about 60% for the general population in Norway and the UK and lower in Greece (54.2%) and in the "health-conscious" sub-cohort from Oxford (46.5%).

In Germany, Denmark and the general population in the UK, no differences were observed in the response rates across centres from the same country. In contrast, in France, a higher rate was observed in Ile-de-France, where the subjects were interviewed immediately after the baseline examination, than in other centres where study participants were re-invited for interviews sometimes more than 2 years afterwards. Within Italy and Spain, lower participation rates were reported in southern centres (Naples and Ragusa, Granada). In the UK, the rate was about 25% lower among the "health-conscious" group (46.5%) compared to the general population group, both in Cambridge and Oxford (60.4% and 63.3%, respectively). One possible explanation is that the "health-conscious" group's initial participation in EPIC was solely by post, so participation in the 24-HDR was their first visit to an EPIC examination centre which was, on average, a greater distance for them. In Greece, the low participation rate (54.2%) was largely due to logistic difficulties of approaching subjects living outside Athens.

Table 2. Participation rate obtained from the EPIC centers and countries[a]

	No. of subjects selected to be approached	No. of subjects interviewed	Participation Rate (%)	Non-response Active (%)	Non-response Passive (%)	Non-response Total (%)
Greece	**5406**	**2930**	**54.2**	**15.3**	**30.5**	**45.8**
Spain	**3741**	**3222**	**86.1**	**9.2**	**4.6**	**13.9**
Granada	722	515	71.3	24.7	4.0	28.7
Murcia	591	548	84.1	3.4	3.9	7.3
Navarra	850	715	91.1	7.8	8.1	15.9
San Sebastian	806	734	86.1	5.2	3.7	8.9
Asturias	772	710	92.0	5.2	2.8	8.0
Italy	**4418**	**3961**	**89.7**	**7.2**	**3.1**	**10.3**
Ragusa	348	306	87.9	6.7	5.4	12.1
Naples	482	403	83.6	12.4	4.0	16.4
Florence	1155	1058	91.6	7.1	1.3	8.4
Turin	1172	1069	91.2	5.4	3.4	8.8
Varese	1261	1125	89.2	7.2	3.6	10.8
France	**6456**	**4854**	**75.2**	**20.5**	**4.3**	**24.8**
Languedoc/ Roussillon	869	625	72.0	22.5	5.5	28.0
Aquitaine	578	443	76.6	19.2	4.2	23.4
Rhône-Alpes	1575	1018	64.6	33.3	6.3	39.6
Bretagne/Pays-de-Loire	803	635	79.1	15.2	4.6	19.8
Ile-de-France	1343	1201	89.4	9.2	1.7	10.9
Alsace-Lorraine	665	480	72.2	22.6	5.2	27.8
Nord-Pas-de-Calais	623	452	72.6	25.0	2.4	27.4
Germany	**4693**	**4299**	**91.6**	**8.2**	**0.2**	**8.4**
Heidelberg	2344	2126	90.7	8.9	0.4	9.3
Potsdam	2349	2173	92.5	7.5	-	7.5
The Netherlands	**5642**	**4585**	**81.4**	**9.7**	*8.9*	**18.6**
Bilthoven	3411	2708	79.4	10.7	9.9	20.6
Utrecht	2231	1877	84.1	8	7.8	15.6
United Kingdom	**1900**	**1117**	**59.0**	**30.0**	**11.0**	**41.0**
Cambridge	905	547	60.4	30.2	9.4	39.6
Oxford gen. Population	640	405	63.3	20.9	15.8	36.7
Oxford health conscious	355	165	46.5	50.7	2.8	53.5
Denmark	**4511**	**3919**	**86.9**	**13.1**	**-**	**13.1**
Copenhagen	3268	2842	87.0	13.0	-	13.0
Aarhus	1243	1077	86.6	13.4	-	13.4
Sweden	**8413**	**6195**	**73.6**	**20.4**	**5.9**	**26.3**
Malmö	4064	3132	77.1	19.0	3.9	22.9
Umeå[b]	4349	3063	70.4	21.8	7.8	29.6
Norway	**2993**	**1819**	**60.8**	**24.0**	**15.2**	**39.2**

[a]Estimates obtained before any exclusion/addition of subjects.
[b]149 individuals were excluded due to a mix up of food frequency questionnaires.

Apart from the Netherlands and Greece, non-participation in the calibration sub-studies was due primarily to the subjects' failure or refusal to respond to the invitation or accept an appointment for the dietary interview ("active" non-response). This represents more than 20% of the total subjects approached in Sweden, France, Norway and the UK, with some variations across centres. In Spain for example, the active non-response rate ranged from 3.4% in Murcia to 24.7% in Granada. Overall, the active non-response was much lower when the subject was approached by surprise just after the main examination than when he/she was contacted afterwards. In Germany and Denmark the non-participation was exclusively active.

In contrast, in the Netherlands and Greece "passive" non-response (i.e. non-participation because it was impossible to get in touch with the subjects) was about the same as, or higher than, active non-response. In the Netherlands, this was because it was often impossible to contact subjects by phone and because staff at the baseline examination centre forgot to refer subjects to the dietitians. In Greece, the passive non-response rate was particularly high (30.5%), which explains the overall low participation rate (~54%). The areas outside Athens were visited by mobile units for short periods only and passive non-response was unavoidably high because of time constraints for approaching and interviewing subjects. In Norway, 15% of the subjects approached could not be interviewed because of passive non-response.

General characteristics and representativeness of the EPIC calibration sub-populations

Tables 3–9 present a series of characteristics of the calibration sub-populations considered for analysis in this supplement. In addition, to estimate the representativeness of the calibration sub-population, we compared this sample to the rest of the EPIC cohort according to certain variables known to influence dietary consumption. In order to take into account the differences in age distribution in the calibration samples, due to the age-stratified sampling strategy, all the results are presented age-adjusted. We tested for significant differences in weight, height and body mass index (BMI) mean estimates between the calibration sample and the rest of the EPIC population. Differences in smoking status, level of education and physical activity at work (categorical variables) were tested using gender- and centre-specific logistic regressions. We modeled the different categorical variables separately as independent covariates, an indicator for distinguishing the calibration sample from the rest of the cohort as a binary outcome, and age as an adjusting variable. Significance was assessed using likelihood ratio statistics, at 95%, 99% and 99.9% levels. Analyses were performed using SAS software (21).

Since the main focus of this paper is the calibration sub-studies, the entire EPIC cohorts, detailed elsewhere (17, 22-24), will be described only for the purpose of comparing the two population groups.

a. *Age and anthropometry*

The calibration sample is composed of middle-aged populations, from 49.3 ± 4.3 years old (Norway) to 58.6 ± 8.4 (Sweden) in women, and from 50.0 ± 7.4 (Bilthoven) and 61.1 ± 7.3 (Sweden) in men (**Table 3 a & b**). Anthropometry varies considerably across countries. Height adjusted for age is about 9–10 cm higher for women in Norway than in Spain and for men in Sweden than in Spain. The same order of difference (~10 kg) is observed for weight among women in France and Greece, whereas a difference in weight of only 5.5 kg is observed in men between Italy and the Netherlands. Spain and Greece report both the lowest heights and the highest BMI in women and men, whereas Italy, the Netherlands, Germany, the UK general population, Sweden and Denmark report about the same BMI in women (25–26) and men (26–27). The lowest BMI (≤24) is observed in the Norwegian women cohort and in highly selected study populations such as women teachers in France and "health-conscious" people in the UK.

When compared to the rest of the EPIC cohorts, the weight, height and BMI means of the calibration population showed statistically significant differences in 20–30% of the sex-specific centres. However, in centres where there was a statistically significant difference, this was usually modest in absolute terms. In most cases, the mean BMI differed between the calibration population and the entire cohort for centres where a statistically significant difference was observed for weight and/or height.

b. *Smoking status*

In this calibration population, the number of never-smokers is about 1.2–2.7 times higher in women than men. For women in Greece, Spain, France and the UK "health conscious" group, never-smokers represent ≥65% of the population (**Table 4a**) and about 35–60% elsewhere. In men, never-smokers represent 24–46% (**Table 4b**). The percentage of ex-smokers varies to a greater extent among women (7–33%) than men (28–46%), as does the percentage of smokers, from less than 9% to ~25% for women and from 21–40% for men, except in the UK (~17%).

In about a quarter of the EPIC centres, the smoking status is not equally distributed between the calibration sub-sample and the entire EPIC cohort. Most of the imbalance is due, however, to differences of only a few percentage points (<5%) across classes. Apart from the Spanish centres, the number of current smokers is equal or lower in the calibration sample than in the entire cohort. In contrast, the number of ex-smokers is higher in the calibration population, except in Umeå, and men in Spain. The number of never-smokers is relatively lower in the calibration sample in women from southern centres (France North-west, Navarra and Greece) whereas it tends to be higher in central and Nordic centres. In men, never-smokers are always equal or over-sampled in the calibration group compared to the entire cohort.

c. Level of education

A common variable in five classes of level of education was used in EPIC (**Tables 5 a & b**). In Malmö, however, where the cohort was recruited before joining EPIC, the education level of 7332 subjects (~25% of the total cohort) was defined differently and these subjects were therefore classified in the closest existing EPIC category (corresponding to "technical school"). Large differences are observed in education levels reflecting gender discrepancies and the diversity of origin of the cohorts (general population, blood donors, teachers, and "health-conscious" groups) (17). For example, the number of subjects who never completed primary school is high in Spain, particularly among women and in the south, and Greece whereas it is nil in the other countries and centres. In contrast, in Italy and France about half of the population, and a third in the Netherlands (women), have a secondary school diploma, and 40–50% of the UK "health conscious" group, French women teachers and men in Germany have a university degree.

In more than a half of the centres, the distribution according to the level of education is not strictly comparable between the calibration and the entire EPIC cohort populations. In certain centres, this difference is due to a few percentage-point differences in distribution across six classes. However, a quite consistent systematic tendency to under-represent the lowest education level classes (i.e. incomplete primary school and primary school) and to over-represent secondary and particularly university is observed.

d. Work-based physical activity

The subjects were asked to report their professional physical activities using a variable in four categories (sedentary, standing, manual work, and heavy manual work) (23, 25). In Malmö, subjects were asked about their physical activity at work as typical professional activity without referring to current occupational status, and a different physical activity questionnaire was used in Norway which was therefore not included in the analysis. In Spain, all participants were classified in one of the categories of work activity independently of employment status, so these variables are not directly comparable with the professional activities reported elsewhere (23).

These differences should not, however, affect the comparison of the calibration sub-sample with the entire cohort because the statistical analysis was stratified by centre.

Except Denmark, Italy and the Netherlands (≤20%) in men only, all countries report a relatively high proportion of subjects with no professional physical activity, particularly women (**Table 6 a & b**). EPIC cohorts have overall moderate professional physical activities, with predominantly sedentary or standing occupations, but a higher proportion of men with manual or heavy manual jobs and a lower number of non-workers compared to women is consistently observed in all EPIC cohorts.

The distribution of professional activities shows statistical differences between the calibration group and entire EPIC cohorts for about 40% of the sex-specific centres. We consistently observed a tendency to under-sample non-workers in most centres and both genders and to over-sample people with a sedentary and/or standing occupation. A more comparable distribution is, however, observed between the study groups when sedentary plus standing, and manual plus heavy manual activities are grouped together. In Bilthoven and the UK "health conscious" group, the discrepancies observed are mainly due to a higher completeness of the calibration (i.e. lower number of missing values) compared to the entire-cohort data.

e. Special diet

The number of study subjects who reported having a special diet during the 24-HDR interview was higher among women (12–39.4%) than men (7–29.7%) (**Table 7**). Apart from the UK "health conscious" group, long-term health problems related to diet (e.g. hyperlipidemia, hypertension, diabetes, stomach or intestinal problems) were the main reason given to explain their usual dietary habits, particularly in Umeå, and to a lesser extent, Greece. Except in Sweden, the number of subjects who reported restricting their dietary intake because of overweight or obesity was in all centres 1.4–5 times higher in women than in men. In the UK, 60% of the women and 74% of the men from the "health-conscious" sample are vegans, ovo-lacto vegetarians or fish eaters who do not consume meat. The number of vegetarians in the other EPIC cohorts is 3.6% or lower

Table 3a. Comparison of mean age and anthropometry, adjusted for age, of the EPIC calibration sub-populations and the rest of the EPIC cohorts: women

	n		Age (yrs)				Height (cm)				
	Calibration	Co-hort	*Calibration*		*Cohort*		*Calibration*		*Cohort*		*P value*[a]
			Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	
Greece	**1374**	**13644**	**57.2**	**9.9**	**54.7**	**11.1**	**157.3**	**6.0**	**156.6**	**6.5**	***
Spain	**1443**	**24088**	**52.9**	**8.3**	**48.4**	**8.3**	**156.7**	**5.7**	**156.1**	**6.2**	-
Granada	300	5750	54.6	8.1	49.3	8.7	156.0	5.6	155.1	6.0	**
Murcia	304	5418	51.6	8.6	48.3	8.3	155.9	5.5	155.5	6.1	
Navarra	271	3900	53.6	7.8	48.4	8.0	157.2	5.5	156.7	6.0	
San Sebastian	244	4009	51.8	8.2	48.1	8.1	158.5	5.7	157.8	6.3	
Asturias	324	5011	52.9	8.4	47.7	8.1	156.2	5.5	156.1	6.3	
Italy	**2512**	**29757**	**54.7**	**7.3**	**50.6**	**8.0**	**158.7**	**6.1**	**158.3**	**6.2**	-
Ragusa	138	3115	50.6	8.3	46.2	7.7	155.7	5.7	156.5	7.2	
Naples	403	4582	54.2	6.7	50.4	7.5	157.2	5.7	156.3	5.9	**
Florence	785	9212	55.3	7.0	51.9	7.6	160.2	6.1	159.9	6.3	
Turin	392	4147	54.2	6.9	50.6	7.5	159.2	6.1	158.8	6.3	
Varese	794	8701	55.3	7.6	50.9	8.4	158.3	5.9	157.9	6.0	
France	**4639**	**68357**	**57.0**	**6.9**	**52.8**	**6.7**	**161.5**	**5.7**	**161.4**	**6.0**	-
South coast	612	9283	57.6	6.8	53.7	6.6	161.4	5.7	161.2	6.2	
South	1396	17035	56.6	7.0	53.0	6.6	161.4	5.7	161.3	6.0	
North-west	622	11041	56.9	6.7	52.8	6.7	160.9	5.6	160.9	5.9	
North-east	2009	30998	57.1	7.0	52.5	6.6	161.8	5.8	161.7	6.1	
Germany	**2150**	**27961**	**51.6**	**8.7**	**49.2**	**9.0**	**163.0**	**6.1**	**162.8**	**6.5**	-
Heidelberg	1087	12530	50.3	8.5	49.4	8.6	163.6	6.1	163.3	6.4	
Potsdam	1063	15431	53.0	8.6	49.1	9.2	162.3	6.0	162.4	6.5	
The Netherlands	**2960**	**23264**	**55.1**	**8.3**	**54.3**	**7.9**	**164.9**	**6.2**	**164.5**	**6.5**	-
Bilthoven	1086	7781	48.9	7.5	48.0	7.5	164.5	6.7	163.9	8.2	**
Utrecht	1874	15483	58.7	6.3	57.5	6.0	164.9	6.1	164.4	7.3	**
United Kingdom	**768**	**45184**	**55.6**	**8.9**	**52.7**	**10.0**	**162.5**	**6.4**	**162.7**	**6.7**	-
Gen. Population	571	19978	56.1	9.0	56.0	8.5	162.2	6.1	161.8	6.3	
Health Conscious	197	25206	54.1	8.7	50.0	10.2	163.1	6.6	163.6	7.2	
Denmark	**1995**	**27880**	**56.8**	**4.4**	**56.8**	**4.4**	**164.1**	**6.0**	**163.9**	**6.0**	-
Copenhagen	1485	19669	57.1	4.4	56.9	4.4	164.3	6.0	164.0	6.1	
Aarhus	510	8211	55.9	4.4	56.4	4.4	163.7	5.8	163.7	5.9	
Sweden	**3285**	**24728**	**58.6**	**8.4**	**54.1**	**8.9**	**164.4**	**6.0**	**164.1**	**6.0**	-
Malmö	1711	15324	61.4	7.8	57.2	7.9	164.3	6.0	164.0	6.2	*
Umeå	1574	9404	55.6	8.0	49.0	7.9	164.0	6.1	163.7	6.2	*
Norway	**1798**	**35428**	**49.3**	**4.3**	**48.2**	**4.3**	**166.8**	**5.7**	**166.5**	**7.5**	-
South & East	1136	19576	49.5	4.3	48.2	4.3	167.1	5.7	166.8	7.5	
North & West	662	15852	49.0	4.3	48.1	4.3	166.3	5.6	166.1	7.4	

[a] *P* value: * <0.05; ** <0.01; *** <0.001.

Table 3a. (continued)

	n		Weight (kg)					BMI (kg/m²)				
	Calibration	Co-hort	*Calibration*		*Cohort*		*P* value[a]	*Calibration*		*Cohort*		*P* value[a]
			Mean	*SD*	*Mean*	*SD*		*Mean*	*SD*	*Mean*	*SD*	
Greece	**1374**	**13644**	**70.9**	**12.7**	**71.6**	**13.8**	-	**28.7**	**5.2**	**29.3**	**5.6**	***
Spain	**1443**	**24088**	**69.2**	**11.2**	**70.2**	**12.2**	-	**28.3**	**4.5**	**28.9**	**4.9**	-
Granada	300	5750	70.6	11.4	71.8	12.1		29.1	4.7	29.9	4.9	**
Murcia	304	5418	71.1	11.2	70.5	12.3		29.4	4.4	29.2	4.9	
Navarra	271	3900	67.5	10.7	69.4	11.7	**	27.4	4.3	28.3	4.7	***
San Sebastian	244	4009	68.0	11.0	68.4	12.2		27.1	4.3	27.5	4.8	
Asturias	324	5011	68.4	11.0	69.9	12.5	*	28.1	4.4	28.7	5.0	*
Italy	**2512**	**29757**	**64.8**	**11.0**	**64.9**	**11.3**	-	**25.8**	**4.3**	**25.9**	**4.4**	-
Ragusa	138	3115	65.3	11.0	66.4	13.9		26.9	4.4	27.1	5.5	
Naples	403	4582	67.0	11.1	66.6	11.6		27.1	4.3	27.3	4.5	
Florence	785	9212	64.8	10.8	64.6	11.1		25.3	4.0	25.3	4.1	
Turin	392	4147	64.5	11.0	64.3	11.3		25.5	4.2	25.5	4.3	
Varese	794	8701	63.7	11.1	64.1	11.3		25.5	4.3	25.7	4.4	
France	**4639**	**68357**	**60.8**	**9.8**	**60.7**	**10.3**	-	**23.3**	**3.5**	**23.3**	**3.7**	-
South coast	612	9283	59.9	9.3	60.0	10.1		23.0	3.4	23.1	3.7	
South	1396	17035	59.9	9.4	60.0	10.0		23.0	3.4	23.1	3.6	
North-west	622	11041	60.2	9.8	60.4	10.4		23.2	3.6	23.3	3.8	
North-east	2009	30998	61.9	10.1	61.3	10.6	*	23.7	3.6	23.5	3.8	*
Germany	**2150**	**27961**	**69.1**	**12.5**	**69.0**	**13.3**	-	**26.1**	**4.6**	**26.0**	**4.8**	-
Heidelberg	1087	12530	68.9	12.5	68.8	13.1		25.8	4.6	25.8	4.8	
Potsdam	1063	15431	69.4	12.6	69.2	13.6		26.3	4.5	26.3	4.9	
The Netherlands	**2960**	**23264**	**70.1**	**11.7**	**69.8**	**12.3**	-	**25.8**	**4.2**	**25.8**	**4.4**	-
Bilthoven	1086	7781	70.9	12.3	70.5	15.1		26.2	4.4	26.3	5.4	
Utrecht	1874	15483	70.2	11.7	70.0	13.8		25.9	4.1	25.9	4.9	
United Kingdom	**768**	**45184**	**66.2**	**11.7**	**66.1**	**12.1**	-	**25.1**	**4.3**	**25.0**	**4.4**	-
Gen. Population	571	19978	67.6	12.4	68.1	12.7		25.7	4.6	26.0	4.7	
Health Conscious	197	25206	62.5	10.9	64.4	11.9	*	23.5	3.9	24.1	4.2	*
Denmark	**1995**	**27880**	**68.9**	**12.2**	**68.8**	**12.4**	-	**25.6**	**4.4**	**25.6**	**4.4**	-
Copenhagen	1485	19669	69.2	12.2	68.9	12.4		25.6	4.4	25.6	4.4	
Aarhus	510	8211	68.0	12.1	68.6	12.4		25.4	4.4	25.6	4.4	
Sweden	**3285**	**24728**	**67.9**	**12.0**	**68.1**	**12.0**	-	**25.1**	**4.3**	**25.3**	**4.3**	-
Malmö	1711	15324	68.1	11.9	67.9	12.2		25.2	4.3	25.3	4.4	
Umeå	1574	9404	67.9	12.4	68.9	12.5	**	25.3	4.4	25.7	4.4	***
Norway	**1798**	**35428**	**67.6**	**12.1**	**68.2**	**15.9**	-	**24.3**	**4.1**	**24.6**	**5.4**	-
South & East	1136	19576	67.2	12.1	68.2	16.0	**	24.1	4.1	24.5	5.5	**
North & West	662	15852	68.1	12.0	68.3	15.8		24.6	4.1	24.7	5.4	

[a] *P* value: * <0.05; ** <0.01; *** <0.001.

Table 3b. Comparison of mean age and anthropometry, adjusted for age, of the EPIC calibration sub-populations and the rest of the EPIC cohorts: men

n			Age (yrs)				Height (cm)				
	Calibration	Co-hort	*Calibration*		*Cohort*		*Calibration*		*Cohort*		*P value*[a]
			Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	
Greece	**1312**	**9350**	**60.5**	**9.8**	**53.1**	**11.4**	**170.5**	**6.6**	**169.8**	**7.0**	***
Spain	**1777**	**13847**	**55.1**	**7.4**	**50.5**	**7.2**	**169.0**	**6.2**	**168.7**	**6.4**	-
Granada	214	1581	58.1	6.7	51.1	8.1	169.4	6.5	167.6	6.5	***
Murcia	243	2440	55.6	7.5	50.2	7.4	167.4	6.1	167.7	6.3	
Navarra	444	3462	56.3	6.8	50.2	6.8	168.6	6.1	168.7	6.3	
San Sebastian	490	3667	51.5	6.8	51.2	7.0	170.4	6.3	169.9	6.3	
Asturias	386	2697	56.2	7.5	49.9	7.0	168.8	6.2	168.4	6.5	
Italy	**1444**	**13655**	**55.2**	**7.0**	**49.9**	**7.5**	**171.9**	**6.7**	**171.3**	**7.0**	-
Ragusa	168	2883	53.5	6.8	48.4	7.0	169.2	6.1	168.3	7.0	
Naples	-	-	-	-	-	-	-	-	-	-	-
Florence	271	3211	54.4	7.3	50.5	7.3	173.1	6.4	173.0	6.6	
Turin	677	5332	55.0	7.0	49.3	7.7	172.0	6.7	171.8	7.1	
Varese	328	2229	57.1	6.2	52.7	7.1	172.1	6.7	171.4	6.5	
France	-	-	-	-	-	-	-	-	-	-	-
South coast	-	-	-	-	-	-	-	-	-	-	-
South	-	-	-	-	-	-	-	-	-	-	-
North-west	-	-	-	-	-	-	-	-	-	-	-
North-east	-	-	-	-	-	-	-	-	-	-	-
Germany	**2268**	**20549**	**54.6**	**7.3**	**52.3**	**7.5**	**175.5**	**6.5**	**175.4**	**6.5**	-
Heidelberg	1033	10896	53.7	7.0	52.5	7.1	175.9	6.5	175.9	6.5	
Potsdam	1235	9653	55.4	7.4	52.2	8.0	175.2	6.5	174.8	6.5	*
The Netherlands	**1024**	**6635**	**50.0**	**7.4**	**48.1**	**7.4**	**177.2**	**7.2**	**176.6**	**8.8**	-
Bilthoven	1024	6635	50.0	7.4	48.1	7.4	177.2	7.2	176.6	8.8	**
Utrecht	-	-	-	-	-	-	-	-	-	-	-
United Kingdom	**518**	**21514**	**57.5**	**8.9**	**55.1**	**9.9**	**176.4**	**6.9**	**175.9**	**7.1**	-
Gen. Population	404	13831	58.1	9.1	57.5	8.5	175.9	6.7	175.0	7.1	*
Health Conscious	114	7683	55.4	7.7	50.9	10.9	177.8	6.9	177.4	7.6	
Denmark	**1923**	**25256**	**56.7**	**4.3**	**56.6**	**4.4**	**176.7**	**6.5**	**176.5**	**6.6**	-
Copenhagen	1356	17390	57.0	4.4	56.7	4.4	176.8	6.5	176.7	6.6	
Aarhus	567	7866	56.0	4.2	56.4	4.4	176.4	6.3	176.1	6.5	
Sweden	**2765**	**18444**	**61.1**	**7.3**	**54.0**	**9.0**	**177.8**	**6.9**	**177.2**	**6.7**	-
Malmö	1421	9642	64.2	6.2	58.9	7.1	177.5	7.0	176.9	7.3	**
Umeå	1344	8802	57.8	6.9	48.7	7.8	177.6	7.0	177.0	6.9	**
Norway	-	-	-	-	-	-	-	-	-	-	-
South & East	-	-	-	-	-	-	-	-	-	-	-
North & West	-	-	-	-	-	-	-	-	-	-	-

[a]*P* value: * <0.05; ** <0.01; *** <0.001.

Table 3b. (Continued)

n	Calibration	Cohort	Weight (kg) Calibration Mean	SD	Cohort Mean	SD	P value[a]	BMI (kg/m²) Calibration Mean	SD	Cohort Mean	SD	P value[a]
Greece	**1312**	**9350**	**82.5**	**12.8**	**81.9**	**13.5**		**28.4**	**4.0**	**28.4**	**4.2**	
Spain	**1777**	**13847**	**81.1**	**10.9**	**81.2**	**11.2**	-	**28.4**	**3.4**	**28.5**	**3.5**	-
Granada	214	1581	83.4	11.4	81.4	11.3	*	29.1	3.7	29.0	3.7	
Murcia	243	2440	79.4	11.4	80.3	11.8		28.3	3.7	28.5	3.8	
Navarra	444	3462	80.8	10.4	82.4	10.8	**	28.4	3.3	29.0	3.4	**
San Sebastian	490	3667	80.9	10.8	81.0	11.0		27.8	3.2	28.0	3.3	
Asturias	386	2697	82.0	11.1	80.8	11.6		28.7	3.4	28.5	3.5	
Italy	**1444**	**13655**	**78.2**	**11.0**	**78.4**	**11.5**	-	**26.5**	**3.4**	**26.7**	**3.5**	-
Ragusa	168	2883	78.2	11.0	78.5	12.7		27.3	3.4	27.7	3.9	
Naples	-	-	-	-	-	-	-	-	-	-	-	-
Florence	271	3211	78.9	11.1	79.0	11.4		26.3	3.3	26.4	3.4	
Turin	677	5332	78.2	11.0	78.3	11.6		26.4	3.2	26.5	3.4	
Varese	328	2229	77.6	11.2	77.9	10.9		26.2	3.4	26.5	3.4	
France	-	-	-	-	-	-	-	-	-	-	-	-
South coast	-	-	-	-	-	-	-	-	-	-	-	-
South	-	-	-	-	-	-	-	-	-	-	-	-
North-west	-	-	-	-	-	-	-	-	-	-	-	-
North-east	-	-	-	-	-	-	-	-	-	-	-	-
Germany	**2268**	**20549**	**83.5**	**12.3**	**83.0**	**12.3**	-	**27.1**	**3.6**	**27.0**	**3.6**	-
Heidelberg	1033	10896	84.0	12.2	83.3	12.2		27.1	3.6	26.9	3.6	
Potsdam	1235	9653	83.3	12.4	82.7	12.4		27.1	3.7	27.1	3.7	
The Netherlands	**1024**	**6635**	**83.8**	**12.7**	**82.7**	**15.5**	-	**26.7**	**3.6**	**26.5**	**4.4**	-
Bilthoven	1024	6635	83.8	12.7	82.7	15.5	**	26.7	3.6	26.5	4.4	
Utrecht	-	-	-	-	-	-	-	-	-	-	-	-
United Kingdom	**518**	**21514**	**79.7**	**11.8**	**79.5**	**12.2**	-	**25.7**	**3.5**	**25.7**	**3.6**	-
Gen. Population	404	13831	81.6	11.7	81.0	12.4		26.4	3.4	26.4	3.6	
Health Conscious	114	7683	73.5	11.4	76.8	12.7	**	23.3	3.3	24.4	3.7	***
Denmark	**1923**	**25256**	**83.1**	**12.4**	**82.9**	**12.6**	-	**26.6**	**3.6**	**26.6**	**3.7**	-
Copenhagen	1356	17390	83.2	12.5	83.0	12.7		26.6	3.7	26.6	3.7	
Aarhus	567	7866	82.7	12.1	82.6	12.4		26.6	3.5	26.6	3.6	
Sweden	**2765**	**18444**	**82.2**	**12.8**	**82.1**	**12.4**	-	**26.0**	**3.7**	**26.1**	**3.6**	-
Malmö	1421	9642	82.2	13.0	82.0	13.6		26.1	3.7	26.2	3.9	
Umeå	1344	8802	82.1	12.9	82.0	12.8		26.0	3.7	26.2	3.7	
Norway	-	-	-	-	-	-	-	-	-	-	-	-
South & East	-	-	-	-	-	-	-	-	-	-	-	-
North & West	-	-	-	-	-	-	-	-	-	-	-	-

[a] *P* value: * <0.05; ** <0.01; *** <0.001.

Table 4a. Comparison of smoking status, adjusted for age, of the EPIC calibration sub-populations and the rest of the EPIC cohorts: women

	Smoking Status								
	Calibration				Cohort				
	Never smoker %	Ex-smoker %	Smoker %	Missing %	Never smoker %	Ex-Smoker %	Smoker %	Missing %	P value[a]
Greece	**70.9**	**6.5**	**18.0**	**4.6**	**74.1**	**5.3**	**15.4**	**5.3**	***
Spain	**69.3**	**11.2**	**19.5**	**0.1**	**72.0**	**9.6**	**18.4**	**0.1**	-
Granada	80.6	6.2	13.2	0.0	77.6	7.5	14.9	0.0	
Murcia	72.0	9.2	18.5	0.3	73.7	8.2	18.1	0.0	
Navarra	60.7	15.1	24.2	0.0	68.4	10.0	21.4	0.1	**
San Sebastian	65.2	16.8	18.0	0.0	69.3	13.3	17.4	0.1	
Asturias	66.8	9.4	23.8	0.0	68.5	10.4	21.1	0.1	
Italy	**54.1**	**21.7**	**24.2**	**0.0**	**52.9**	**19.8**	**25.8**	**1.4**	-
Ragusa	51.6	19.4	28.9	0.0	53.8	18.1	25.5	2.7	
Naples	46.4	21.2	32.4	0.0	41.5	18.3	40.2	0.0	
Florence	46.5	25.2	28.1	0.1	47.6	24.0	27.8	0.6	
Turin	57.3	24.0	18.7	0.0	53.2	20.6	19.8	6.4	
Varese	64.6	17.2	18.2	0.0	64.2	16.4	19.1	0.2	
France	**67.5**	**20.1**	**7.1**	**5.3**	**66.4**	**18.3**	**8.6**	**6.8**	-
South coast	71.6	15.4	8.4	4.6	66.4	17.2	9.4	7.0	
South	66.2	21.0	7.2	5.7	67.1	18.0	8.2	6.6	
North-west	65.3	22.0	8.5	4.2	67.7	17.9	7.8	6.7	*
North-east	68.7	18.3	7.4	5.6	65.5	18.8	8.9	6.8	
Germany	**52.3**	**28.9**	**18.8**	**0.0**	**53.7**	**27.8**	**18.5**	**0.0**	-
Heidelberg	47.9	31.3	20.9	0.0	48.3	30.6	21.1	0.0	
Potsdam	57.3	26.6	16.0	0.1	58.0	25.5	16.5	0.1	
The Netherlands	**41.7**	**33.1**	**24.8**	**0.4**	**39.6**	**33.1**	**26.6**	**0.6**	-
Bilthoven	33.7	31.9	33.8	0.6	34.4	30.5	34.9	0.1	
Utrecht	46.0	33.8	19.9	0.3	42.3	34.4	22.5	0.8	**
United Kingdom	**61.4**	**28.4**	**8.8**	**1.4**	**56.3**	**30.3**	**10.4**	**3.0**	-
Gen. Population	59.9	28.5	9.8	1.8	52.7	28.0	12.8	6.5	*
Health Conscious	64.6	28.3	6.5	0.5	59.2	32.1	8.5	0.3	
Denmark	**47.9**	**26.7**	**24.9**	**0.4**	**43.0**	**24.5**	**31.9**	**0.6**	-
Copenhagen	49.0	26.4	24.2	0.3	42.2	24.3	32.8	0.7	***
Aarhus	44.2	27.9	27.3	0.6	45.2	24.9	29.5	0.4	
Sweden	**53.7**	**24.7**	**20.8**	**0.8**	**49.2**	**24.0**	**26.1**	**0.6**	-
Malmö	46.2	30.0	23.7	0.0	44.1	27.6	28.3	0.1	
Umeå	60.0	16.9	21.5	1.6	57.8	18.1	22.5	1.5	**
Norway	**34.9**	**33.4**	**26.5**	**5.2**	**33.5**	**29.1**	**31.8**	**5.7**	
South & East	35.9	34.2	24.9	5.0	33.7	29.1	31.5	5.7	***
North & West	33.4	32.1	29.0	5.4	33.2	29.0	32.1	5.6	

[a] *P* value: * <0.05; ** <0.01; *** <0.001.

Table 4b. Comparison of smoking status, adjusted for age, of the EPIC calibration sub-populations and the rest of the EPIC cohorts: men

	Smoking Status								
	Calibration				Cohort				
	Never smoker %	Ex-smoker %	Smoker %	Missing %	Never smoker %	Ex-Smoker %	Smoker %	Missing %	P value[a]
Greece	**26.2**	**33.5**	**34.0**	**6.3**	**23.2**	**30.2**	**38.4**	**8.2**	***
Spain	**32.5**	**27.5**	**39.9**	**0.1**	**28.7**	**30.9**	**40.3**	**0.1**	-
Granada	31.1	39.4	29.4	0.0	27.4	39.4	33.1	0.1	
Murcia	28.2	28.0	43.8	0.0	24.3	38.0	37.7	0.1	**
Navarra	31.0	18.4	50.6	0.0	31.2	24.5	44.3	0.0	*
San Sebastian	32.0	26.3	41.6	0.0	30.4	26.6	42.9	0.1	
Asturias	32.6	32.4	34.7	0.3	28.3	33.5	38.2	0.1	
Italy	**28.5**	**42.6**	**26.5**	**2.4**	**25.6**	**39.4**	**29.0**	**5.9**	-
Ragusa	23.5	44.3	31.7	0.6	22.0	37.4	34.9	5.7	
Naples	-	-	-	-	-	-	-	-	-
Florence	27.0	44.7	25.4	3.0	25.0	42.9	27.8	4.3	
Turin	27.5	42.5	26.9	3.1	26.9	37.2	27.1	8.7	
Varese	34.9	39.8	23.7	1.5	27.9	42.3	28.1	1.7	
France	-	-	-	-	-	-	-	-	-
South coast	-	-	-	-	-	-	-	-	-
South	-	-	-	-	-	-	-	-	-
North-west	-	-	-	-	-	-	-	-	-
North-east	-	-	-	-	-	-	-	-	-
Germany	**30.3**	**46.2**	**23.4**	**0.0**	**30.1**	**45.1**	**24.7**	**0.0**	-
Heidelberg	32.7	44.4	23.0	0.0	29.7	44.3	26.0	0.0	
Potsdam	28.4	47.8	23.8	0.0	30.5	46.1	23.3	0.1	
The Netherlands	**24.3**	**38.0**	**37.1**	**0.6**	**25.1**	**36.2**	**38.5**	**0.2**	-
Bilthoven	24.3	38.0	37.1	0.6	25.1	36.2	38.5	0.2	
Utrecht	-	-	-	-	-	-	-	-	-
United Kingdom	**39.4**	**41.9**	**17.2**	**1.5**	**38.5**	**41.4**	**15.2**	**4.9**	-
Gen. Population	38.1	42.7	17.2	2.0	31.9	43.3	17.3	7.5	
Health Conscious	40.5	36.0	23.6	0.0	50.7	37.6	11.3	0.3	
Denmark	**29.0**	**37.9**	**32.3**	**0.7**	**25.2**	**36.6**	**37.4**	**0.8**	-
Copenhagen	29.5	38.4	31.6	0.5	25.1	36.0	38.0	0.9	***
Aarhus	28.5	36.9	33.4	1.2	25.6	38.0	36.0	0.4	
Sweden	**45.7**	**32.5**	**20.7**	**1.1**	**38.7**	**35.5**	**24.6**	**1.1**	-
Malmö	30.0	43.0	27.0	0.0	27.8	43.3	28.7	0.1	
Umeå	57.0	24.9	15.8	2.3	51.6	26.3	19.8	2.2	**
Norway	-	-	-	-	-	-	-	-	-
South & East	-	-	-	-	-	-	-	-	-
North & West	-	-	-	-	-	-	-	-	-

[a]*P* value: * <0.05; ** <0.01; *** <0.001..

Table 5a. Comparison of education level, adjusted for age, of the EPIC calibration sub-populations and the rest of the EPIC cohorts: Women

	Education level												
	Calibration						Cohort						
	Incomplete primary school %	Primary school %	Technical school %	Secondary school %	University %	Missing %	Incomplete primary school %	Primary school %	Technical school %	Secondary school %	University %	Missing %	P value[a]
Greece	**20.9**	**32.2**	**7.5**	**22.6**	**16.2**	**0.7**	**27.9**	**37.9**	**5.6**	**16.8**	**10.4**	**1.3**	*****
Spain	**28.6**	**41.8**	**6.9**	**7.4**	**10.0**	**5.3**	**35.9**	**39.0**	**5.2**	**5.2**	**9.1**	**5.7**	**-**
Granada	38.7	33.2	4.7	5.5	6.2	11.7	47.1	25.6	3.7	4.1	8.9	10.6	**
Murcia	40.9	22.4	4.9	6.2	16.0	9.5	45.8	26.0	3.4	3.8	11.3	9.6	**
Navarra	15.7	62.3	7.5	4.1	9.7	0.7	29.3	50.7	5.7	5.2	7.4	1.8	***
San Sebastian	20.7	47.7	12.2	10.8	7.5	1.2	22.9	53.2	9.4	5.4	7.7	1.4	*
Asturias	24.0	47.1	6.2	10.1	10.2	2.5	28.0	47.6	5.1	8.0	9.1	2.1	
Italy	**0.9**	**24.0**	**10.9**	**49.4**	**14.5**	**0.2**	**1.8**	**28.1**	**10.9**	**44.7**	**13.0**	**1.6**	**-**
Ragusa	0.6	28.9	5.7	49.7	15.1	0.0	3.7	21.6	4.4	50.2	17.2	2.9	
Naples	1.6	18.8	9.3	51.9	18.3	0.0	2.8	24.4	6.8	49.5	16.5	0.0	
Florence	1.0	20.8	10.2	49.5	18.0	0.5	1.1	26.1	12.1	42.8	17.0	0.8	***
Turin	1.0	18.3	13.2	55.8	11.7	0.0	1.5	20.7	13.7	47.6	10.1	6.4	
Varese	0.6	32.3	12.7	44.2	10.1	0.1	1.4	37.8	12.7	40.9	6.9	0.4	**
France	**0.4**	**9.5**	**0.0**	**46.5**	**39.8**	**3.8**	**0.4**	**10.9**	**0.0**	**49.1**	**35.5**	**4.1**	**-**
South coast	0.4	10.4	0.0	47.1	38.2	3.9	0.4	10.7	0.0	48.3	36.3	4.3	
South	0.4	8.0	0.0	47.9	38.9	4.8	0.3	10.7	0.0	50.1	34.9	4.0	**
North-west	0.2	14.5	0.0	52.2	30.7	2.4	0.5	11.9	0.0	52.6	30.5	4.4	
North-east	0.5	10.4	0.0	45.2	40.3	3.6	0.4	10.8	0.0	47.4	37.3	4.0	**

[a]*P* value: * <0.05; ** <0.01; *** <0.001. [b]For the Malmö cohort, which existed before EPIC, 7332 subjects were defined differently from the EPIC variables and were classified in the closest EPIC category corresponding to "technical school".

Table 5a. (Continued)

	Education level												
	Calibration						Cohort						
	Incom-plete primary school	Primary school	Techni-cal school	Secon-dary school	Univer-sity	Missing	Incom-plete primary school	Primary school	Techni-cal school	Secon-dary school	Univer-sity	Missing	*P* value[a]
	%	%	%	%	%	%	%	%	%	%	%	%	
Germany	**0.6**	**23.5**	**41.6**	**8.5**	**25.6**	**0.2**	**0.6**	**23.4**	**41.5**	**7.9**	**26.5**	**0.1**	**-**
Heidelberg	0.3	25.1	41.5	9.0	23.8	0.2	0.6	26.8	39.8	8.4	24.2	0.1	
Potsdam	0.7	20.5	44.0	7.2	27.4	0.2	0.6	20.8	42.8	7.4	28.3	0.1	
The Netherlands	**0.0**	**19.6**	**33.1**	**29.5**	**17.6**	**0.2**	**0.0**	**20.2**	**31.8**	**29.9**	**17.2**	**1.0**	**-**
Bilthoven	0.0	13.2	34.9	28.5	22.9	0.5	0.0	16.7	34.6	28.6	19.1	1.0	**
Utrecht	0.0	22.6	31.8	30.4	15.1	0.1	0.0	22.1	30.3	30.4	16.2	0.9	
United Kingdom	**0.0**	**12.8**	**31.0**	**9.8**	**26.2**	**20.2**	**0.0**	**14.4**	**27.3**	**10.0**	**26.4**	**22.0**	**-**
Gen. Population	0.0	17.9	34.9	8.4	17.4	21.4	0.0	29.8	30.5	6.7	12.6	20.4	***
Health Conscious	0.0	0.0	17.6	14.5	51.1	16.8	0.0	0.0	25.0	12.8	39.0	23.2	**
Denmark	**0.0**	**27.2**	**47.8**	**13.5**	**11.4**	**0.0**	**0.0**	**31.8**	**46.2**	**11.6**	**10.1**	**0.2**	**-**
Copenhagen	0.0	26.0	47.8	13.8	12.5	0.0	0.0	29.1	47.4	11.9	11.3	0.3	*
Aarhus	0.0	31.3	47.4	12.8	8.4	0.0	0.0	38.2	43.4	10.9	7.5	0.1	*
Sweden	**0.5**	**27.9**	**29.2**	**13.0**	**29.0**	**0.3**	**0.5**	**37.1**	**27.5**	**12.1**	**22.4**	**0.4**	**-**
Malmö	0.9	33.3	32.3	6.6	26.7	0.2	0.8	39.1	30.0	7.0	22.8	0.3	***
Umeå	0.0	18.8	24.2	24.0	32.6	0.4	0.0	34.0	23.4	20.2	21.7	0.7	***
Norway	**0.8**	**17.4**	**32.9**	**34.6**	**14.3**	**0.0**	**1.5**	**22.0**	**36.0**	**28.2**	**12.2**	**0.0**	**-**
South & East	0.6	16.0	33.0	35.6	14.9	0.0	1.4	19.1	36.3	30.1	13.1	0.0	***
North & West	1.1	20.1	32.7	32.6	13.4	0.0	1.6	25.6	35.6	26.0	11.2	0.0	***

[a]*P* value: * <0.05; ** <0.01; *** <0.001. [b]For the Malmö cohort, which existed before EPIC, 7332 subjects were defined differently from the EPIC variables and were classified in the closest EPIC category corresponding to "technical school".

Table 5b. Comparison of education level, adjusted for age, of the EPIC calibration sub-populations and the rest of the EPIC cohorts: Men

	Education level												
	Calibration						Cohort						
	Incom-plete primary school %	Primary school %	Techni-cal school %	Secon-dary school %	Univer-sity %	Missing %	Incom-plete primary school %	Primary school %	Techni-cal school %	Secon-dary school %	Univer-sity %	Missing %	*P* value[a]
Greece	**12.0**	**30.3**	**17.4**	**15.5**	**24.3**	**0.5**	**14.5**	**38.0**	**17.1**	**13.2**	**13.7**	**3.5**	*******
Spain	**19.7**	**38.6**	**15.5**	**9.0**	**14.2**	**2.9**	**24.8**	**37.1**	**12.4**	**7.7**	**14.7**	**3.4**	**-**
Granada	21.4	22.8	6.4	9.9	33.0	6.5	30.8	25.0	4.9	8.8	19.8	10.8	***
Murcia	33.3	26.2	6.1	5.6	19.8	9.1	29.4	27.1	6.8	7.9	22.1	6.6	
Navarra	17.5	46.2	15.6	10.9	8.9	0.9	25.7	42.1	11.1	8.2	11.7	1.2	***
San Sebastian	17.5	42.5	21.6	8.0	9.0	1.4	20.1	43.4	19.0	6.0	9.8	1.7	
Asturias	18.0	39.4	15.3	9.7	16.3	1.3	22.3	37.9	14.7	8.5	15.3	1.4	
Italy	**0.4**	**15.0**	**18.6**	**53.1**	**12.6**	**0.3**	**0.5**	**16.3**	**14.0**	**51.6**	**14.1**	**3.5**	**-**
Ragusa	1.8	19.6	9.8	49.5	18.7	0.6	1.1	19.0	8.1	50.2	17.1	4.4	
Naples	-	-	-	-	-	-	-	-	-	-	-	-	-
Florence	0.2	13.3	17.2	50.8	17.7	0.7	0.2	17.3	11.2	55.2	14.1	2.0	*
Turin	0.3	11.8	21.5	55.0	11.3	0.1	0.3	12.1	17.1	50.3	14.8	5.4	**
Varese	0.0	21.7	19.4	52.8	5.8	0.3	0.4	20.9	18.3	51.4	8.8	0.3	
France	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
South coast	-	-	-	-	-	-	-	-	-	-	-	-	-
South	-	-	-	-	-	-	-	-	-	-	-	-	-
North-west	-	-	-	-	-	-	-	-	-	-	-	-	-
North-east	-	-	-	-	-	-	-	-	-	-	-	-	-

[a] *P* value: * <0.05; ** <0.01; *** <0.001. [b] For the Malmö cohort, which existed before EPIC, 7332 subjects were defined differently from the EPIC variables and were classified in the closest EPIC category corresponding to "technical school".

Table 5b. (Continued)

	Education level												
	Calibration						Cohort						
	Incom-plete primary school	Primary school	Techni-cal school	Secon-dary school	Univer-sity	Missing	Incom-plete primary school	Primary school	Techni-cal school	Secon-dary school	Univer-sity	Missing	*P* value[a]
	%	%	%	%	%	%	%	%	%	%	%	%	
Germany	**0.5**	**24.3**	**28.7**	**4.7**	**41.7**	**0.1**	**0.6**	**24.4**	**27.5**	**5.2**	**42.2**	**0.1**	**-**
Heidelberg	0.5	31.7	29.9	5.3	32.3	0.3	0.6	30.8	26.7	5.5	36.4	0.1	
Potsdam	0.5	17.4	28.3	4.1	49.7	0.0	0.6	17.2	28.4	4.9	48.8	0.1	
The Netherlands	**0.0**	**12.5**	**41.3**	**17.6**	**28.3**	**0.3**	**0.0**	**14.0**	**40.0**	**18.2**	**27.1**	**0.7**	**-**
Bilthoven	0.0	12.5	41.3	17.6	28.3	0.3	0.0	14.0	40.0	18.2	27.1	0.7	
Utrecht	-	-	-	-	-	-	-	-	-	-	-	-	-
United Kingdom	**0.0**	**8.5**	**31.4**	**10.5**	**30.5**	**19.1**	**0.0**	**16.6**	**29.3**	**9.7**	**28.3**	**16.0**	**-**
Gen. Population	0.0	12.1	36.0	10.8	23.8	17.3	0.0	25.1	36.5	8.8	15.8	13.9	***
Health Conscious	0.0	0.0	17.5	7.1	50.0	25.4	0.0	0.0	16.4	11.5	52.3	19.8	
Denmark	**0.0**	**30.4**	**30.2**	**9.3**	**30.1**	**0.1**	**0.0**	**35.1**	**29.1**	**7.6**	**27.9**	**0.3**	**-**
Copenhagen	0.0	29.0	30.0	9.3	31.7	0.1	0.0	33.0	29.3	8.0	29.4	0.4	**
Aarhus	0.0	34.0	30.5	9.1	26.4	0.0	0.0	39.8	28.8	6.8	24.5	0.1	*
Sweden	**0.1**	**33.2**	**26.1**	**16.5**	**23.8**	**0.3**	**0.5**	**40.7**	**22.6**	**15.9**	**19.8**	**0.5**	**-**
Malmö	0.1	43.2	20.6	12.0	23.9	0.2	0.9	45.5	19.5	11.7	22.2	0.3	*
Umeå	0.0	24.8	27.0	23.7	24.1	0.3	0.0	35.5	26.4	20.5	16.9	0.7	***
Norway	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
South & East	-	-	-	-	-	-	-	-	-	-	-	-	-
North & West	-	-	-	-	-	-	-	-	-	-	-	-	-

[a]*P* value: * <0.05; ** <0.01; *** <0.001. [b]For the Malmö cohort, which existed before EPIC, 7332 subjects were defined differently from the EPIC variables and were classified in the closest EPIC category corresponding to "technical school".

Table 6a. Comparison of work-based physical activity, adjusted for age, of the EPIC calibration sub-populations and the rest of the EPIC cohorts: women

	Work-based physical activity												
	Calibration						Cohort						*P* value[b]
	Non-worker	Sedentary occupation	Standing occupation	Manual work	Heavy manual work	Missing	Non-worker	Sedentary occupation	Standing occupation	Manual work	Heavy manual work	Missing	
	%	%	%	%	%	%	%	%	%	%	%	%	
Greece	**58.5**	**19.1**	**16.8**	**5.1**	**0.0**	**0.5**	**61.0**	**14.4**	**13.3**	**8.8**	**0.1**	**2.5**	***
Spain	-	**15.0**	**80.2**	**1.7**	**0.2**	**0.0**	**2.7**	**12.7**	**82.6**	**1.9**	**0.1**	**2.8**	-
Granada	-	8.3	87.8	0.7	0.0	0.0	1.0	11.9	85.9	1.2	0.1	3.2	
Murcia	-	17.9	78.1	2.0	1.1	0.0	1.9	14.0	82.0	1.8	0.3	1.0	
Navarra	-	13.4	81.5	2.1	0.0	0.0	3.6	10.5	84.2	1.6	0.1	2.9	
San Sebastian	-	19.7	76.3	3.0	0.0	0.0	0.5	15.7	81.3	2.5	0.0	1.0	
Asturias	-	14.7	78.6	1.2	0.3	0.0	6.8	11.3	79.1	2.8	0.1	5.3	
Italy	**48.6**	**29.4**	**13.1**	**6.1**	**2.2**	**0.6**	**51.1**	**27.4**	**12.5**	**5.3**	**1.9**	**1.7**	-
Ragusa	41.6	33.1	19.8	2.3	3.2	0.0	43.2	31.5	17.1	3.2	2.1	2.9	
Naples	52.2	32.6	7.1	5.9	2.2	0.0	59.0	27.2	5.0	6.4	2.4	0.0	
Florence	46.6	31.9	14.5	4.8	1.6	0.6	46.0	31.9	14.3	5.2	1.5	1.1	
Turin	51.5	23.0	15.2	7.0	3.2	0.0	49.7	25.2	11.7	4.0	2.5	6.8	
Varese	49.6	27.3	12.0	8.0	1.9	1.1	55.9	22.4	13.2	6.3	1.8	0.5	***
France	**27.4**	**17.4**	**48.6**	**1.8**	**0.0**	**4.9**	**31.7**	**16.4**	**44.8**	**2.0**	**0.0**	**5.1**	-
South coast	33.6	14.5	47.5	1.2	0.0	3.3	35.9	15.2	41.5	2.0	0.0	5.4	***
South	28.8	16.3	48.3	1.7	0.0	4.9	32.8	15.4	44.9	2.1	0.0	4.7	***
North-west	30.2	14.8	47.4	2.1	0.0	5.5	34.0	13.7	45.5	2.2	0.0	4.6	***
North-east	25.5	19.1	48.1	2.2	0.0	5.1	29.0	18.2	45.5	2.0	0.0	5.3	***

[a]In Spain, all participants were classified in one of the categories of work activity independently from the employment status.

[b]*P* value: * <0.05; ** <0.01; *** <0.001. m: missing.

Table 6a. (Continued)

	Work-based physical activity												
	Calibration						Cohort						
	Non-worker	Sedentary occupation	Standing occupation	Manual work	Heavy manual work	Missing	Non-worker	Sedentary occupation	Standing occupation	Manual work	Heavy manual work	Missing	*P* value[b]
	%	%	%	%	%	%	%	%	%	%	%	%	
Germany	**34.8**	**37.4**	**24.7**	**2.4**	**0.1**	**0.6**	**32.5**	**40.1**	**23.9**	**2.9**	**0.1**	**0.5**	**-**
Heidelberg	34.3	35.1	26.2	3.2	0.2	1.0	33.8	36.1	25.7	3.3	0.1	1.0	
Potsdam	34.2	40.9	23.3	1.3	0.2	0.1	31.4	43.3	22.5	2.5	0.1	0.2	**
The Netherlands	**49.5**	**17.5**	**16.0**	**9.3**	**4.9**	**2.9**	**48.0**	**17.7**	**14.3**	**8.3**	**4.9**	**6.8**	**-**
Bilthoven	45.4	19.5	16.1	8.0	3.2	7.9	47.8	15.9	10.0	5.0	2.6	18.7	***
Utrecht	50.9	16.7	16.2	10.3	5.9	0.1	48.5	18.1	16.4	10.0	6.1	0.9	
United Kingdom	**37.9**	**32.0**	**20.9**	**7.3**	**0.0**	**2.0**	**39.8**	**25.8**	**20.2**	**6.1**	**0.2**	**7.9**	**-**
Gen. Population	41.9	29.0	19.2	7.6	0.0	2.3	44.1	16.7	15.9	8.0	0.1	15.2	***
Health Conscious	38.2	32.3	24.6	4.0	0.0	1.0	35.9	33.3	23.8	4.6	0.3	2.0	
Denmark	**30.8**	**33.4**	**16.7**	**17.9**	**1.2**	**0.1**	**27.2**	**32.3**	**17.6**	**20.6**	**1.6**	**0.7**	**-**
Copenhagen	30.5	35.3	15.4	17.3	1.5	0.1	25.2	35.3	16.8	20.2	1.6	1.0	***
Aarhus	32.6	27.7	20.1	19.4	0.2	0.0	31.9	25.4	19.6	21.5	1.6	0.1	
Sweden	**24.8**	**27.6**	**32.6**	**12.3**	**2.1**	**0.6**	**33.4**	**23.3**	**28.0**	**12.6**	**2.1**	**0.6**	**-**
Malmö	34.5	31.1	28.3	5.7	0.1	0.4	40.9	27.2	26.3	5.0	0.2	0.4	***
Umeå	13.0	20.4	42.7	19.1	4.0	0.9	20.3	17.2	31.2	25.3	5.1	0.9	***
Norway	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
South & East	m	m	m	m	m	m	m	m	m	m	m	m	m
North & West	m	m	m	m	m	m	m	m	m	m	m	m	m

[a]In Spain, all participants were classified in one of the categories of work activity independently from the employment status.

[b]*P* value: * <0.05; ** <0.01; *** <0.001.

m: missing.

Table 6b. Comparison of work-based physical activity, adjusted for age, of the EPIC calibration sub-populations and the rest of the EPIC cohorts: men

	Work-based physical activity												
	Calibration						Cohort						
	Non-worker	Sedentary occupation	Standing occupation	Manual work	Heavy manual work	Missing	Non-worker	Sedentary occupation	Standing occupation	Manual work	Heavy manual work	Missing	*P* value[b]
	%	%	%	%	%	%	%	%	%	%	%	%	
Greece	**26.2**	**27.5**	**20.8**	**20.1**	**1.7**	**3.7**	**27.1**	**26.3**	**17.0**	**22.2**	**2.0**	**5.5**	***
Spain[a]	-	**33.7**	**31.0**	**25.1**	**5.9**	**0.0**	**4.8**	**34.5**	**34.1**	**21.1**	**5.5**	**4.3**	-
Granada	-	40.3	32.1	17.6	6.5	0.0	4.3	36.4	37.3	14.3	7.7	3.5	
Murcia	-	33.6	37.4	13.4	10.5	0.0	6.2	38.8	33.7	13.0	8.4	5.1	
Navarra	-	31.6	32.7	28.8	4.0	0.0	3.9	33.9	33.6	24.1	4.5	2.9	
San Sebastian	-	34.1	26.6	35.4	3.2	0.0	1.2	32.6	31.0	31.3	3.8	0.8	
Asturias	-	32.4	34.2	16.4	8.8	0.0	10.1	32.3	37.2	15.3	5.0	8.3	
Italy	**19.4**	**40.8**	**21.1**	**12.5**	**5.1**	**1.0**	**19.6**	**36.8**	**21.4**	**12.8**	**5.1**	**4.3**	-
Ragusa	8.8	41.5	25.0	10.5	12.9	1.2	11.3	35.1	27.0	11.8	9.7	5.1	
Naples	-	-	-	-	-	-	-	-	-	-	-	-	-
Florence	15.3	45.9	19.5	13.7	4.6	1.1	18.4	41.1	19.8	13.2	5.5	2.0	
Turin	21.6	39.9	22.2	11.0	3.9	1.3	21.2	38.0	19.4	11.9	2.9	6.6	
Varese	27.6	36.3	17.3	15.3	3.2	0.3	28.5	30.3	21.3	15.3	3.9	0.7	
France	-	-	-	-	-	-	-	-	-	-	-	-	-
South coast	-	-	-	-	-	-	-	-	-	-	-	-	-
South	-	-	-	-	-	-	-	-	-	-	-	-	-
North-west	-	-	-	-	-	-	-	-	-	-	-	-	-
North-east	-	-	-	-	-	-	-	-	-	-	-	-	-

[a]In Spain, all participants were classified in one of the categories of work activity independently from the employment status.
[b]*P* value: * <0.05; ** <0.01; *** <0.001.

Table 6b. (Continued)

	Work-based physical activity												
	Calibration						Cohort						
	Non-worker	Sedentary occupation	Standing occupation	Manual work	Heavy manual work	Missing	Non-worker	Sedentary occupation	Standing occupation	Manual work	Heavy manual work	Missing	*P* value[b]
	%	%	%	%	%	%	%	%	%	%	%	%	
Germany	**28.4**	**43.3**	**22.0**	**5.6**	**0.7**	**0.0**	**25.3**	**45.0**	**21.8**	**7.1**	**0.7**	**0.1**	**-**
Heidelberg	23.5	47.4	22.8	5.4	0.9	0.0	21.4	48.5	22.2	6.9	0.9	0.1	
Potsdam	32.7	39.8	21.2	5.8	0.5	0.0	29.8	41.0	21.3	7.3	0.5	0.1	
The Netherlands	**20.0**	**37.1**	**16.5**	**9.0**	**8.4**	**8.9**	**20.7**	**27.7**	**11.4**	**8.2**	**6.0**	**26.0**	**-**
Bilthoven	20.0	37.1	16.5	9.0	8.4	8.9	20.7	27.7	11.4	8.2	6.0	26.0	***
Utrecht	-	-	-	-	-	-	-	-	-	-	-	-	-
United Kingdom	**35.0**	**30.7**	**13.2**	**14.6**	**3.4**	**3.1**	**35.0**	**28.8**	**12.6**	**12.8**	**3.6**	**7.2**	**-**
Gen. Population	37.6	27.3	14.1	14.0	3.8	3.2	37.2	21.9	10.7	15.4	4.6	10.2	
Health Conscious	29.9	41.8	6.9	18.3	0.4	2.6	30.5	41.6	16.1	8.2	1.8	1.7	*
Denmark	**16.6**	**42.1**	**16.8**	**18.2**	**6.3**	**0.0**	**15.3**	**39.1**	**16.8**	**20.1**	**8.1**	**0.6**	**-**
Copenhagen	16.3	43.3	16.4	17.7	6.2	0.0	14.8	40.7	15.9	19.8	8.0	0.8	*
Aarhus	17.9	38.9	17.8	19.3	6.1	0.0	16.3	35.6	18.8	20.9	8.4	0.0	
Sweden	**29.0**	**23.4**	**27.2**	**17.3**	**2.4**	**0.6**	**33.3**	**23.5**	**25.6**	**14.0**	**3.0**	**0.6**	**-**
Malmö	43.8	29.9	16.7	7.8	1.7	0.1	42.6	27.6	20.2	6.8	2.3	0.4	**
Umeå	18.8	20.6	33.7	22.4	3.3	1.1	22.1	18.7	32.1	22.6	3.8	0.7	***
Norway	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
South & East	-	-	-	-	-	-	-	-	-	-	-	-	-
North & West	-	-	-	-	-	-	-	-	-	-	-	-	-

[a]In Spain, all participants were classified in one of the categories of work activity independently from the employment status.
[b]*P* value: * <0.05; ** <0.01; *** <0.001.

Table 7. Number (%) and type of special diets reported in the EPIC calibration sub-populations

	Women						Men					
	Reported special diet		Type of special diet reported				Reported special diet		Type of special diet reported			
			Obesity	Vegetarian[a]	Diseases	Other			Obesity	Vegetarian[a]	Diseases	Other
	n	%	%	%	%	%	n	%	%	%	%	%
Greece	**424**	**29.3**	**4.1**	**0.4**	**20.2**	**4.6**	**382**	**27.8**	**1.2**	**0.3**	**21.0**	**5.3**
Spain	**306**	**20.5**	**5.6**	**0.1**	**12.1**	**2.7**	**300**	**16.4**	**1.5**	**0.1**	**12.2**	**2.7**
Granada	75	23.4	3.8	0.0	15.0	4.7	60	26.4	2.6	0.4	18.5	4.9
Murcia	49	15.5	5.7	0.3	7.3	2.2	24	9.8	1.2	0.0	6.5	2.0
Navarra	52	19.0	5.5	0.0	11.7	1.8	95	21.0	1.1	0.0	18.4	1.6
San Sebastian	24	9.8	2.9	0.0	4.5	2.5	30	6.1	0.8	0.0	2.4	2.8
Asturias	106	31.0	9.1	0.0	19.6	2.3	91	22.2	2.4	0.0	16.9	2.9
Italy	**434**	**16.9**	**6.3**	**0.2**	**7.5**	**3.0**	**134**	**9.2**	**1.7**	**0.3**	**5.8**	**1.4**
Ragusa	11	8.0	4.4	0.0	0.7	2.9	7	4.1	1.2	0.0	1.8	1.2
Naples	96	23.4	11.7	0.5	8.5	2.7	-	-	-	-	-	-
Florence	136	16.9	6.3	0.4	8.7	1.5	32	11.6	1.5	0.7	8.3	1.1
Turin	45	11.4	4.8	0.0	4.3	2.3	54	7.9	2.2	0.4	4.3	1.0
Varese	146	18.0	4.6	0.0	8.4	5.1	41	12.5	1.2	0.0	8.8	2.4
France	**1129**	**23.8**	**5.7**	**0.4**	**14.4**	**3.3**	-	-	-	-	-	-
South coast	176	28.0	7.6	0.2	17.0	3.2	-	-	-	-	-	-
South	251	17.8	3.1	0.4	10.8	3.5	-	-	-	-	-	-
North-west	213	32.6	3.8	0.3	22.6	5.8	-	-	-	-	-	-
North-east	489	23.8	7.6	0.6	13.3	2.3	-	-	-	-	-	-

[a] *"Vegetarian" includes vegans, ovo-lacto vegetarians and fish eaters (no meat eaters).*

Table 7. (Continued)

	Women						Men					
	Reported special diet		Type of special diet reported				Reported special diet		Type of special diet reported			
			Obesity	Vegetarian[a]	Diseases	Other			Obesity	Vegetarian[a]	Diseases	Other
	n	%	%	%	%	%	n	%	%	%	%	%
Germany	**432**	**19.7**	**3.0**	**2.1**	**13.5**	**1.1**	**528**	**22.4**	**2.2**	**1.0**	**18.2**	**1.0**
Heidelberg	245	22.1	2.8	3.6	15.3	0.4	240	22.5	2.1	2.1	18.2	0.2
Potsdam	187	17.2	3.1	0.6	11.7	1.8	288	22.2	2.2	0.2	18.2	1.6
The Netherlands	**680**	**22.4**	**5.9**	**0.5**	**11.1**	**4.9**	**72**	**7.0**	**1.2**	**0.0**	**3.8**	**2.0**
Bilthoven	129	11.7	3.2	0.5	5.0	3.1	72	7.0	1.2	0.0	3.8	2.0
Utrecht	551	28.4	7.4	0.5	14.5	6.0	-	-	-	-	-	-
United Kingdom	**318**	**39.4**	**2.6**	**17.2**	**12.4**	**7.2**	**158**	**29.7**	**0.9**	**18.2**	**7.7**	**2.8**
Gen. Population	155	26.3	2.7	1.5	13.2	8.8	56	13.6	0.7	2.0	8.3	2.7
Health Conscious	163	75.1	2.3	59.9	10.1	2.8	102	84.3	1.7	73.6	5.8	3.3
Denmark	**242**	**12.0**	**2.1**	**0.2**	**8.8**	**1.0**	**146**	**7.6**	**0.9**	**0.0**	**6.3**	**0.3**
Copenhagen	200	13.3	1.5	0.2	10.4	1.3	123	9.0	0.8	0.0	7.8	0.4
Aarhus	42	8.2	4.1	0.0	3.9	0.2	23	4.1	1.2	0.0	2.8	0.0
Sweden	**1073**	**31.2**	**0.8**	**0.8**	**22.4**	**7.2**	**729**	**25.5**	**0.7**	**0.2**	**19.0**	**5.5**
Malmö	382	22.0	0.1	0.6	11.8	9.5	242	16.9	0.4	0.4	8.5	7.6
Umeå	691	40.6	1.5	1.0	33.3	4.9	487	34.1	1.1	0.1	29.6	3.4
Norway	**359**	**19.2**	**1.5**	**0.8**	**13.5**	**3.4**	-	-	-	-	-	-
South & East	233	19.7	1.7	1.1	13.3	3.6	-	-	-	-	-	-
North & West	126	18.4	1.2	0.3	13.7	3.2	-	-	-	-	-	-

[a] *"Vegetarian" includes vegans, ovo-lacto vegetarians and fish eaters (no meat eaters).*

Logistics and methodological issues of the 24-hour dietary recall method

Some of the characteristics of the reference dietary calibration method used are reported in **Table 8**.

a. Time interval between baseline dietary assessment and 24-HDR measurements

The time interval between dietary measurements varies from 1 day (or a few days) to several years. In the Netherlands, Germany and Denmark, most of the interviews were conducted at the same time as the baseline examination or shortly afterwards. In the UK, Italy, Greece and Spain, the interval between the two dietary measurements was between 12.4 and 25 months whereas in France, for logistic reasons, and Sweden, where the cohort existed before joining EPIC, the interval was as high as between 31.7 and 34.1 months.

b. Duration of the recalled day

The period to be covered during the recalled dietary interview was defined as the individual's time between waking up on the recalled day to waking up on the following day (interview day). This procedure was chosen instead of the time period from midnight to midnight to facilitate memory retrieval during the interview. Whatever the centre or country, the mean time interval was always about 24 hours.

c. Interview duration

The average duration of the 24-HDR interviews was 31.1 ± 13.3 minutes, and ranged from 27 ± 11 in Denmark to 39 ± 20 in the UK. The variations observed across centres may be explained by differences in dietary habits, total number of food items reported and the proportion of mixed recipes, which usually require more time to process than single food items. In the UK, the average duration was almost twice as high as the mean shortest interview time (22 ± 7 minutes in Ragusa), particularly among the "health-conscious" study population 43 ± 28 minutes).

Table 8. Characteristics of the 24-hour dietary recall measurements

	n	Time interval between dietary measurements (months)[a]		Time period recalled[b] *(hours)*		Interview duration (minutes)		Number of food items reported[c]	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Greece	**2686**	**23.7**	**13.7**	**24.1**	**0.8**	**27**	**16**	**15.3**	**5.3**
Spain	**3220**	**24.9**	**10.4**	**24.1**	**1.2**	**31**	**10**	**26.2**	**7.2**
Granada	514	27.0	12.9	24.0	0.8	31	11	27.2	6.9
Murcia	547	21.3	8.7	24.0	1.1	26	8	26.5	7.3
Navarra	715	31.1	7.1	24.1	1.3	34	9	25.3	6.2
San Sebastian	734	19.5	9.1	24.1	1.4	30	10	28.5	7.1
Asturias	710	25.6	9.9	24.1	1.1	34	10	23.7	7.4
Italy	**3956**	**20.1**	**12.4**	**24.1**	**0.9**	**29**	**11**	**25.0**	**6.9**
Ragusa	306	21.3	10.0	24.0	0.9	22	7	22.2	6.0
Naples	403	34.2	11.7	24.0	0.9	27	10	21.3	5.9
Florence	1056	16.9	15.7	24.0	0.9	28	10	24.6	6.1
Turin	1069	19.4	9.2	24.1	1.1	33	10	28.7	7.5
Varese	1122	18.6	8.5	24.1	0.9	30	13	23.9	6.2
France	**4639**	**31.7**	**9.3**	**24.1**	**0.9**	**29**	**11**	**27.2**	**6.7**
South coast	612	32.1	8.1	24.1	0.8	33	13	27.5	7.1
South	1396	27.5	8.9	24.1	0.9	26	8	26.2	6.3
North-west	622	33.2	8.9	24.1	0.9	32	9	28.5	6.8
North-east	2009	34.1	8.9	24.1	0.9	30	11	27.3	6.7

[a]Time interval between baseline dietary assessment and 24-hour diet recall measurements.
[b]*Mean time period covered from wake-up on the recalled day to wake-up on the following day.*
[c]Food + recipe items.

Table 8. (Continued)

	n	Time interval between dietary measurements (months)[a]		Time period recalled[b] *(hours)*		Interview duration (minutes)		Number of food items reported[c]	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Germany	**4418**	**0.5**	**1.8**	**24.0**	**1.1**	**35**	**17**	**23.2**	**6.3**
Heidelberg	2120	1.0	2.1	24.0	1.2	35	15	23.4	6.6
Potsdam	2298	-0.1	1.2	24.1	1.1	36	18	23.0	6.1
The Netherlands	**3984**	**0.5**	**1.5**	**24.1**	**1.2**	**33**	**12**	**27.7**	**7.7**
Bilthoven	2110	0.2	0.5	24.0	1.3	31	13	26.7	7.7
Utrecht	1874	0.9	2.1	24.2	1.0	35	10	28.9	7.6
United Kingdom	**1286**	**12.4**	**11.0**	**24.1**	**0.8**	**39**	**20**	**30.4**	**8.3**
General population	975	11.9	10.8	24.2	0.8	37	17	30.8	8.1
"Health conscious"	311	13.9	11.5	24.0	0.5	43	28	28.8	8.7
Denmark	**3918**	**0.1**	**0.5**	**24.0**	**1.2**	**27**	**11**	**24.1**	**7.0**
Copenhagen	2841	0.1	0.5	24.0	1.2	27	11	23.5	6.9
Aarhus	1077	0.1	0.4	24.0	1.2	25	11	25.5	7.0
Sweden	**6050**	**34.1**	**16.0**	**24.1**	**1.2**	**33**	**13**	**25.3**	**7.1**
Malmö	3132	30.0	14.9	24.0	1.1	31	12	24.9	7.2
Umeå	2918	38.5	15.9	24.1	1.3	35	14	25.7	7.0
Norway	**1798**	**12.8**	**3.5**	**24.0**	**1.6**	**30**	**12**	**23.3**	**6.4**
South & East	1136	12.8	3.5	24.0	1.6	31	13	23.5	6.6
North & West	662	12.9	3.5	24.0	1.6	30	11	22.9	6.1

[a]Time interval between baseline dietary assessment and 24-hour diet recall measurements.
[b]Mean time period covered from wake-up on the recalled day to wake-up on the following day.
[c]Food + recipe items.

d. Number of food items

The mean number of items reported per interview varied twofold across countries, from 15.3 ± 5.3 (Greece) to 30.4 ± 8.3 (UK). However, when these two countries are disregarded, the total number of food items reported varied much less 23.2 ± 6.3 in Germany to 27.2 ± 6.7 in France (women only). The variation between centres from the same countries was small except for Turin (Italy) and Asturias (Spain) where the number of food items reported was, respectively, higher (28.7 ± 7.5) and lower (23.7 ± 7.4) than in the other local centres.

e. Coverage of weekdays and seasons

The optimal coverage of weekdays, particularly Fridays and Saturdays, was restricted both by a low participation rate for interviews performed during weekends and by the logistic problems of approaching and interviewing the subjects during non-working days. **Table 9** shows that Fridays were highly under-represented in Germany (4.8%), Norway (7.7%), Denmark (8.3%), France (8.3%) and far below the expected 14.3% corresponding to an equal distribution of the seven weekdays. For Saturdays, the under-sampling was much lower than for Fridays, except in Denmark (Copenhagen mainly) where it was below 9%. This is probably because data concerning Saturdays were mostly collected on Mondays (i.e. during a working day) with a 48-hour time interval. In contrast, the interviews covering Fridays were always obtained during a weekend, Saturdays (24-hour interval) or Sundays (48-hour interval), which decreased the participation rate because of the logistic problems of interviewing subjects during non-working days. In the other countries, the same tendency to under-sample Fridays (and Saturdays) is observed, although to a lesser extent.

In certain countries, the interviews collected according to seasons tend to be under-sampled in summer and, to a lesser extent, autumn (Table 9). The 24-hour dietary recalls collected in summer were under-sampled by about 45–50% (Denmark, France and Norway) and up to 70% in Greece. In the other countries, both under-sampling and oversampling were observed but to a lesser degree. However, when the four seasons are grouped into two classes (spring/summer and autumn/winter), the coverage of seasons is much better balanced, except for Germany, where spring/summer tended to be over-sampled, whereas in Norway and Denmark it tended to be under-sampled by about 20–25%.

Table 9. Weekday and seasonal distribution obtained in the EPIC calibration sub-studies

	Weekday							Seasons			
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Spring	Summer	Autumn	Winter
	%	%	%	%	%	%	%	%	%	%	%
Greece	**14.4**	**14.5**	**15.4**	**15.0**	**12.6**	**13.1**	**15.0**	**36.1**	**6.5**	**25.9**	**31.5**
Spain	**16.4**	**15.8**	**16.0**	**12.5**	**12.4**	**13.2**	**13.6**	**28.5**	**26.5**	**20.6**	**24.5**
Granada	21.6	20.4	18.9	15.6	4.5	6.6	12.5	28.2	38.9	13.2	19.7
Murcia	14.4	15.0	14.4	14.1	14.1	14.6	13.4	28.3	21.6	25.1	25.1
Navarra	17.5	15.4	14.8	11.1	13.7	14.7	12.9	29.9	22.0	22.5	25.6
San Sebastian	14.6	14.2	15.8	12.5	13.8	14.4	14.7	25.6	32.6	16.8	25.1
Asturias	15.1	15.4	16.5	10.3	14.2	14.2	14.4	30.1	19.6	24.4	25.9
Italy	**16.0**	**16.5**	**16.4**	**15.6**	**11.2**	**10.2**	**14.3**	**30.5**	**18.1**	**25.4**	**26.0**
Ragusa	13.7	23.5	22.6	17.0	7.2	3.6	12.4	46.7	12.1	28.8	12.4
Naples	18.4	13.9	20.6	17.1	6.0	7.9	16.1	22.3	8.9	15.9	52.9
Florence	14.9	14.7	14.5	14.5	14.0	12.6	14.9	31.6	17.1	29.9	21.4
Turin	15.4	16.0	15.3	16.8	12.3	9.3	14.9	32.5	22.4	19.5	25.7
Varese	17.2	17.7	15.9	14.4	10.4	11.3	13.1	25.9	20.0	29.4	24.7
France	**18.6**	**18.3**	**17.1**	**15.5**	**8.3**	**11.2**	**11.1**	**34.5**	**13.0**	**22.7**	**29.8**
South coast	17.3	19.3	16.3	14.2	10.6	10.6	11.6	33.3	4.6	25.8	36.3
South	17.8	17.5	16.9	16.8	8.0	11.7	11.5	37.7	16.3	22.3	23.8
North-west	17.7	15.9	16.4	17.2	9.7	11.4	11.7	29.1	7.2	18.2	45.5
North-east	19.8	19.2	17.7	14.4	7.4	10.9	10.6	34.3	15.2	23.5	27.1
Germany	**20.2**	**21.6**	**17.6**	**13.2**	**4.8**	**11.1**	**11.6**	**29.7**	**33.1**	**16.1**	**21.1**
Heidelberg	22.2	21.4	17.3	14.5	1.2	11.0	12.4	22.0	39.3	17.6	21.1
Potsdam	18.3	21.8	17.8	12.1	8.1	11.1	10.8	36.9	27.4	14.7	21.1

Table 9. (Continued)

	Weekday							Seasons			
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Spring	Summer	Autumn	Winter
	%	%	%	%	%	%	%	%	%	%	%
The Netherlands	**15.2**	**15.3**	**14.7**	**13.6**	**13.0**	**13.8**	**14.5**	**24.6**	**30.6**	**22.8**	**22.0**
Bilthoven	15.9	16.5	14.7	13.3	11.4	13.7	14.6	26.3	28.9	24.3	20.6
Utrecht	14.4	13.9	14.8	13.9	14.8	13.9	14.3	22.8	32.6	21.1	23.5
United Kingdom	**17.0**	**15.0**	**16.5**	**14.7**	**11.3**	**11.8**	**13.7**	**31.7**	**21.4**	**24.5**	**22.4**
General population	15.3	13.3	15.7	15.3	12.9	12.4	15.1	31.4	21.9	25.1	21.6
"Health conscious"	22.5	20.3	19.0	12.9	6.1	10.0	9.3	32.8	19.9	22.5	24.8
Denmark	**19.5**	**21.4**	**16.6**	**14.9**	**8.3**	**8.7**	**10.8**	**22.8**	**12.7**	**24.6**	**40.0**
Copenhagen	18.5	23.2	16.7	15.2	8.2	7.7	10.4	25.8	12.9	21.2	40.1
Aarhus	22.0	16.5	16.3	14.0	8.5	11.1	11.7	14.9	12.0	33.4	39.7
Sweden	**15.1**	**15.2**	**14.7**	**14.6**	**12.1**	**14.1**	**14.3**	**26.3**	**24.2**	**19.4**	**30.2**
Malmö	15.8	15.5	14.8	14.4	10.5	14.1	14.9	24.2	18.7	25.2	31.8
Umeå	14.3	14.8	14.5	14.8	13.9	14.0	13.7	28.5	30.1	13.1	28.3
Norway	**16.9**	**17.2**	**17.0**	**12.3**	**7.7**	**11.4**	**17.5**	**25.8**	**13.5**	**30.3**	**30.5**
South & East	17.3	17.5	17.5	13.4	6.9	10.3	17.2	25.1	13.3	30.6	31.1
North & West	16.3	16.6	16.2	10.4	9.2	13.1	18.1	26.9	13.8	29.8	29.6

Discussion

The EPIC calibration sub-studies were set up in order to improve the comparability of dietary data across the participating centres. The calibration concept imposes a number of requirements, which include the following: 1) calibration sub-populations must be representative of the EPIC cohorts; 2) the common reference method for dietary intake assessment must provide correct estimates of mean population intakes; 3) random errors in the reference measurements, i.e. variations not structurally related to subjects' true intake levels, must be statistically independent of (i.e., not correlated with) random errors in the dietary questionnaire assessments used for the full cohort.

For the above requirements to be met, much depends on practical, logistic and methodological issues. In order to obtain the necessary representative population, a high participation rate must be achieved from the individuals invited to take part in the calibration sub-study. In our studies, about 70% of the study centres reported a participation rate above 75%. This response rate was consistently better when the subjects were recruited immediately after baseline examination than in centres where the subjects had to be re-invited at a later date. Important logistic constraints to re-approaching the subjects, as suggested by a high passive non-response rate, were observed in Greece and, to a lesser extent, in the Netherlands. It can be expected, however, that a passive non-response is random with regard to relevant subject characteristics.

Apart from study logistics, variations in the participation rates across study centres may also be partially explained by differences in social attitude and culture. In particular, the comparatively low response rates from representative samples of general populations (UK general population and Norway) or from an atypical population group (the "health-conscious" group from Oxford) suggest that a number of other uncontrolled factors may determine the subjects' participation rate, as observed in the SENECA study (26).

The sampling procedures for the calibration sub-studies were stratified by age group and gender and the sample size requirement was weighted by the expected numbers of cases of cancer in age–gender categories over 10 years of follow-up. This relative weighting will increase the precision of the statistical calibration procedure when it is used to correct relative risk estimates for biases induced by errors in the baseline dietary questionnaire assessments (7,9). Within strata of age and gender, however, the aim was to obtain the participation of a random, fully representative sample of cohort members in the calibration studies.

In most centres, after adjustment for age, no significant differences in height, weight, BMI and smoking status were observed between the calibration sample and the rest of the individual cohort for either men or women. Greater differences were observed for level of education and physical activity. The large and heterogeneous study populations involved in the analysis may explain the higher likelihood of detecting statistically significant differences. Indeed, the actual differences observed were modest in most centres. However, we observed a slight tendency to under-represent current and never-smokers compared to ex-smokers and subjects with a low education level in the calibration sample compared to the rest of the cohort. Non-workers were also under-represented compared to sedentary and/or standing professional occupations in the calibration compared to the cohort in both genders. A higher completeness (i.e. a lower number of missing values) of the calibration compared to the entire cohort dataset also explains some differences between distributions, particularly for work-based physical activity in Bilthoven and the UK general population. Although most of the discrepancies in the distributions of these categorical variables were due to differences of only a few percentage points across classes, this might also suggest a possible selection/sampling bias that should not be completely disregarded, particularly in certain study centres.

In order to investigate further whether the observed differences in subjects' characteristics influence dietary estimates from the calibration sub-samples as representative of the entire cohorts, we compared the centre mean dietary intakes obtained from the baseline dietary methods between the calibration and the rest of the cohort. Dietary intakes estimated from baseline assessment methods were used in this analysis because they were the only dietary measurements available from *all* the EPIC study subjects (24-hour dietary recalls were collected from only 5–12% of the EPIC cohorts). The statistical analysis was stratified by centre in order to control for differences in baseline dietary methods used across EPIC and the dietary comparison was made for 16 main food groups, using the same EPIC-SOFT classification system across centres. Overall, 89% of the centre–sex–food group combinations considered show a mean difference of less than ± 10% (69% had a difference within ± 5%). However, 59% of the differences above ± 10% were observed in only four centres (UK "health-conscious" group, Ragusa, Granada, and Umeå) out of the 24 centres involved in this analysis. The UK "health-conscious" sub-group alone represented about one-fourth of these values, probably because of the low participation rate, the small size of the calibration sample and the lack of representativeness of the different sub-components of this group (i.e. vegans, vegetarians, and fish eaters) compared to the rest of the "health-conscious" cohort. For Granada and Umeå, the relatively low response rate (~70%) and statistically significant difference in distribution between the calibration and the rest of the cohort for anthropometric measurements (women in both centres), smoking status and physical activity at work (Umeå), suggest that they may not be strictly representative of the entire cohort, but further investigations are required. For Ragusa, no explanation was found to explain the systematic differences in mean estimates observed in about one-third of the combinations.

This analysis will be presented in greater detail elsewhere and further explanatory statistical analyses will consider the impact of imperfect representativeness of the calibration sub-samples observed in certain centres, particularly the "health-conscious" group, when calibrating individual dietary questionnaire measurements.

Logistic constraints in performing interviews during weekends were reported in several EPIC countries. They were partially overcome by conducting interviews for Saturdays on Mondays, allowing a 48-hour time interval. This, however, made it impossible to distinguish whether observed variations in average food intakes between Saturdays and other weekdays reflected the true differences or whether they were the result of bias because of the increased time elapsed (48 hours instead of 24), which may have affected the subjects' memory and capacity to report their diet. Fridays, for which interviews could only be performed during the weekend, were frequently under-sampled compared to the other weekdays. Collecting dietary interviews by telephone, using an adapted version of the EPIC-SOFT programme as successfully experimented in Norway, may be a promising alternative to improve the coverage of all weekdays and seasons in future (27). However, the practical difficulty of obtaining an equal distribution of 24-HDR according to days of the week and seasons and the confirmation that a high day-to-day variation for different food groups such as meat, fish and alcohol exists (28-30) suggest that adjustments for imperfect distributions of season and particularly day of the week are needed in statistical analyses on diet and when applying the calibration.

In most situations, the 24-HDRs were collected either at the time of baseline examination, or after re-contacting the subjects. In several centres, cohorts existed before they joined the EPIC network. In other centres, where baseline recruitment had started relatively early, the calibration studies were initiated several years later because the EPIC-SOFT programme had not yet been finalized. In these centres, subjects were re-contacted up to 3 years after their baseline examination and dietary questionnaire assessment. A somewhat longer time interval between the baseline dietary questionnaire assessment and 24-HDR for the calibration studies may have the advantage of reducing correlations between random errors of the two measurements, to the extent that such correlations depend on whether measurements were collected over a short interval of time. A disadvantage, however, is that over longer time intervals more subjects may have changed their diet because of age, development of disease, or other changes in life status. The calibration studies were set up to correct for between-centre differences in the effects of errors in the baseline dietary questionnaire assessments. This correction will have to rely on the assumption that changes in true mean intake level over time did not substantially affect the validity of the 24-HDR measurements for inferences about between-centre differences in subjects' true habitual intake level at the time of recruitment.

The time needed to perform the interviews with the EPIC-SOFT programme (~30 minutes) was quite comparable across centres and compatible with the cost and logistic constraints of large nutritional studies. This includes both the time needed to perform the dietary interview and the automatic data entry. More time, varying across centres, was needed, however, to update incomplete 24-hour dietary recalls after the interview. The degree of standardization of 24-HDR measurements for use in calibration sub-studies has been reported elsewhere (19). Overall, 24-HDR measurements were reasonably well-standardized across the interviewers involved in the calibration studies, although within certain centres an interviewer or gender effect was observed. The extent of systematic under-reporting associated with 24-HDR measurements and its main determinants, discussed elsewhere in this supplement, will give further insights into the relative validity of mean 24-HDR measurements (31).

This was the first time that calibration sub-studies had been set up in a large multi-centre European study. These studies showed that, despite some inherent methodological and logistic constraints, such a study design works relatively well in practice and can provide valuable additional measurements for better interpreting results from multi-centre epidemiological studies on diet and risk of chronic disease. In addition, the overall results suggest that, after adjustment for age, the calibration samples are fairly representative of the entire cohorts and that dietary intakes estimated from these sub-samples should reasonably be interpreted as representative of the main cohorts in most of the EPIC centres.

Acknowledgements

The work described in the paper was carried out with the financial support of the "Europe Against Cancer" Programme of the European Commission (SANCO); Ligue contre le Cancer (France); Société 3M (France); Mutuelle Générale de l'Education Nationale; Institut National de la Santé et de la Recherche Médicale (INSERM); Institut Gustave Roussy; German Cancer Aid; German Cancer Research Center; German Federal Ministry of Education and Research; Danish Cancer Society; Health Research Fund (FIS) of the Spanish Ministry of Health; the Spanish Regional Governments of Andalucia, Asturia, Basque Country, Murcia and Navarra; Cancer Research UK; Medical Research Council, UK; the Stroke Association, UK; British Heart Foundation; Department of Health, UK; Food Standards Agency, UK; the Wellcome Trust, UK; Greek Ministry of Health; Greek Ministry of Education; Italian Association for Research on Cancer; Italian National Research Council; Dutch Ministry of Public Health, Welfare and Sports; Dutch Prevention Funds; LK Research Funds; Dutch ZON (Zorg Onderzoek Nederland); World Cancer Research Fund (WCRF); Swedish Cancer Society; Swedish Scientific Council; Regional Government of Skane, Sweden; Norwegian Cancer Society; the Norwegian Research Council. Partial support for the publication of this supplement was provided by the Centre de Recherche et d'Information Nutritionnelles (CERIN). In addition, we wish to thank all study participants for their cooperation and all the interviewers who participated to the field work studies in each EPIC centre.

References

1. Willett W. Future Research Directions. In: Willett W, ed. *Nutritional Epidemiology*, New York: Oxford University Press, 1990, 380–5.
2. White E, Hunt JR, Casso D. Exposure measurement in cohort studies: the challenges of prospective data collection. *Epidemiol. Rev.* 1998; **20**: 43–56.
3. Willett WC, Coldditz GA. Approaches for conducting large cohort studies. *Epidemiology* 1998; **20**: 91–9.
4. Prentice RL. Design issues in cohort studies. *Stat. Methods Med. Res.* 1995; **4**: 273–92.
5. Friedenreich CM. Methodologic issues for pooling dietary data. *Am. J. Clin. Nutr.* 1994; **59**: S251–2.
6. Rosner B, Willett WC, Spiegelman D. Correction of logistic relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat. Med.* 1989; **8**: 1051–69.
7. Plummer M, Clayton D, Kaaks R. Calibration in multi-centre cohort studies. *Int. J. Epidemiol.* 1994; **23**: 419–26.
8. Kaaks R, Plummer M, Riboli E, Estève J, van Staveren WA. Adjustment for bias due to errors in exposure assessments in multi-centre cohort studies on diet and cancer: a calibration approach. *Am. J. Clin. Nutr.* 1994; **59**: S245–50.
9. Kaaks R, Riboli E, van Staveren WA. Calibration of dietary intake measurements in prospective cohort studies. *Am. J. Epidemiol.* 1995; **142**: 548–56.
10. Kipnis V, Carroll RJ, Freedman LS, Li L. Implications of a new dietary measurement error model for estimation of relative risk: application to four calibration studies. *Am. J. Epidemiol.* 1999; **150**: 642–51.

11. Thompson FE, Moler JE, Freedman LS, Clifford CK, Stables GJ, Willett WC. Register of dietary assessment calibration-validation studies: a status report. *Am. J. Clin. Nutr.* 1997; **65**: S1142–7.
12. Stram DO, Hankin JH, Lynne R, Wilkens LR, Pike MC, Monroe KR, Park S, Henderson BE, Nomura AMY, Earle ME, Nagamine FS, Kolonel LN. Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. *Am. J. Epidemiol.* 2000; **151**: 358–70.
13. Riboli E, Elmståhl S, Saracci R, Gullberg B, Lindgärde F. The Malmö Food Study: validity of two dietary assessment methods for measuring nutrient intakes. *Int. J. Epidemiol.* 1997; **26**: S161–71.
14. Overvad K, Tjønneland A, Haraldsdóttir J, Bang S, Ewertz, Møller-Jensen O. Development of a semi-quantitative food frequency questionnaire to assess food, energy and nutrient intake in Denmark. *Int. J. Epidemiol.* 1991; **20**: 906–12.
15. Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, Sneyd MI, Key TJA, Roe L, Day NE. Comparison of dietary assessment methods in nutritional epidemiology: weighted records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. *Br. J. Nutr.* 1994; **72**: 619–43.
16. Margetts BM, Pietinen P, Riboli E. EPIC: European Prospective Investigation into Cancer and Nutrition: validity studies on dietary assessment methods. *Int. J. Epidemiol.* 1997; **26** (Suppl. 1): S1–189.
17. Riboli E, Hunt K, Slimani N, Ferrari P, Norat T, Fahey M, Charrondière R, Hémon B, Casagrande B, Vignat J, Overvad K, Tjønneland A, Clavel-Chapelon F, Wahrendorf J, Boeing H, Trichopoulos D, Trichopoulou A, Vineis P, Palli D, Bueno-de-Mesquita HB, Peeters PHM, Lund E, Engeset D, González CA, Barricarte A, Berglund G, Hallmans G, Day N, Key T, Kaaks R, Saracci R. EPIC: study population and data collection. *Public Health Nutr.* 2002; (this supplement)
18. Slimani N, Deharveng G, Charrondière RU, van Kappel AL, Ocké MC, Welch A, Lagiou A, van Liere M, Agudo A, Pala V, Brandstetter B, Andrén C, Stripp C, van Staveren WA, Riboli E. Structure of the standardized computerized 24-hour diet recall interview used as reference method in the 22 centres participating in the EPIC project. *Comput. Methods Programs Biomed.* 1999; **53**: 251–66.
19. Slimani N, Ferrari P, Ocké M, Welch A, Boeing H, van Liere M, Pala V, Amiano P, Lagiou A, Mattisson I, Stripp C, Engeset D, Charrondière R, Buzzard M, van Staveren W, Riboli E. Standardization of the 24-hour diet recall calibration method used in the European Prospective Investigation into Cancer and Nutrition (EPIC): general concepts and preliminary results. *Eur. J. Clin. Nutr.* 2000; **54**: 900–17.
20. Kaaks R, Riboli E, van Staveren W. Sample size requirements for calibration studies of dietary intake measurements in prospective cohort investigations. *Am. J. Epidemiol.* 1995; **142**: 557–65.
21. SAS Institute. *SAS/STATR User's Guide Version 8*. SAS Manual 4(6). Cary, NC: SAS Institute, 1999–2000.

22. Haftenberger M, Lahmann PH, Panico S, Gonzalez CA, Seidell JC, Boeing H, Giurdanella MC, Krogh V, Bueno-de-Mesquita HB, Peeters PHM, Skeie G, Hjartåker A, Rodriguez M, Quirós JR, Berglund G, Janlert U, Khaw KT, Spencer EA, Overvad K, Tjønneland A, Clavel-Chapelon F, Tehard B, Miller AB, Klipstein-Grobusch K, Benetou V, Kiriazi G, Riboli E, Slimani N. Overweight, obesity and body fat distribution in individuals aged 50–64 years involved in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Public Health Nutr.* 2002; (this supplement)
23. Haftenberger M, Schuit AJ, Tormo MJ, Boeing H, Wareham N, Bueno-de-Mesquita HB, Kumle M, Hjartåker A, Chirlaque MD, Ardanaz E, Andren C, Lindahl B, Peeters PHM, Allen NE, Overvad K, Tjønneland A, Clavel-Chapelon F, Linseisen J, Bergmann MM, Trichopoulou A, Lagiou P, Salvini S, Panico S, Riboli E, Ferrari P, Slimani N. Physical activity of subjects aged 50–64 years involved in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Public Health Nutr.* 2002; (this supplement)
24. Klipstein-Grobusch K, Slimani N, Krogh V, Keil U, Boeing H, Overvad K, Tjønneland A, Clavel-Chapelon F, Thiebaut A, Lineisen J, Schulze MB, Lagiou P, Papadimitrou A, Saieva, Veglia F, Bueno-de-Mesquita HB, Peeters PHM; Kumle M, Brustad M, Martinez-García C, Barricarte A, Berglund G, Weinehall L, Mulligan A, Allen N, Ferrari P, Riboli E. Trends in self-reported past alcoholic beverage consumption and ethanol intake from 1950 to 1995 observed in eight European countries participating to the European Investigation into Cancer and Nutrition (EPIC) project. *Public Health Nutr.* 2002; (this supplement)
25. Pols MA, Peeters PH, Ocké MC, Slimani N, Bueno-de-Mesquita HB, Collette HJA. Estimation of reproducibility and relative validity of the questions included in the EPIC physical activity questionnaires. *Int. J. Epidemiol.* 1997; **26**: S181–9.
26. SENECA Investigators. Assessment of bias in the SENECA study. *Eur. J. Clin. Nutr.* 1996; **50**(Suppl. 2): S4–S8
27. Brustad M, Skeie G, Braaten T, Slimani N, Lund E. Comparison of telephone versus face-to-face interview in the assessment of dietary intake by the 24-hour recall EPIC-SOFT programme – the Norwegian calibration study. *Eur. J. Clin. Nutr.* (in press).
28. Linseisen J, Kesse E, Slimani N, Bueno-de-Mesquita BH, Ocké MC, Skeie G, Kumle M, Dorronsoro Iraeta M, Morote Gómez P, Janzon L, Stattin P, Welch A, Spencer EA, Overvad K, Tjønneland A, Clavel-Chapelon F, Miller AB, Klipstein-Grobusch K, Lagiou P, Kalapothaki V, Masala G, Giurdanella MC, Norat T, Riboli E. Meat consumption in the EPIC cohorts – results from the 24-hour dietary recalls. *Public Health Nutr.* 2002; (this supplement)
29. Welch AA, Lund E, Amiano P, Dorronsoro M, Brustad M, Kumle M, Rodriguez M, Lasheras C, Janzon L, Jansson J, Luben R, Spencer EA, Overvad K, Tjønneland A, Clavel-Chapelon F, Linseisen J, Klipstein-Grobusch K, Benetou V, Zavitsanos X, Tumino R, Galasso R, Bueno-de-Mesquita HB, Ocké MC, Charrondière UR, Slimani N. Variability of fish consumption within the 10 European countries participating in the EPIC Study. *Public Health Nutr.* 2002; (this supplement)

30. Sieri S, Agudo A, Kesse E, Klipstein-Grobusch K, San-José B, Welch AA, Krogh V, Luben R, Allen N, Overvad K, Tjønneland A, Clavel F, Thiébaut A, Miller A, Boeing H, Kolyva M, Saieva C, Celentano E, Ocké MC, Peeters PHM, Brustad M, Kumle M, Dorronsoro M, Fernandez Feito A, Mattisson I, Weinehall L, Slimani N, Riboli E. Patterns of alcohol consumption in 10 European countries participating in the EPIC project. *Public Health Nutr.* 2002; (this supplement)
31. Ferrari P, Slimani N, Ciampi A, Trichopoulos D, Naska A, Lauria C, Veglia F, Bueno-de-Mesquita HB, Ocké MC, Brustad M, Braaten T, Tormo MJ, Amiano P, Mattisson I, Johansson I, Welch A, Davey G, Overvad K, Tjønneland A, Clavel-Chapelon F, Thiebaut A, Linseisen J, Boeing H, Hemon B, Riboli E. Evaluation of mis-reporting in the 24-hour diet recalls in EPIC. *Public Health Nutr.* 2002; (this supplement)

Chapter 6

Group level validation of protein intakes estimated by 24-hour diet recall and dietary questionnaires against 24-hour urinary nitrogen in the EPIC calibration study

Nadia Slimani, Sheila Bingham, Shirley Runswick, Pietro Ferrari, Nicholas E Day, Ailsa Welch, Tim Key, Antony Miller, Heiner Boeing, Sabina Sieri, Fabrizio Veglia, Dominico Palli, Salvatore Panico, Rosario Tumino, Bas Bueno-de-Mesquita, Marga C Ocké, Françoise Clavel-Chapelon, Antonia Trichopoulou, Wija van Staveren, Elio Riboli

Cancer Epidemiol. Biomark. Prev. (submitted)

Abstract

A calibration approach was developed to correct for systematic between-cohort dietary measurement errors in the European Prospective Investigation into Cancer and Nutrition (EPIC), a large multi-centre cohort study. Standardized 24-hour diet recalls (24-HDRs) from a sub-sample of the study population were used as reference calibration measurements. In order to validate these measurements used for between-cohort calibration, group mean nitrogen (and total energy) intakes obtained with 24-HDRs were compared against mean 24-hour urinary nitrogen. Similar analyses using nitrogen and energy intake data from different dietary questionnaires (DQs) used at baseline were conducted in order to estimate the effect of the calibration approach. The study conducted between 1995 and 1999 involves 1103 volunteers of both genders from 12 centres participating in EPIC. The partial Pearson's correlation adjusted for gender at the group mean level between urinary and dietary nitrogen or energy intake are high with both dietary methods, when both genders are considered together. A higher correlation was observed between mean urinary and dietary nitrogen with 24-HDRs (0.84), particularly after exclusion of outliers (0.90), than with DQs (0.72) (N=22). Inversely, the correlation is higher with DQs (0.82) than with 24-HDRs (0.69; 0.77 after exclusion) when energy intakes were considered. Furthermore, the calculated β regression coefficients were not significantly different from the ideal value of 1 when mean nitrogen from 24-HDR or DQ were regressed on urinary estimates, suggesting that overall the systematic bias across centres is of uniform magnitude. Group level correlations were systematically higher in men than women with both dietary methods and for nitrogen and total energy intake. Overall, both dietary methods ranked centres well when compared to urinary nitrogen measurements. However, in contrast to 24-HDRs, errors in DQs tend to vary in both directions (under- and over-reporting). This observation may have implications on the dietary measurement error characteristics and support the potential benefit of the between-cohort calibration.

Introduction

A new generation of sophisticated validation and calibration designs using several reference dietary and/or biological measurements has been proposed to address the problem of measurement errors in dietary assessment methods used to investigate the relationship between dietary exposure and different diseases (1-3). Biological markers of absolute quantitative dietary intakes such as urinary nitrogen (4) or doubly-labelled water for estimating total energy expenditure in free living people (5) are suggested as the two most accurate independent reference measurements for use in validation (or calibration) studies. In the absence of changes in nitrogen or energy balance, these markers provide a reliable estimate of individual or population nitrogen (protein) intakes and total energy expenditure (energy intakes). Of particular importance is the fact that they fulfil the fundamental statistical requirements of a reference method in validation or calibration studies, i.e. their measurement errors are independent of both the true unknown actual dietary intakes and of errors in the dietary measurements to be validated (6). Single measures have been used to verify estimated population intakes of nitrogen by various dietary methods (7-11) whereas repeat measures are needed to verify individual estimates (12-14).

Although studies using biological markers as a reference to validate individual measurements on restricted numbers of subjects are increasingly reported in the literature, nothing has been published on the validity of dietary measurements used as reference calibration in large nutritional surveys. In such studies, characterised by study populations from different geographical areas or ethnic groups, the calibration of methods used to assess diet in epidemiological studies has been proposed for two main purposes: 1) to correct for systematic over- or under-estimation of the true intakes attributable to the different dietary questionnaires used at the population level 2) to possibly correct for attenuation bias in the relative risk due to random errors in dietary measurements at the individual level, (15-17). In practice, calibration involves applying, in addition to the dietary questionnaires used to estimate individual intakes, a second highly standardized dietary method in a representative sub-sample as a common reference measurement across study populations. For calibration at the population level (between-cohort calibration), one of the key statistical requirements is that the reference dietary method (e.g. 24-hour diet recalls, dietary records or biological markers) must be unbiased and provide a correct estimate of the mean of group populations.

This paper presents the results of a study designed to validate the mean of single 24-hour diet recall measurements, used as reference method to calibrate population mean estimates, against mean single 24-hour urinary nitrogen in the European Prospective Investigation into Cancer and Nutrition (EPIC). EPIC is a network of prospective cohort studies with approximately half a million subjects recruited from 23 European centres in 10 Western European countries (France, Italy, Spain, United Kingdom, Germany, the Netherlands, Greece, Sweden, Denmark and Norway) (18). Diet has been measured for all participants using validated dietary questionnaires (14, 19-21), and mean nitrogen (and total energy) intakes obtained from these EPIC questionnaires were also compared both to the 24-hour dietary recall and urinary measurements.

Material and methods

Subjects

The EPIC cohort includes almost 500,000 subjects from 10 countries who completed baseline dietary and other lifestyle questionnaires; 36,900 of them also underwent a single 24-hour diet recall interview which was used as the reference calibration method (18). A convenient sub-sample from the calibration in 9 EPIC centres (or country): Paris (France), Varese, Turin, Florence and Ragusa (Italy), Greece where the sample was recruited from different geographical regions, Cambridge and Oxford (United Kingdom), Heidelberg and part of the participants in Postdam (Germany) were re-contacted and asked to collect a 24-hour urine specimen between 1995 and 1999. Those who agreed to participate signed a consent form. In Bilthoven (the Netherlands) a strict random sample from the calibration sample was drawn from the subjects approached during 1997. In Potsdam (Germany) and Naples (Italy) the subjects involved in this study were all from the EPIC cohorts but not from the calibration sample. In Potsdam (Germany) most of the participants came from a previous validation study that was conducted among cohort members where repeated 24-hour diet recalls and 24-hour urines were collected using the same standard procedure as in the calibration study (22). In Naples, the subjects were selected from the whole cohort and asked to collect a 24-hour urine sample, and a 24-hour diet recall when they brought the urine collection back to the centre. A total of 12 EPIC centres were thus finally involved in the statistical analysis. Initially the sample included 1386 volunteers, both men and women, but a number of them were excluded for reasons which will be explained later. Ultimately the data from 1103 subjects were used for the statistical analysis.

Dietary data

Information on usual *individual* dietary intakes was assessed at baseline for each subject entering the EPIC cohorts using exhaustive quantitative dietary questionnaires (France, the Netherlands, Germany, Greece and Italy, except Naples) or a semi-quantitative food frequency questionnaire (UK, Naples) developed and validated in each participating country (14,20). These questionnaires contain 158 to 266 food and recipe items, and the portion sizes were estimated either by photographs or standard units, except in the UK and Naples where standard portions were used.

In addition, a single 24-hour diet recall (24-HDR) interview was used as the EPIC reference calibration measurement and collected from a representative stratified sample of each cohort (23). In contrast to the baseline dietary questionnaires, the 24-hour diet recall interviews were highly standardized across countries, using a computerized programme (EPIC-SOFT) with the same structure and interview procedure across countries. The aim of the calibration was to estimate *mean* population rather than individual dietary intakes, and only one 24-HDR measurement was collected from study subjects. Information on all foods and beverages consumed during the previous day was collected, entered and coded automatically according to common rules. Food portion sizes were estimated using a common picture book containing sets of photos of 140 foods and recipes (24) and other available methods such as standard units and household measurements. Trained dieticians conducted all the interviews face-to-face. More details on the concept of standardization and the structure of EPIC-SOFT are described in detail elsewhere (25,26).

Food composition tables

In the absence of a standardized European nutrient database (27,28), the average nitrogen and energy intakes were calculated using country-specific food composition tables for both dietary methods. A recent review on the comparability of food composition tables available in countries participating in EPIC (29) suggested that both protein and energy values reported in the food composition tables used in this analysis are reasonably comparable for between-country comparisons.

24-hour urine sample

One or two days prior to the urine collection, the subjects were instructed on how to collect a complete 24-hour urine sample following a standardized protocol (13). Each subject was provided with two 2-litre bottles for urine collection, each containing 5g of boric acid as a preservative.

In order to evaluate the completeness of the 24-hour urine collections, the subjects were given three 80 mg tablets of p-amino benzoic acid (PABA) to take at mealtimes during the course of the urine collection (4). PABA is absorbed by simple diffusion throughout the small intestines, and its metabolites are actively secreted by the renal tubule. It is therefore completely and quantitatively excreted in urine over the 24-hour period, and the completeness of collection is estimated by measuring the PABA metabolite recovery in urine. After collection, the 24-hour urine was stored at –20° C at the local centre, then packed in dry ice and shipped within 24hours to a central laboratory in Cambridge where all the chemical analyses were performed according to standard procedures (13). In urine, ototal nitrogen (N) was determined by the Kjeldahl technique (Tecator 1002; Perstorp Analytical, Bristol, Avon), and PABA was measured by colorimetry (30). 24-hour urine collections containing between 85% to 110% of the PABA marker were considered complete. Urine specimens containing less than 70% PABA recovery (when fewer than 3 tablets may have been taken) and above 110% (when subjects may have taken drugs that interfere with the colorimetric analysis) were excluded from the analysis. Further urine samples (70-84% PABA recovery) were used after correction for urinary electrolytes up to the expected values at 93% of PABA recovery (31). To allow for extra-renal losses in estimating 24-hour N excretion, each individual 24-hour urine N was divided by 0.81 (4).

Energy intake

In order to estimate the impact on the mean of extreme reported energy intakes which were physiologically implausible, high under- or over-reporters were defined according to Goldberg et al. (32). Different exclusion cut-off points were used for 24-HDRs (one single observation per individual) and for the DQ (long-term usual diet) to take into account the different random day-to-day variations in dietary intakes and physical activity assumed for these two dietary methods which are on different time scales (32).

Statistical analysis

We first compared the mean nitrogen (and energy) intakes estimated from the dietary questionnaires (DQ), 24-hour diet recalls (24-HDRs) and urine. In order to achieve asymptotic normality of the distributions, the individual nitrogen (and energy) measurements obtained from the dietary questionnaires (DQ and 24-HDRs) and the urine were log-transformed, and the geometric means and their confidence intervals were calculated.

The statistical analysis was stratified by gender and centre to take into account the different study designs and populations involved and any centre-related residual confounding. Although the total number of subjects was low in Ragusa and Naples, it was decided not to merge these two Southern centres, because different baseline (DQ) dietary measurements were used. In order to compare urinary nitrogen with the nitrogen (and energy) estimates from the two dietary methods, independently of each other, non-calibrated dietary questionnaire means were used in this analysis. The degrees of under- or over-reporting of nitrogen in dietary measurements were expressed as the ratio of the mean of 24-HDRs or DQ intakes divided by mean urinary nitrogen. For total energy, this ratio was computed as the mean total energy intakes estimated by each dietary method divided by the mean of individually calculated basal metabolic rates (BMR) multiplied by a constant Physical Activity Level (PAL) of 1.55, used as a conservative lower value for sedentary populations (33).

A t-test was used to determine statistically the differences between, respectively, the mean nitrogen DQs and 24-HDRs and the reference mean urinary nitrogen estimates. The t-test was performed on the means of log-transformed individual values, and p-values < 0.05 were considered statistically significant.

We then performed an ecological analysis between mean centre nitrogen and energy intakes estimated from DQs and 24-HDRs, and urinary nitrogen. Pearson correlation coefficients and associated 95% confidence intervals were calculated. Knowing that one of the purposes of the EPIC between-cohort calibration was to correct for centre differences in dietary measurement errors, our objective was to estimate the ability of the two dietary methods to discriminate between mean intakes across centres, taking urinary nitrogen as the reference. In addition, we performed linear regression analysis to study the relationship between the centre means for urine nitrogen and dietary nitrogen (energy intakes), when dietary measurements (24-hour diet recalls or dietary questionnaires) were used as the dependent variable. A test was performed to indicate whether the slope of the regression line was significantly different from 1. All the analyses were performed using the SAS statistical package, version 8 (SAS Institute Inc., Cary, NC, USA).

Results

Between 80-96% of the urine samples in Paris, Italy (except Naples), the Netherlands, Oxford and Potsdam and slightly less in the Heidelberg (78%) and Cambridge (72%) were considered complete with >85% recovery of PABA (**Table 1**). In Greece, only ~20% of the urine specimens were considered complete. After correction of incomplete but usable samples (70-84% PABA recovery), all centres had a PABA recovery between 82-98% except Greece (49%).

Forty-five subjects without one of the two dietary measurements (i.e. 24-HDR or DQ) were excluded from the statistical analyses, as were individuals with total energy/BMR ratio from the DQ in the top and bottom 1% of the cohort distribution by centre and gender, a routine exclusion made on the EPIC baseline questionnaire data. However, such exclusions were not performed on 24-HDRs. Furthermore, in order to compare only subjects with a balanced protein-energy metabolism, the subjects on a diet during the 24-HDR interview (N =52) were also excluded from the statistical analysis. Overall, a total of 283 subjects (20.4%) were excluded, and a final sample of 1103 was used for the statistical analysis (Table 1).

Table 2 a & b summarises some general characteristics of the study populations and the time interval between dietary and urinary measurements. This study population involved middle-aged people (~ 53 to 62 years old), except in the Netherlands, and the women in Germany and men in Ragusa where they were younger. Overall, there were more women (~ 59%) than men, reflecting the different gender distributions in EPIC (~ 70% women). The range in average BMI was narrow across countries in men (25.6-26.8 kg/m^2) but broader in women (23.2-27.5 kg/m^2). Greece had a much greater BMI in both men (29.0 kg/m^2) and women (30.0 kg/m^2).

Table 1. PABA recovery and type and number of subjects excluded from the statistical analysis

	France	Italy					UK		NL	Germany		Greece	Total
	Paris	Varese	Turin	Florence	Ragusa	Naples	Cambridge	Oxford	Bilthoven	Heidelberg	Potsdam		
- Initial samples numbers	137	61	48	54	21	27	340	98	223	95	155	127	**1386**
- Gender (% of women)	100	44.3	50.0	74.1	47.6	100	53.2	54.1	59.6	53.7	42.6	53.5	**59.0**
- Satisfactory PABA Recovery (%)[a]	81.8	83.6	95.8	87.0	81.0	66.7	71.8	85.7	81.5	77.9	85.2	20.5	**74.7**
- Satisfactory PABA recovery, after correction (%)[b]	91.2	88.5	97.9	98.1	85.7	81.5	86.8	95.9	85.2	89.5	94.8	48.8	**86.0**
Type and number of exclusions performed before the statistical analysis													
- Incomplete PABA recovery [c]	12	7	1	1	3	5	44	3	33	10	6	64	**189**
- 24-HDR not available	0	0	0	0	0	0	0	0	2	0	14	1	**17**
- *Dietary Questionnaire not available*[d]	2	3	0	0	0	1	13	2	2	2	2	1	**28**
- On a diet	17	1	4	4	0	1	1	9	9	2	3	1	**52**
- Total exclusion[e]	30[e] *21.9%*	*10[e]* *16.4%*	*5* *10.4%*	*5* *9.3%*	*3* *14.3%*	*7* *25.9%*	*58* *17.1%*	*14* *14.3%*	*46* *20.6 %*	*14* *14.7%*	*25* *16.1%*	*66[e]* *52.0%*	**283[e]** ***20.4%***
- Final samples (per center)	107	51	43	49	18	20	282	84	177	81	130	61	**1103**

[a]: : Corresponds to PABA recovery between 85% to 110%

[b]: : Correction of the usable urinary samples with PABA recovery between 70-84%, up to the 93% PABA recovery expected nitrogen values, when only one urine sample is used

[c]: : Corresponds to PABA recovery below 70% and above 110%

[d] : Or dietary questionnaires with an energy intake/BMR ratio within 1% of both sides of EPIC cohort distribution per country and gender

[e] : One subject in Paris and one in Varese were excluded from the total exclusion, not counted twice, because of both incomplete PABA recovery and dieting.

Table 2: Study population characteristics after exclusions, and time intervals between urinary and 24-hour diet recall (24HDR) and dietary questionnaire (DQ) measurements (mean and standard deviations are shown)

	FRANCE	ITALY					UK	
	Paris	Varese	Turin	Florence	Ragusa	Naples	Cambridge	Oxford
MEN								
Number	-	32	22	13	11	-	128	38
Age (years)	-	60.0 ± 3.9	55.3 ± 8.1	57.2 ± 7.3	48.7 ± 4.6	-	56.6 ± 9.4	59.7 ± 8.7
Height (cm)	-	173.6 ± 5.0	176.0 ± 8.3	169.5 ± 6.6	171.8 ± 4.4	-	174.7 ± 7.2	178.1 ± 7.1
Weight (Kg)	-	78.9 ± 8.8	79.5 ± 10.1	73.7 ± 9.2	77.0 ± 9.4	-	78.6 ± 10.6	78.8 ± 11.2
Body mass index (kg/m^2)	-	26.2 ± 2.3	25.6 ± 2.8	25.7 ± 3.2	26.1 ± 3.3	-	25.8 ± 3.2	24.8 ± 3.0
Time interval 1 (days) [a]	-	0	0	0	0	-	-3.6 ± 24.6	-4.2 ± 1.8
Time interval 2 (days) [b]	-	-334.7 ± 132.7	-853.9 ± 556.1	-665.4 ± 243.1	-622.7 ± 350.9	-	-541.8 ± 404.1	-521.9 ± 301.9
WOMEN								
Number	107	19	21	36	7	20	154	46
Age (years)	57.0 ± 7.0	54.8 ± 9.5	55.7 ± 6.7	58.1 ± 6.7	57.9 ± 6.6	52.9 ± 6.3	55.0 ± 8.8	55.4 ± 9.7
Height (cm)	161.6 ± 6.2	159.5 ± 6.0	163.1 ± 6.9	160.3 ± 5.7	155.7 ± 5.7	159.8 ± 4.9	162.2 ± 6.5	162.5 ± 6.0
Weight (Kg)	60.5 ± 9.0	61.8 ± 7.7	62.8 ± 10.3	66.0 ± 14.8	66.7 ± 10.3	66.1 ± 9.8	67.9 ± 9.5	68.0 ± 11.7
Body mass index (kg/m^2)	23.2 ± 3.1	24.4 ± 3.4	23.6 ± 3.4	25.5 ± 4.5	27.5 ± 4.1	25.9 ± 3.7	25.0 ± 3.5	25.9 ± 5.1
Time interval 1 (days) [a]	-0.7 ± 1.0	0	0.3 ± 1.5	0	0	0	-5.9 ± 3.2	-4.8 ± 1.7
Time interval 2 (days) [b]	-1311.4 ± 219.9	-797.1 ± 474.7	-602.4 ± 548.7	-573.8 ± 380.4	-560.6 ± 536.1	-1807.2 ±66.0	-487.1 ± 376.1	-533.0 ± 365.5

Table 2: (Continued)

	NL	GERMANY		GREECE
	Bilthoven	**Heidelberg**	**Potsdam**	
MEN				
Number	72	38	77	30
Age (years)	50.3 ± 11.9	52.6 ± 7.1	55.7 ± 8.2	61.8 ± 11.5
Height (cm)	178.7 ± 7.2	175.9 ± 6.5	174.5 ± 6.7	167. ± 3 7.6
Weight (Kg)	85.5 ± 11.2	81.6 ± 14.8	81.0 ± 9.6	81.3 ± 9.3
Body mass index (kg/m^2)	26.8 ± 4.0	26.3 ± 3.9	26.7 ± 3.3	29.0± 2.6
Time interval 1 (days) [a]	-3.5 ± 2.0	-3.0 ± 5.8	0.3 ± 2.5	0.5 ± 1.5
Time interval 2 (days) [b]	-12.0 ± 17.2	-40.6 ± 55.3	-46.0 ± 35.3	-367.3 ± 282.4
WOMEN				
Number	105	43	53	31
Age (years)	46.8 ± 10.0	50.7 ± 8.9	52.7 ± 9.0	58.1 ± 10.4
Height (cm)	165.9 ± 7.6	163.2 ± 5.9	161.1 ± 5.5	156. ± 6 6.1
Weight (Kg)	69.3 ± 11.0	65.6 ± 12.2	67.5 ± 12.1	73.5 ± 11.7
Body mass index (kg/m^2)	25.2 ± 3.8	24.7 ± 5.1	26.1 ± 4.8	30.0± 4.7
Time interval 1 (days) [a]	-3.6 ± 1.9	-1.3 ± 4.8	0.22 ± 1.6	0.3 ± 1.2
Time interval 2 (days) [b]	-8.3 ± 17.5	-41.7 ± 47.0	-49.4 ± 59.4	-301.8 ± 249.0

[a] Time interval 1 : time interval between the 24-hour diet recall measurements and the 24-hour urine collection
[b] Time interval 2 : time interval between the baseline individual dietary questionnaire and the 24-hour urine collection. For Italy, the time interval was estimated using the date of collection of non dietary EPIC data (the date of collection of dietary data is missing)

The first purpose of our analysis was to validate the mean 24-hour diet recall versus 24-hour urinary measurements. The urine samples were therefore collected at the same time as the 24-hour diet recall interview or within a maximum of 6 days afterwards. The time interval between the two dietary questionnaire collections, and hence between the DQ and urine measurements, is longer and varies across countries. In Bilthoven and the German centres, the 24-HDRs were collected from a few days to less than two months after the baseline dietary questionnaire. In the other centres, 24-HDRs and urine samples were obtained 11 to 21 months later and more than 2 to 5 years after in Turin (men), France and Naples.

Table 3 a & b reports the geometric means and the ratio between nitrogen (and total energy) estimated from either the 24-HDR or the DQs and 24-hour urine stratified by centre and gender. The reported ratios, indicating the degree of systematic under- or over-reporting of dietary nitrogen and total energy intakes compared to the reference measurements are presented with and without exclusion of extreme under- or over-reporters according to Goldberg et al. (32). The results presented after exclusion were obtained using the Goldberg cut-off based on either 24-HDR <u>or</u> DQ specific values (not both).

The mean N 24-HDR/N urinary ratio ranged from 0.76 (Greece) to 0.99 (Ragusa) for men and from 0.54(Greece) to 0.92 (Paris) for women. The t-test consistently showed a statistically significant difference between log-transformed means in all centres except in men in Varese, Turin and Ragusa (Italy) and women in Ragusa, with average under-reporting of 13% in men and 19% in women when Greece was omitted. The exclusion according to Goldberg's 24-HDR cut-off did not change the significance of the results except for men in the Netherlands and Heidelberg and women in France, where the difference between means is no longer statistically different from zero after the exclusion of 13 to 18% of subjects. For dietary questionnaires, the mean N DQ/N urinary ratio ranged from 0.75 (Heidelberg) to 0.90 (Oxford) for men and 0.76 (Heidelberg) to equal or above 1.04 (Oxford, Naples, France, Florence, Cambridge) for women. The difference between means is not statistically significant for men in Florence and Ragusa and for women in France, Oxford and all Italian centres. The exclusion according to Goldberg (percent exclusion indicated in brackets) changes the significance of the results in men in Turin (23%), Greece (33%) and Cambridge (47%), and in women in Heidelberg (47%), Potsdam (32%) and Greece (32%).

Table3a. Comparison of geometric means (SE intervals) and ratios of mean nitrogen (and total energy) estimated from 24-H urine, 24-Hour diet recall (24HDR) and Dietary questionnaire (DQ): Men

	FRANCE	ITALY					UK	
	Paris	Varese	Turin	Florence	Ragusa	Naples	Cambridge	Oxford
- N° of subjects	-	32	22	13	11	-	128	38
- Mean nitrogen in urine (gr.)	-	17.84 [16.53-19.26]	17.48 [15.96-19.14]	16.85 [14.92-19.03]	19.66 [17.05-22.66]	-	15.05 [14.46-15.67]	15.71 [14.80-16.68]
- Mean nitrogen in 24HDRs gr.)	-	17.27 [15.49-19.25]	14.93 [13.03-17.11]	13.58 [11.52-16.00]	19.53 [15.01-25.41]	-	12.79 [12.10-13.52]	13.49 [12.26-14.83]
- Mean nitrogen in DQ (gr.)	-	15.52 [14.03-17.17]	14.84 [13.50-16.32]	14.48 [12.47-16.80]	17.51 [14.63-20.95]	-	13.25 [12.69-13.83	14.07 [13.20-14.99
- Ratio N 24HDR/N urinary [a]	-	0.97 (0.99)	0.85 (0.94)	0.81 (0.92*)	0.99 (1.00)	-	0.85 (0.94**)	0.86 (0.94**)
- Ratio N 24HDR/N urinary (*after exclusion*) [a]	-	0.96 (0.99)	0.88 (0.95)	0.81 (0.92*)	0.99 (1.00)	-	0.88 (0.95**)	0.86 (0.94*)
- Ratio N DQ/ N urinary [a]	-	0.87 (0.95*)	0.85 (0.94*)	0.86 (0.95)	0.89 (0.96)	-	0.88 (0.95**)	0.90 (0.96*)
- Ratio N DQ/ N urinary (*after exclusion*) [a]	-	0.88 (0.96*)	0.87 (0.95)	0.83 (0.94)	0.88 (0.96)	-	0.98 (0.99)	0.91 (0.97*)
- Ratio 24HDR energy intake/energy requirement [a]	-	1.07 (1.01)	0.96 (0.99)	0.90 (0.99)	0.96 (1.00)	-	0.86 (0.98**)	0.95 (0.99)
- Ratio 24HDR energy intake/energy requirement (*after exclusion*) [a]	-	1.03 (1.00)	0.98 (1.00)	0.90 (0.99)	0.96 (1.00)	-	0.90 (0.99**)	0.95 (0.99)
- Ratio DQ energy intake/ energy requirement	-	0.93 (0.99)	0.90 (0.99*)	0.99 (1.00)	1.09 (1.01)	-	0.78 (0.97**)	0.87 (0.98**)
- Ratio DQ energy intake/energy requirement (*after exclusion*) [a]	-	0.95 (0.99)	0.96 (0.99)	0.94 (0.99)	1.08 (1.01)	-	0.92 (0.99**)	0.90 (0.99**)
- Extreme reporters excluded from the 24HDRs (%)	-	2 (6.3)	1 (4.5)	0 (0.0)	0 (0.0)	-	13 (10.2)	2 (5.3)
- Extreme reporters excluded from the DQs (%)	-	6 (18.8)	5 (22.7)	3 (23.0)	2 (18.2)	-	60 (46.8)	8 (21.1)

Table3a: Men (Continued)

	NL	GERMANY		GREECE
	Bilthoven	Heidelberg	Potsdam	
- N° of subjects	72	38	77	30
- Mean nitrogen in urine (gr.)	17.01 [16.22-17.84]	15.94 [14.02-18.12]	15.86 [15.15-16.60]	15.19 [13.63-16.92]
- Mean nitrogen in 24HDRs gr.)	14.83 [13.21-16.65]	12.48 [10.68-14.57]	13.03 [12.05-14.08]	10.53 [8.84-12.55]
- Mean nitrogen in DQ (gr.)	13.96 [13.05-14.92]	11.96 [11.13-12.85]	13.58 [12.84-14.36]	12.10 [10.93-13.39]
- Ratio N 24HDR/N urinary [a]	0.87 (0.95*)	0.78 (0.91*)	0.82 (0.93**)	0.69 (0.86**)
- Ratio N 24HDR/N urinary (*after exclusion*) [a]	0.97 (0.99)	0.88 (0.95)	0.84 (0.94**)	0.76 (0.90*)
- Ratio N DQ/ N urinary [a]	0.82 (0.93**)	0.75 (0.90**)	0.86 (0.94**)	0.80 (0.92**)
- Ratio N DQ/ N urinary (*after exclusion*) [a]	0.90 (0.96*)	0.80 (0.92*)	0.91 (0.96*)	0.87 (0.95)
- Ratio 24HDR energy intake/energy requirement [a]	0.89 (0.99*)	0.88 (0.98*)	0.88 (0.98**)	0.70 (0.95**)
- Ratio 24HDR energy intake/energy requirement (*after exclusion*) [a]	0.98 (1.00)	0.95 (0.99)	0.91 (0.99**)	0.81 (0.97**)
- Ratio DQ energy intake/ energy requirement	0.82 (0.98**)	0.85 (0.98**)	0.90 (0.99**)	0.83 (0.98**)
- Ratio DQ energy intake/energy requirement (*after exclusion*) [a]	0.95 (0.99)	0.92 (0.99*)	0.94 (0.99*)	0.92 (0.99)
- Extreme reporters excluded from the 24HDRs (%)	*10 (13.9)*	7 (18.4)	7 (9.1)	7 (23.3)
- Extreme reporters excluded from the DQs (%)	26 (36.1)	11 (28.9)	19 (24.7)	10 (33.3)

- [a] The ratios correspond to the ratios of geometric means. The ratios in brackets correspond to the ratios of means of log-transformed measurements. A T-test of statistical significance of mean differences between urinary and dietary measurements (24_HDR and DQ) has been calculated on the means of log-transformed values. The symbols * and * * stand, respectively, for significance at 95% and 99% of the difference between the two mean estimates.

Table3b: Comparison of geometric means (SE intervals) and ratios of mean nitrogen (and total energy) estimated from 24-H urine, 24-Hour diet recall (24HDR) and Dietary questionnaire (DQ): Women

	FRANCE	ITALY					UK	
WOMEN	Paris	Varese	Turin	Florence	Ragusa	Naples	Cambridge	Oxford
- N° of subjects	107	19	21	36	7	20	154	46
- Mean nitrogen in urine (gr.)	13.71 [13.14-14.32]	13.87 [12.64-15.21]	14.95 [13.25-16.85]	13.47 [12.47-14.54]	13.25 [11.31-15.52]	14.88 [13.40-16.53]	12.48 [12.05-12.92]	11.44 [10.87-12.04]
- Mean nitrogen in 24HDRs (gr.)	12.57 [11.82-13.37]	9.02 [7.87-10.33]	12.15 [10.64-13.88]	11.12 [9.81-12.62]	11.55 [8.84-15.08]	11.54 [10.07-13.22]	10.49 [9.97-11.04]	8.39 [7.69-9.15]
- Mean nitrogen in DQ (gr.)	14.73 [13.87-15.64]	12.37 [10.79-14.20]	12.71 [11.44-14.12]	14.50 [13.29-15.82]	11.25 [8.62-14.69]	15.73 [14.28-17.32]	13.62 [13.08-14.18]	11.88 [11.08-12.74]
- Ratio N 24HDR/N urinary [a]	0.92 (0.97*)	0.65 (0.84**)	0.81 (0.92*)	0.83 (0.93*)	0.87 (0.95)	0.78 (0.91**)	0.84 (0.93**)	0.73 (0.87**)
- Ratio N 24HDR/N urinary (*after exclusion*) [a]	0.93 (0.97)	0.69(0.86**)	0.78 (0.92*)	0.86 (0.94*)	0.87 (0.95)	0.75 (0.89**)	0.90 (0.96**)	0.75 (0.88**)
- Ratio N DQ/ N urinary [a]	1.07 (1.03)	0.89 (0.96)	0.85 (0.94)	1.08 (1.03)	0.85 (0.94)	1.06 (1.02)	1.09 (1.03**)	1.04 (1.02)
- Ratio N DQ/ N urinary (*after exclusion*) [a]	1.02 (1.01)	0.90 (0.96)	0.81 (0.92*)	1.08 (1.03)	0.92 (0.97)	1.00 (1.00)	1.14 (1.05**)	1.09 (1.04)
- Ratio 24HDR energy intake/energy requirement [a]	0.93 (0.99*)	0.81 (0.97**)	0.99 (1.00)	0.89 (0.98*)	0.89 (0.98)	0.89 (0.99)	0.77 (0.97**)	0.74 (0.96**)
- Ratio 24HDR energy intake/energy requirement (*after exclusion*) [a]	0.96 (0.99)	0.88 (0.98*)	1.00 (1.00)	0.92 (0.99)	0.89 (0.98)	0.86 (0.98**)	0.86 (0.98**)	0.83 (0.98**)
- Ratio DQ energy intake/ energy requirement	1.05 (1.00)	0.98 (1.00)	0.93 (0.99)	1.03 (1.00)	0.91 (0.99)	1.11 (1.01)	0.93 (0.99**)	0.85 (0.98**)
- Ratio DQ energy intake/energy requirement (*after exclusion*) [a]	1.00 (1.00)	0.99 (1.00)	0.96 (0.99)	1.01 (1.00)	1.01 (1.00)	1.01 (1.00)	0.98 (1.00)	0.91 (0.99**)
- Extreme reporters excluded from the 24HDRs (%)	*14 (13.1)*	3 (15.8)	2 (9.5)	2 (5.5)	0 (0.0)	1 (5.0)	29 (18.8)	11 (23.9)
- Extreme reporters excluded from the DQs (%)	29 (27.1)	5 (26.3)	6 (28.6)	8 (22.2)	3 (42.9)	6 (30.0)	40 (26.0)	20 (43.5)

Table3b: Women (continued)

WOMEN	NL	GERMANY		GREECE
	Bilthoven	Heidelberg	Potsdam	
- N° of subjects	105	43	53	31
- Mean nitrogen in urine (gr.)	13.47 [12.85-14.13]	12.66 [11.47-13.97]	12.14 [11.43-12.89]	12.04 [10.93-13.28]
- Mean nitrogen in 24HDRs (gr.)	11.81 [11.02-12.66]	10.32 [9.16-11-63]	9.69 [8.64-10.87]	6.50 [5.23-8.06]
- Mean nitrogen in DQ (gr.)	11.23 [10.70-11.78]	9.67 [8.95-10.46]	10.75 [9.98-11.59]	10.48 [9.58-11.47]
- Ratio N 24HDR/N urinary [a]	0.88 (0.95**)	0.82 (0.92*)	0.80 (0.91**)	0.54 (0.75**)
- Ratio N 24HDR/N urinary (*after exclusion*) [a]	0.91 (0.96*)	0.81 (0.92**)	0.87 (0.94*)	0.74 (0.88*)
- Ratio N DQ/ N urinary [a]	0.83 (0.93**)	0.76 (0.89**)	0.89 (0.95*)	0.87 (0.94*)
- Ratio N DQ/ N urinary (*after exclusion*) [a]	0.91 (0.96**)	0.84 (0.93)	0.94 (0.98)	0.95 (0.98)
- Ratio 24HDR energy intake/energy requirement [a]	0.87 (0.98**)	0.90 (0.99)	0.84 (0.98**)	0.60 (0.93**)
- Ratio 24HDR energy intake/energy requirement (*after exclusion*) [a]	0.92 (0.99**)	0.92 (0.99*)	0.91 (0.99*)	0.80 (0.97**)
- Ratio DQ energy intake/ energy requirement	0.83 (0.98**)	0.85 (0.98**)	0.89 (0.98**)	0.83 (0.98**)
- Ratio DQ energy intake/energy requirement (*after exclusion*) [a]	0.93 (0.99**)	1.00(1.00)	0.96 (0.99)	0.94 (0.99)
- Extreme reporters excluded from the 24HDRs (%)	11 (10.5)	6 (14.0)	7 (13.2)	14 (45.2)
- Extreme reporters excluded from the DQs (%)	36 (34.3)	20 (46.5)	17 (32.1)	10 (32.2)

- [a] The ratios correspond to the ratios of geometric means. The ratios in brackets correspond to the ratios of means of log-transformed measurements.

A T-test of statistical significance of mean differences between urinary and dietary measurements (24_HDR and DQ) has been calculated on the means of log-transformed values. The symbols * and * * stand, respectively, for significance at 95% and 99% of the difference between the two mean estimates.

The energy intake/BMR ratio calculated from 24-HDRs ranges from 0.86 to 0.96, except in Varese (1.07) where it is higher but not statistically significant, and lower in Greece (0.70). In women, the ratios are from 0.74 (Oxford) to 0.93 (Paris), and 0.60 in Greece. The mean difference is not statistically significant in the Italian centres (except Varese and Florence in women), in Oxford (men) and in women in Paris, and after exclusion of men in Bilthoven and Heidelberg.

In the dietary questionnaire, the energy intake/BMR ratio ranged from 0.78 (Cambridge) to 1.09 (Ragusa) for men and 0.83 (Bilthoven, Greece) to 1.11 (Naples) for women. The test comparison showed no statistically significant difference between means in France and in either gender in Italy, except Florence in men. After Goldberg exclusions, the means were no longer statistically significant for men in Florence (23% exclusions), Bilthoven (36%) and Greece (33%), and women in Cambridge (26%), Greece and Potsdam (32%) and Heidelberg (47%).

Figures 1 & 2 show the Pearson's correlation at the group mean level, between urinary nitrogen and nitrogen or total energy intake from 24-HDRs and DQs. The results presented include both genders using a partial Pearson's correlation adjusted for sex (N=22). The correlation between mean centre urinary and dietary nitrogen is higher with 24-HDRs (0.84 [0.65-0.93]) than DQs (0.72 [0.33-0.85]). When values for women in Varese and Greece in both genders, which deviate from the others, are excluded, the partial correlation between urinary nitrogen and 24-HDR means goes up to 0.90 [0.77-0.96]. Inversely, a higher correlation with DQs (0.82 [0.60-0.92]) than 24-HDRs (0.69 [0.38-0.86], 0.77 [0.52-0.90], after exclusion of Greece and women in Varese) was observed between total energy intakes and urinary nitrogen. When genders are considered separately, the ecological correlation appears higher in men than in women with both dietary methods **(see annex 1 a,b)**. Whereas the same correlation is observed in DQ (0.91 [0.66-0.98]) and 24-HDRs (0.94 [0.76-0.99]) in men, the correlation is better with the 24-HDRs (0.69 [0.19-0.91]) than with the DQs (0.57 [-0.006-0.86]) in women, particularly when Varese and Greece are excluded (0.84 [0.51-0.95]). For total mean energy, the correlation with urinary nitrogen is quite comparable with both dietary methods (0.68 [0.17-0.90] vs 0.65 [0.12-0.89]) in women, but higher for men with DQs (0.93 [0.72-0.98] than with 24-HDRs (0.70 [0.13-0.92])

The regression analysis showed that the β regression coefficients are not significantly different from the ideal value of 1 when regressing mean nitrogen from 24-HDR or DQ (after adjustment for sex) on urinary estimates, both when all sex and centres are considered together and when analysed separately, except for estimates from 24-HDRs in men. Lower slopes were consistently observed with DQ measurements. This test could not be performed in order to compare energy intakes (Kcal) and urinary nitrogen estimates (grams), because the two measurements are expressed on different scales.

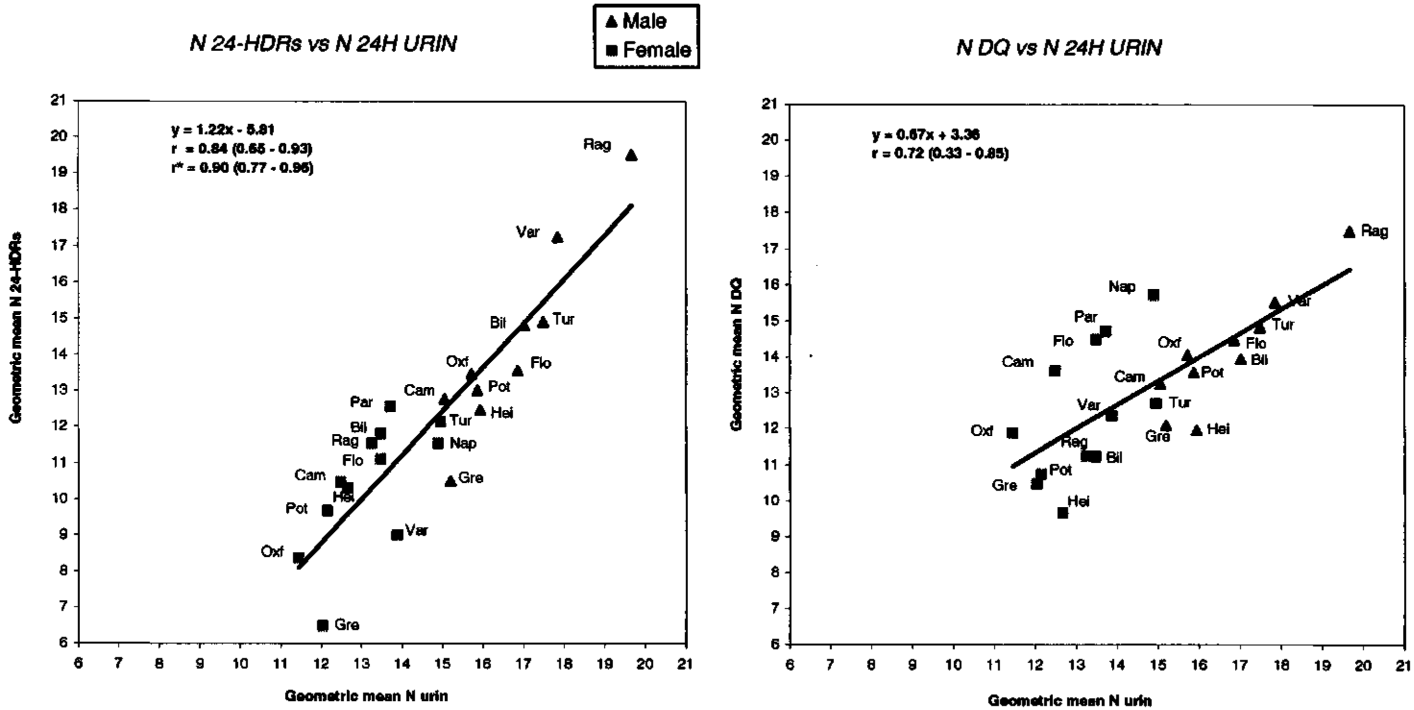
Discussion

The principal aim of this analysis was to evaluate the validity of mean nitrogen intake and indirectly total energy intake estimated from 24-HDRs and determine whether the method is able to rank the centres correctly when compared to mean urinary nitrogen. If this is the case, the use of 24-HDRs for between-cohort calibration of EPIC DQs would be justified. The same dietary variables, obtained from the non-standardized dietary methods (DQs) applied on all EPIC individuals at baseline, were also compared to the reference urinary measurements in order to estimate the additional benefit of the calibration approach in EPIC to correct systematic mean population bias.

As far as we know, this is the first time that dietary measurements used for between-cohort calibration have been validated using an independent biological reference measurement such as urinary nitrogen. The convenience sample used in the analysis is, however, not strictly representative of the EPIC cohorts, neither does it include all EPIC centres. Any extrapolation to the entire EPIC cohort or to the general population needs to be made with caution. Furthermore, in most centres it was not possible to estimate the response rate and the reasons for non-participation, which may have biased the results.

Although the time interval between the collection of urine specimens and dietary data, particularly dietary questionnaires varied considerably across centres, no statistically significant correlation was observed between the centre mean time interval and the difference in mean urinary and 24-HDR nitrogen measurements. This suggests the effect of time interval measurements on correlations at population levels in this study were, at the most, modest.

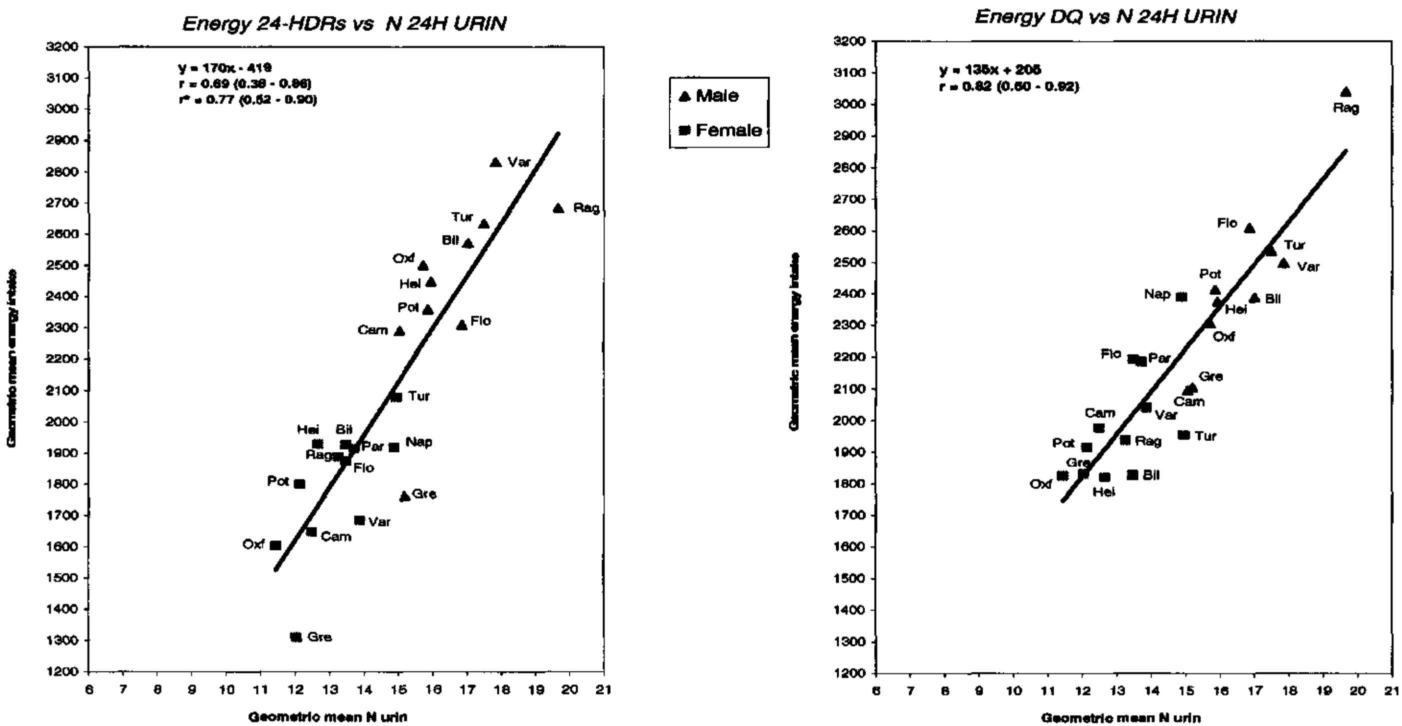
Figure 1: Pearson's ecological correlation (confidence intervals) adjusted for sex between centre geometric means of 24-hour urinary nitrogen and mean nitrogen from 24-hour diet recalls (24-HDRs) or dietary questionnaires (DQ) (N=22)



*After exclusion of Greece and Varese

Par : Paris	*Flo* : Florence	*Var* : Varese	*Tur* : Turin	*Rag* : Ragusa	*Nap* : Naples
Cam : Cambridge	*Oxf* : Oxford	*Bil* : Bilthoven	*Hei* : Heidelberg	*Pot* : Potsdam	*Gre* : Greece

Figure 2: Pearson's ecological correlation (confidence intervals) adjusted for sex between centre geometric means of 24-hour urinary nitrogen and total energy from 24-hour diet recalls (24-HDRs) or dietary questionnaires (DQ) (N=22)



*After exclusion of Greece and Varese

Par : Paris	*Flo* : Florence	*Var* : Varese	*Tur* : Turin	*Rag* : Ragusa	*Nap* : Naples
Cam : Cambridge	*Oxf* : Oxford	*Bil* : Bilthoven	*Hei* : Heidelberg	*Pot* : Potsdam	*Gre* : Greece

Higher gender-partial Pearson correlations between centre mean urinary nitrogen and nitrogen intakes were observed with 24-HDR compared to DQs, when both sexes and centres were considered together. When genders were considered separately, urinary nitrogen was well correlated with both dietary measurements in men while in women the correlation was higher with 24-HDRs than with DQ, particularly when Varese and Greece with outlier values were excluded. Despite the evidence of underreporting in mean estimates reported in Tables 3a and b, and the relatively low within- and between-subjects physiological variations associated with both nitrogen and total energy estimates reported in the literature (34-37), both dietary methods ranked centres well when compared to urine measurements, particularly in men. These correlations did not change significantly when centre means were adjusted further for age and/or BMI. In addition, the β coefficients estimated by regressing mean nitrogen intakes from 24-HDR and DQ (after adjustment for gender) on mean urinary nitrogen were not statistically different from the ideal value of 1, nor in gender-specific analyses, except for men for 24-HDRs. This suggests that, although mean centre/country dietary nitrogen intakes are not correct in absolute values, the order and magnitude of the systematic bias (generally under-estimation) is quite comparable across centres with both dietary measurements. In the absence of a completely unbiased reference dietary method, a constant under- (or over-) estimation can still fulfil one of the basic statistical requirements of re-scaling group mean values obtained with non-standardized DQs (38). Furthermore, it suggests that possible methodological errors in nitrogen intake estimation (e.g. lack of standardized food composition tables and different baseline dietary questionnaires) might be quite comparable across centres and/or are small compared to the actual differences in mean nitrogen (or total energy) intakes across centres.

Although urinary nitrogen, corrected for extra-renal losses, is a good reference measurement for estimating under-reporting of nitrogen (protein) dietary intakes and identifying gross under-reporters, it does not provide a precise estimate of the overall degree of under- or over-reporting of total energy intakes (39). Since doubly labelled water measurements were not collected in our study, urinary nitrogen was also used as an indirect and surrogate reference of total energy intakes. This assumption is based on the relatively high physiological correlation existing between the intakes of total energy and its main sources, particularly protein. We considered that some insights, even indirect, of the correlation at the centre mean level between total energy intake and urinary nitrogen could provide useful information for understanding the nature and magnitude of the overall measurement errors.

Indeed a non-neutral differential bias of food and nutrient measurements has been reported in the literature when (high) under-reporters are compared to other population groups (13, 40-44). In most of these studies (13, 40-42, 44) under-reporting is usually associated with a relative increase in protein intakes, when expressed as a percentage of total energy, which suggests that proteins are relatively less likely to be grossly under- or over-estimated than other macro-nutrients. This was confirmed in a small observational study where it was observed that snacks, in contrast to main meals, were not well reported, and that, although total energy was under-estimated by about 14%, protein intakes were correctly estimated (43). Studying the validity of protein intake measurements only may lead to under-estimation of the overall actual bias, particularly if differential under- or over-reporting in different foods are suspected. Overall, our results show that the 24-HDRs and DQs used in EPIC rank centres by mean total energy intakes in close agreement with mean urinary nitrogen when all centres are considered together, and in men (0.87) when the genders are considered separately. In women, the correlation with energy intakes is lower with both the DQs (0.64) and the 24-HDRs (0.67 all centres; 0.74 after exclusion of two outliers).

The gender differences suggested by our results may be explained by different possible factors.

1) Two centres (Greece, Varese in Italy) were identified as outliers, which affected the Pearson correlation calculated between 24-HDR and reference measurements in women. For example, the correlation between mean urinary nitrogen and nitrogen estimated from 24-HDRs increased from 0.69 to 0.84 after exclusion of these two centres. In contrast to Varese, where the problem is observed only in women, Greek subjects of both sexes systematically under-report by ≥ 30% with the 24-HDRs but to a lesser extent with DQs. We re-analyzed the Greek 24-HDR in detail, and no systematic errors in the dietary interview procedure was found which could explained this high under-reporting. A low level of completeness of urine samples estimated with the PABA check (20%) compared to the other EPIC centres suggest that a possible effect of subjects' behavior cannot be excluded, as indicated by other studies on EPIC data (45).

2) We observed that men under-report protein intake systematically with both dietary methods compared to urinary measurements, whereas women systematically under-report with 24-HDRs but over- or under-report in DQ depending on the centre. In Bilthoven, Germany, Greece and certain Italian centres, women under-report whereas in the two British centres, France, Florence and Naples, they tend to over-report by about 7-9%, although the difference is not always statistically significant from zero.

3) As suggested by the reference urinary measurements, the range of mean nitrogen intakes is quite narrow (~ 23% in both genders) compared to other dietary components, but does not overlap between women (11.44 to 14.8 g) and men (15.1 to 19.7 g). Apart from Greece, the same figure is observed with 24-HDRs whereas for DQs several centres have averages contained within the range of the opposite gender. Compared to urine measurements, the range of consumption observed is, however, greater with both dietary methods, highlighting random errors in the measurements, particularly with DQs in women.

4) The absolute range of variation, particularly of total energy intake, is narrower in women so random measurement errors are more likely to distort the correlation in women than in men. This may be due to the greater range of average work and leisure time activities in men than in women which has been observed across EPIC centres (46). Using dietary and doubly-labelled water as reference measurements, Black et al. (47) showed that women tend to report on average 11% less energy than men, after weight, height and age (i.e. for gender physiological differences) are controlled for.

5) These results may also suggest that women are more likely, consciously or unconsciously, to under- (or mis-) report their diet than men. This is supported by other analyses on the EPIC data (45) and some other studies (48-50). In respiratory room conditions, it has been shown that there are no gender differences in energy expenditure, after body composition and physical activity are controlled for (51). Goldberg's approach, expressed as the percentage of subjects with non-plausible physiological total energy values at the individual level, was used as an indirect (but not exhaustive) empirical indicator of the overall quality of the data and as a possible criterion for exclusion of extreme values. In our analysis, the exclusion according to Goldberg's cut-off showed that under- (or over-) estimation of total energy or nitrogen intakes of the mean groups were unlikely to be due exclusively to subjects reporting extreme levels of consumption. In most cases, the exclusions according to Goldberg did not change the statistical significance of the test difference between mean dietary nitrogen and total energy intakes and mean reference measurements, except when a high proportion of subjects were excluded, particularly with DQs. The lower number of subjects outside plausible physiological ranges according to the Goldberg cut-off suggests that 24-HDRs are less prone to such extreme values than DQs which, in addition, vary to a greater extent across centres. Although our results suggest a more systematic problem of measurement errors not specifically attributable, in most cases, to extreme mis-reporters, the lack of sensitivity of the current Goldberg cut-off points recently reported by Black et al. (52) cannot be disregarded as a possible cause.

In conclusion, the analysis shows overall comparable results using 24-HDR or DQs to rank centres according to mean level nitrogen or total energy intakes. Surprisingly, despite the lack of standardisation, the dietary questionnaires considered in this analysis provide relatively good ranking of centres according to their average mean total nitrogen or energy intakes. However, in contrast to 24-HDRs, errors in DQs are not only different in magnitude across countries but also in direction (over- and under-estimation), particularly in women. This may cause problems of misclassification, when pooling different individual questionnaire data without any correction for systematic bias in the baseline measurements, which need to be investigated further. Future validation studies on calibration measurements should be designed to have greater statistical power and include more representative samples from all EPIC centres, particularly those from the Nordic countries which joined EPIC later and used different dietary questionnaires (semi-quantitative FFQs) than the more extensive quantitative dietary questionnaires used in this analysis (Riboli et al., 2002). Furthermore studies using doubly labelled water measurements are needed to better evaluate the validity of total energy intakes.

Acknowledgements

The work described in the paper was carried out with the support of the "Europe Against Cancer" Programme of the European Commission (SANCO); Ligue contre le Cancer (France); Société 3M (France); Mutuelle Générale de l'Education Nationale; Institut National de la Santé et de la Recherche Médicale (INSERM); German Cancer Aid; German Cancer Research Center; German Federal Ministry of Education and Research; Cancer Research UK; Medical Research Council, UK; the Stroke Association, UK; British Heart Foundation; Department of Health, UK; Food Standards Agency, UK; the Wellcome Trust, UK; Greek Ministry of Health; Greek Ministry of Education; Italian Association for Research on Cancer; Italian National Research Council. In addition, we wish to thank all study participants for their cooperation and all the persons who participated to the field work studies in each EPIC centre.

References

1 Thompson FE, Moler JE, Freedman LS, Clifford CK, Stables GJ, Willett WC. Register of dietary assessment calibration-validation studies: a status report. Am J Clin Nutr 1997;65:S1142-7.

2 Kaaks R, Riboli E. The role of multi-centre cohort studies in studying the relation between diet and cancer. Cancer Lett 1997;114:263-70.

3 Stram DO, Hankin JH, Lynne R et al. Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. Am J Epidemiol 2000;151:358-70.

4 Bingham SA, Cummings JH. Urine nitrogen as an independent validatory measure of dietary intake: a study of nitrogen balance in individuals consuming their normal diet. Am J Clin Nutr 1985;42:1276-89.

5 Schoeller DA. Measurement of energy expenditure in free-living human using doubly labeled water. J Nutr 1988;118:1278-89.

6 Kaaks R, Riboli E, Sinha R. Biochemical markers of dietary intakes. In: Toniolo P, Boffetta P, Shuker D, Rothman N, Hulka B, Pearce N, eds. Application of biomarkers in cancer epidemiology (IARC Sci Publ No 142). Lyon: International Agency for Research on Cancer, 1997:103-26.

7 Steen B, Isaksson B, Svanborg A. Intake of energy and nutrients and meal habits in 70 year old males and females in Gothenburg Sweden: a population study. Acta Med Scand 1977;611:S39-86

8 Warnold I, Carlgren G, Krotkiewski M. Energy expenditure and body composition during weight reduction in hyperplastic obese women. Am J Clin Nutr 1978;31:750-63.

9 Isaksson B. Urinary nitrogen as a validity test in dietary surveys. Am J Clin Nutr 1980;33:4-12.

10 Bingham SA, Wiggins HS, Englyst H et al. Methods and validity of the dietary assessments in four Scandinavian populations. Nutr Cancer 1982;4:23-33.

11 Hughes JM, Smithers G, Gay C et al. The British National Diet and Nutrition Survey of people aged 65 years and over: protocol and feasibility study. Proc Nutr Soc 1995;54:631-43.

12 Van Staveren WA, de Boer JA, Burema J. Validity of the dietary history method. Am J Clin Nutr 1985;42:554-9.

13 Bingham SA, Cassidy A, Cole T et al. Validation of weighed records and other methods of dietary assessment using the 24 h urine nitrogen technique and other biological markers Br J Nutr 1995;73:531-50.

14 Margetts BM, Pietinen P, Riboli E, eds. EPIC: European Prospective Investigation into Cancer and Nutrition: Validity Studies on Dietary Assessment Methods. Int J Epidemiol 1997;26(Suppl. 1):S1-189.

15 Rosner B, Willett WC, Spiegelman D. Correction of logistic relative risk estimates and confidence intervals for systematic within-person measurement error. Stat Med 1989;8:1051-69.

16 Plummer M, Clayton D, Kaaks R. Calibration in multi-centre cohort studies. Int J Epidemiol 1994;23:419-26.

17 Kaaks R, Plummer M, Riboli E, Estève J, van Staveren WA. Adjustment for bias due to errors in exposure assessments in multi-centre cohort studies on diet and cancer: a calibration approach. Am J Clin Nutr 1994;59:S245-50.

18 Riboli E, Hunt K, Slimani N et al. EPIC: Study populations and data collection. Publ Health Nutr (in press).

19 Overvad K, Tjønneland A, Haraldsdóttir J, Bang S, Ewertz M, Møller-Jensen O. Development of a semi-quantitative food frequency questionnaire to assess food, energy and nutrient intake in Denmark. Int J Epidemiol 1991;20:906-12.

20 Bingham SA, Gill C, Welch A et al. Comparison of dietary assessment methods in nutritional epidemiology: weighted records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. Br J Nutr 1994;72:619-43.

21 Riboli E, Elmståhl S, Saracci R, Gullberg B, Lindgärde F. The Malmö Food Study: Validity of two dietary assessment methods for measuring nutrient intakes. Int J Epidemiol 1997;26:S161-71.

22 Kroke A, Klipstein-Grobusch K, Voss S, Moseneder J, Thielecke F, Noack R, Boeing H. Validation of a self-administrated food-frequency questionnaire administrated in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with a doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. Am J Clin Nutr 1999; 70(4):439-47

23 Slimani N, Kaaks R, Ferrari P et al. EPIC calibration study: Rationale, design and population characteristics. Publ Health Nutr (in press).

24 van Kappel AL, Amoyel J, Slimani N, Vozar B, Riboli E. EPIC-SOFT Picture book for estimation of food portion sizes. Lyon: International Agency for Research on Cancer, 1994.

25 Slimani N, Deharveng G, Charrondière RU et al. Structure of the standardized computerized 24-hour diet recall interview used as reference method in the 22 centres participating in the EPIC project. Comput Meth Programs Biomed 1999;53:251-66.

26 Slimani N, Ferrari P, Ocké M et al. Standardization of the 24-hour diet recall calibration method used in the European Prospective Investigation into Cancer and Nutrition (EPIC): general concepts and preliminary results. Eur J Clin Nutr 2000;54:900-17.

27 Slimani N, Charrondière UR, van Staveren W, Riboli E. Standardization of food composition databases for the European Prospective Investigation into Cancer and Nutrition (EPIC): General theoretical concept. J Food Comp Anal 2000;13:567-84.

28 Charrondière UR, Vignat J, Møller A. The European Nutrient Database (ENDB) for Nutritional Epidemiology. J Food Comp Anal (in press).

29 Deharveng G, Charrondière R, Slimani N, Riboli E, Southgate DAT. Comparison of food composition tables available in the nine European countries participating in EPIC. Eur J Clin Nutr 1999;53:60-79.

30 Bingham SA, Williams R, Cole TJ, Price CP, Cummings JH. Reference values for analytes of 24-h urine collections known to be complete. Ann Clin Biochem 1988;25:610-9.

31 Johansson G, Bingham S, Vahter M. A method to compensate for incomplete 24 h urine collections in nutritional epidemiology studies. Publ Health Nutr 1999;2:587-91.

32 Goldberg GR, Black AE, Jebb SA et al. Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. Eur J Clin Nutr 1991;45:569-81.

33 FAO/WHO/UNU. Energy and protein requirements. Report of a joint expert consultation (WHO Tech Rep Series No 724). Geneva: WHO, 1985.

34 Beaton GH, Milner J, McGuire V, Feather TE, Little JA. Source of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr 1983;37:986-95.

35 Bingham SA. The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. Nutr Abstr Rev (Series A) 1987;57,705-42.

36 Nelson M, Black AE, Morris JA, Cole TJ. Between- and within-subject variation in nutrient intake from infancy to old age: Estimating the number of days required to rank dietary intakes with desired precision. Am J Clin Nutr 1989 50:155-67.

37 Black AE, Cole TJ. Within- and between-subject variation in energy expenditure measured by the doubly-labelled water technique: implications for validating reported dietary energy intake. Eur J Clin Nutr 2000;54:386-394.

38 Riboli E, Kaaks R. Invited commentary: the challenge of multi-center cohort studies in the search for diet and cancer. Am J Epidemiol 2000;151:371-4.

39 Black AE, Bingham SA, Johansson G, Coward WA. Validation of dietary intakes of protein and energy against 24 hour urinary N and DLW energy expenditure in middle-aged women, retired men and post-obese subjects: comparisons with validation against presumed energy requirements. Eur J Clin Nutr 1997;51:405-13.

40 Rutishauser IH. Is dietary underreporting macronutrient specific?. Eur J Clin Nutr 1995;49:219-20.

41 Pryer JA, Vrijheid M, Nichols R, Kiggins M, Elliott P. Who are the 'low energy reporters' in the dietary and nutritional survey of British adults?. Int J Epidemiol 1997;26:146-54.

42 Price GM, Paul AA, Cole TJ, Wadsworth ME. Characteristics of the low-energy reporters in a longitudinal national dietary survey. Br J Nutr 1997;77:833-51.

43 Poppitt SD, Swann D, Black AE, Prendice AM. Assessment of selective under-reporting of food intake by both obese and non-obese women in a metabolic facility. Int J Obesity 1998;22:303-11.

44 Voss S, Kroke A, Klipstein-Grobusch K, Boeing H. Is macronutrient composition of dietary intake data affected by underreporting? Results from the EPIC-Potsdam Study. European Prospective Investigation into Cancer and Nutrition. Eur J Clin Nutr 1998;52:119-26.

45 Ferrari P, Slimani M, Ciampi A et al. Evaluation of mis-reporting in the 24-hour diet recalls in EPIC. Publ Health Nutr (in press)

46 Haftenberger M, Schuit AJ, Tormo MJ et al. Physical activity of subjects aged 50 to 64 years involved in the European Prospective Investigation into Cancer and Nutrition (EPIC). Publ Health Nutr (in press)

47 Black AE, Coward WA, Cole TJ, Prentice AM. Human energy expenditure in affluent societies: an analysis of 574 doubly-labelled water measurements. Eur J Clin Nutr 1996;50:72-92.

48 Johnson RK, Goran MI, Poehlman ET. Correlates of over- and underreporting of healthy older men and women. Am J Clin Nutr 1984;59;1286-90.

49 Klesges RC, Eck LH, Ray JW. Who underreports dietary intake in a dietary recall? Evidence from the second National Health and Nutrition Examination Survey. J Consult Clin Psychol 1995;63,438-44.

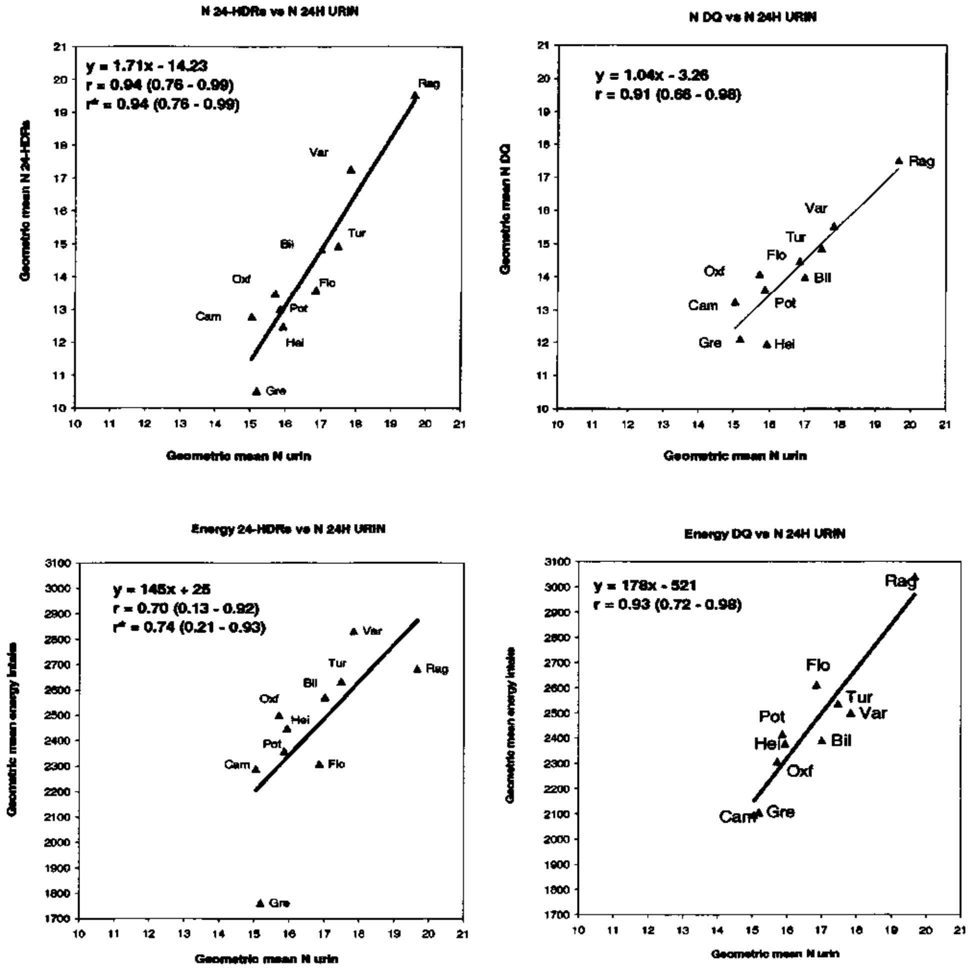
50 Briefel RR, Sempos CT, McDowel MA, Chien S, Alaimo K. Dietary method research in the Third National Health and Nutrition Examination Survey: Underreporting of energy intake. Am J Clin Nutr 1997;65:S1203-9.

51 Klausen B, Toubro S, Astrup A. Age and sex effects on energy expenditure. Am J Clin Nutr1997;65:895-907

52 Black AE. The sensitivity and specificity of the Goldberg cut-off for EI:BMR for identifying diet reports of poor validity. Eur J Clin Nutr 2000;54:394-404

Annex 1a: Pearson's ecological correlation (confidence intervals) between center geometric means of 24-hour urinary nitrogen and mean nitrogen (or total energy) from 24-hour diet recalls (24-HDRs) or dietary questionnaires (DQ)

Men (N=10)

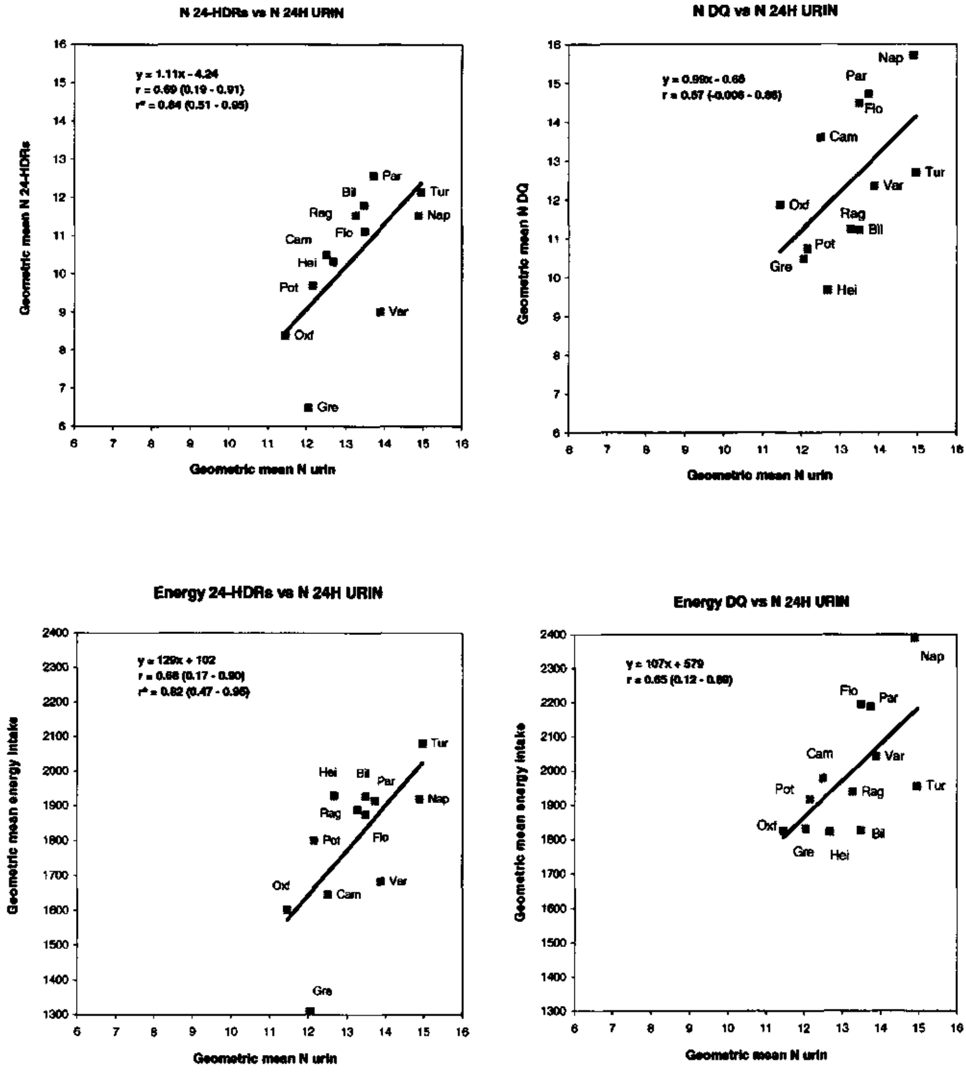


* After exclusion of Greece

Par : Paris	*Flo* : Florence	*Var* : Varese	*Tur* : Turin	*Rag* : Ragusa	*Nap* : Naples
Cam : Cambridge	*Oxf* : Oxford	*Bil* : Bilthoven	*Hei* : Heidelberg	*Pot* : Potsdam	*Gre* : Greece

Annex 1b: Pearson's ecological correlation (confidence intervals) between center geometric means of 24-hour urinary nitrogen and mean nitrogen (or total energy) from 24-hour diet recalls (24-HDRs) or dietary questionnaires (DQ)

Women (N=12)



* After exclusion of Greece

Par : Paris	*Flo* : Florence	*Var* : Varese	*Tur* : Turin	*Rag* : Ragusa	*Nap* : Naples
Cam : Cambridge	*Oxf* : Oxford	*Bil* : Bilthoven	*Hei* : Heidelberg	*Pot* : Potsdam	*Gre* : Greece

Chapter 7

Diversity of dietary patterns observed in the EPIC project

Nadia Slimani, Michael Fahey, Ailsa A Welch, Elisabet Wirfält, Connie Stripp, Erik Bergström, Jakob Linseisen, Mathias B Schulze, Christina Bamia, Y Chloptsios, Fabrizio Veglia, Salvatore Panico, Bas H Bueno-de-Mesquita, Marga C Ocké, Magritt Brustad, Eiliv Lund, Carlos A González, A Barcos, Göran Berglund, Anna Winkvist, Angela Mulligan, Paul Appleby, Kim Overvad, Anne Tjønneland, Françoise Clavel-Chapelon, Emmanuelle Kesse, Pietro Ferrari, Wija A Van Staveren, Elio Riboli

Public Health Nutrition (in press)

Abstract

The objective of this study was to describe the diversity in dietary patterns existing across centres/regions participating in EPIC. Single standardized 24-hour dietary recall measurements, collected for a sub-sample of 35,955 men and women aged 35–74 years and participating in the EPIC nested calibration study, were considered in the analysis. A graphic multi-dimensional representation was used to compare adjusted mean consumption of 22 food groups. Although wide differences were observed across centres, the countries participating in EPIC are characterized by specific dietary patterns. Overall, Italy and Greece have a dietary pattern characterized by plant foods (except potatoes) and a lower consumption of animal and processed foods compared to the other EPIC countries. France and particularly Spain have more heterogeneous dietary patterns, with a relatively high consumption of both plant foods and animal products. Apart from characteristics specific to vegetarian groups, the UK "health-conscious" group shares with the UK general population a relatively high consumption of tea, sauces, cakes, soft drinks (women), margarine and butter. In contrast, the diet in the Nordic countries, the Netherlands, Germany and the UK general population is relatively high in potatoes and animal, processed and sweetened/refined foods, with proportions varying across countries/centres. In these countries, consumption of vegetables and fruit is similar to, or below, the overall EPIC mean, and is low for legumes and vegetable oils. Overall, dietary patterns were similar for men and women, although there were large gender differences for certain food groups. There are considerable differences in food group consumption and dietary patterns among the EPIC study populations. This large heterogeneity should be an advantage when investigating the relationship between diet and cancer and formulating new aetiological hypotheses related to dietary patterns and disease.

Introduction

Of all the environmental exposures diet is a *universal* exposure comprising a complex mixture of different compounds that varies over time, space and according to a number of historical, ethnic, religious, agricultural, socio-economic and psychological factors, at the individual and population levels (1). The lack of unbiased dietary methods and the difficulty of measuring long-term individual cumulative dietary exposure, when measurements are obtained at fixed point(s) in time of the subject's life, are important methodological limitations for estimating diet accurately as an exposure (2,3). These drawbacks may partly explain the lack of consistency in estimates of the association between diet and disease, particularly with regard to cancer (4), and the attenuation in relative risk estimates of the actual relationship between diet and the outcome diseases due to random errors in dietary measurements (5-7). Another emerging possible explanation, however, is the methodologies used so far to analyse diet–disease relationships that do not fully capture the nature and complexity of diet. For decades, the traditional practice has been to consider separately, in univariate analyses, the role of single foods, food groups or nutrients in relation to disease risk in order to distinguish their possible specific individual effects. The high inter-correlation between food (groups) and/or nutrient intakes and the difficulty of controlling for a number of possible dietary and other confounders in an analysis (8-11) may, however, also partly explain the distortion in the estimates of the association observed between specific dietary components and cancer in particular (12). In addition, current scientific evidence on carcinogenic mechanisms and the results from experimental and epidemiological studies suggest, for most cancers, multi-factorial causes involving several dietary and other factors, with possible synergetic or antagonist effects (13).

These considerations have led to increased interest in multivariate or multi-dimensional analyses of so-called dietary patterns which could better demonstrate that different foods and food groups are often consumed by individuals according to a reproducible pattern, despite large within- and between-subject variations (14,15). Various statistical techniques (e.g. factor or cluster analysis) (16) and other more innovative approaches (17) exist and are now increasingly used for studying the association between dietary patterns and cancers and other chronic diseases (18-22). Despite various methodological limitations (e.g. lack of stability and specificity of dietary patterns, and subjective selection of variables) (16,23), this area of exploratory research opens promising perspectives for a better understanding of the relationship between dietary exposure and chronic diseases (24,25).

The aim of this paper is to describe and highlight the diversity in dietary patterns existing across 27 centres/regions participating in the European Prospective Investigation into Cancer and Nutrition (EPIC) using a graphic multi-dimensional comparison of the mean consumption of 22 common food groups obtained from highly standardized computerized 24-hour dietary recall interviews.

Materials and methods

EPIC is a large cohort representing about half a million individuals in 10 western European countries (Denmark, France, Germany, Greece, Italy, Norway, Spain, Sweden, The Netherlands, United Kingdom) (26). The choice of study populations was not intended to provide representative samples, but was determined by practical and logistical considerations to obtain high participation and long-term follow-up (26). The study subjects were either population-based (Bilthoven in the Netherlands, Greece, Germany, Sweden, Denmark, Norway, Spain, Italy, Cambridge and a small part of the Oxford cohort from the UK), participants in breast cancer screening (Utrecht in the Netherlands and Florence, Italy), or teachers and school workers in France. In Oxford, most of the cohort was recruited among subjects with an interest in health and/or vegetarian eating habits and were self-defined vegans (i.e. consuming no animal products), ovo-lacto vegetarians, fish eaters (i.e. consumers of fish but not meat), and meat eaters. Blood donors were also recruited in different proportions in certain Italian and Spanish centres. In France, Norway, Utrecht (The Netherlands) and Naples (Italy) only women were recruited. Further details on the EPIC project are detailed elsewhere (26).

Information on usual individual dietary intakes was obtained using different dietary assessment methods developed and validated in each participating country (27-30). In order to adjust for possible systematic over- or underestimation in dietary intake measurements and correct for attenuation bias in relative risk estimates, a calibration approach was adopted. A single 24-hour dietary recall (24-HDR) was collected from a random sample of 5–12% (1.5% in the UK) of the EPIC cohorts, weighted according to the cumulative numbers of cancer cases expected per fixed age and sex stratum (31). Interviews were scheduled with the intention of obtaining an equal distribution of 24-HDRs according to day of the week and season.

The results presented in this paper are based on single 24-HDRs collected between 1995 and 2000 from 35,955 subjects (13,031 men and 22,924 women) of the 36,900 who participated in the calibration studies nested in the EPIC cohorts.

The statistical analysis was restricted to subjects aged 35–74 years; 945 subjects outside this range were excluded. In the UK, the "health-conscious" group and the subjects recruited from the general population both in Cambridge and Oxford (general population group) were considered as two separate population groups. In France and Norway, where the study subjects were scattered all over the country, the groups were sub-divided into, respectively, four and two geographical regions. Consequently, results for 27 centres/regions are presented in this analysis. However, for convenience, the term "centres" will be used to name both EPIC administrative centres and regions. Further details on the calibration study design and characteristics of the study participants are given elsewhere (31).

Computerized 24-hour dietary recall interview software (EPIC-SOFT) was developed to assess dietary intakes reported across the EPIC centres in a standardized manner (32,33). The same structure and translated interface was used in all 10 countries in order to standardize dietary interviews among the 90 interviewers involved in the calibration study. All the 24-hour dietary recalls were collected during a face-to-face interview, except in Norway where the interviews were conducted by telephone (34). Methods of quantification of portion sizes were standardized between countries using photographs, household measures depicted in pictures and standard units. For the telephone interviews, this material was mailed to the subjects in advance. Foods were classified according to the common EPIC-SOFT classification in 17 main groups and 124 sub-groups.

All main groups of the EPIC-SOFT classification covering the different components of diet in all countries were considered in the analysis, except "miscellaneous" and "soups". However, the groups "dairy products" ("milk" and "other dairy products"), "meat and meat products" ("fresh meat" and "processed meat"), "non-alcoholic beverages" ("tea", "coffee", "soft-drinks", "fruit & vegetable juices") and "fat" ("vegetable oils", "butter" and "margarine") were considered at the sub-group level as indicated in brackets. Twenty-two main (sub-) groups of the common EPIC-SOFT classification were finally considered in this analysis (**Table 1**). These food groups were ordered according to their plant or animal origins and their degree of food processing to facilitate the comparison of dietary patterns across centres. For the comparison of alcoholic beverages, which vary greatly in type (e.g. beer, wine, spirits) and alcohol content across EPIC study centres (35), we expressed alcohol consumption in grams of ethanol rather than millilitres of alcoholic beverages. Ethanol intake was computed using country-specific nutritional values.

The first purpose of the calibration study was to obtain good estimates of mean food intakes at the population level, and a single 24-HDR measurement was collected from each study participant. Our analysis is therefore focused on the comparison of *mean* dietary intakes. In order to compare dietary patterns across the 27 EPIC centres, crude and adjusted mean food consumption (g) for the ith food group, m(i), was calculated by sex and centre for each of the 22 food (sub-) groups. A standard multiple linear regression model was used to adjust for age, day of the week and season in order to correct for different distributions of these factors observed across the EPIC centres (31). Total energy intake was also added to the model, but as this did not change the overall dietary patterns in most centres, these results are not reported here. Overall EPIC mean consumption, M(i), was also calculated for the same food groups, by sex, as the arithmetic mean of the centre mean intakes. To express variation in centre mean intakes from the overall EPIC mean, percentage food intake relative to the EPIC mean was calculated for each food group, by sex and centre as: [100% * m(i)] / M(i)].

Table 1. Definitions and contents of the EPIC-SOFT food (sub-) groups considered in the analysis

Food groups	Definition and content
Vegetables	Leafy, fruiting, root, grain, pod and stalk vegetables, mushroom, allium, cruciferous, sprouts and mixed salad/vegetables
Fruits	Fresh fruits, nuts, seeds, stewed fruit, mixed fruits, and olives
Potatoes	Potatoes and potato products, except crisps
Legumes	Dried peas, lentils and beans, except soya
Cereals and cereal products	Flour, flakes, starches, pasta, rice and other grains, bread, crispbread, rusks, breakfast cereals, salty and aperitif biscuits, dough and pastry (puff, short-crust, pizza)
Cakes	Cakes, pies, pastries, puddings (non-milk-based), dry cakes, biscuits
Sugar and confectionery	Sugar, jam, marmalade, honey, chocolate and products, candy bars, confetti/flakes, drops, boiled sweets, chewing gum, nougat, cereal bars, marzipan, syrup, water ice
Added fats	
Vegetable oils	Vegetable oils
Margarines	Margarines, mixed dairy margarines, baking fat
Butter	Butter, herbal butter, butter concentrate
Dairy products	
Milk	Liquid milk (e.g. cow, goat), processed milk (condensed, dried), whey
Dairy products	Milk beverages, yogurt, cheeses, cream desserts, puddings (milk-based), dairy creams, ice cream
Meat and meat products	
Fresh meat	Beef, veal, pork, lamb/mutton, horse, goat, poultry, game, and offal
Processed meat	Processed meat from red meat or poultry (e.g. ham, bacon, sausages, pates etc.)
Eggs	Eggs (e.g. chicken, turkey, duck, goose, quail) and egg products, except if used for bread and bakery products
Fish and shellfish	Fish and fish products, crustaceans and molluscs
Sauces	Tomato sauces, dressing sauces, mayonnaises and similars
Non-alcoholic beverages	
Tea	Tea (with and without caffeine), iced tea: infusion, powder, instant beverage
Coffee	Coffee (with and without caffeine): infusion, powder, instant beverage
Soft-drinks	Carbonated/soft/isotonic drinks, diluted syrups
Fruit/vegetable juices	Fruit and/or vegetable juices and nectars, freshly squeezed juices: pure or diluted with water
Alcoholic beverages	Expressed as ethanol

A multi-dimensional graphic representation of the relative food intakes was used to illustrate differences in dietary patterns by centre and gender; their corresponding values are provided in **Annex 1**. EPIC means, used as the common denominator to calculate deviations, are indicated in each figure by a reference circle of radius 100%. If the relative consumption of a food group is above 100%, it indicates that the given centre is characterized by a relatively high consumption of that food group compared to the reference EPIC mean, and *vice versa* when the relative intake is below 100%. The same scale was used in the graphs for all countries and both genders (0–250%). The end peaks of means exceeding 250% are not reported in the graphs, but are indicated in Annex 1.

Results

Italy

In all the Italian centres, diet is dominated by plant foods, particularly cereals and cereal products, fruits and vegetable oils, when compared to the EPIC mean (**Figures 1 a & b**). In Varese however, intake of vegetable oils was much lower compared to the other centres. Several additional food groups characterize certain Italian centres and highlight further specific local dietary patterns. Legume consumption varies widely and is consistently much higher in women (except Turin) and in southern centres (Ragusa, Naples and Florence). Sauces, essentially tomato-based in these centres, are strong characteristics of diet in Ragusa and Varese in both genders but not in the other Italian centres, where consumption is around or below the EPIC mean. Alcohol is a characteristic of diet in men in Varese (124%) and to a lesser degree in Turin in both genders, compared to the other centres. Whereas fresh meat consumption is about 10–20% higher than the EPIC mean in men in Varese and in women from Florence, Turin and Ragusa, it is closer or below the EPIC mean elsewhere.

In contrast to fruit, vegetables are a relatively less important dietary characteristic in Italy, where consumption is around or below the EPIC mean, except in Turin (~135%) and to a lesser extent in Florence (~112%) in both genders. Consumption of potatoes is relatively low but similar across centres. Milk, butter and processed meat consumption is around or below the EPIC mean and follows a geographical gradient, with ~30% to ~60% higher intakes in northern centres compared to southern centres, and *vice versa* for fish and eggs (in men). Coffee, tea, soft drinks and juices are relatively rarely consumed in most centres and margarines and butter (in southern centres) are not consumed.

Greece

The Greek diet is strongly characterized by plant food groups such as vegetable oils, legumes, and vegetables in both genders when compared to the EPIC mean (**Figures 2 a & b**). In contrast to Italy and Spain, vegetable consumption is much higher than fruit consumption, with the latter below the EPIC mean in both genders. Consumption of cereal products is around the EPIC mean for both men and women. Consumption of animal products is around the mean for fish in men and below for fish in women, as are dairy products, fresh meat, eggs and milk for women. Alcohol (particularly among women), meat, juices, sugar products, cakes, soft drinks, potatoes and sauces and, to a greater extent, butter, coffee, margarine, processed meats, tea (<30%) are not important components of the Greek diet, when compared to the EPIC mean.

Spain

Spain has a complex dietary pattern characterized by a diet rich in both plant foods (legumes, vegetable oils, fruits and vegetables, except in Asturias) and animal food groups, particularly sea foods, eggs and milk, but with important local variations (**Figures 3 a & b**). Consumption of legumes, fish and eggs are important dietary characteristics in all Spanish centres, but particularly in northern centres. Milk is relatively highly consumed all over Spain in women, but greater centre differences are observed in men. High consumption of fresh (and processed) meat and alcohol (in men) is more specific of northern centres. Potato consumption is within ±10–15% of the EPIC mean, with 30–35% differences across centres. Consumption of cereal products, juices, soft drinks, cakes and sugar products is relatively low in all centres, except in Murcia where cake intake is around the EPIC mean. Tea, butter, margarine, coffee, and sauces are not characteristic of the diet in any of the Spanish centres when compared to the EPIC means.

France

In France, the four large geographical regions cover the entire country, including inland, Atlantic and Mediterranean areas. However, the same main food groups (i.e. butter, dairy products, fresh meats, alcohol and, to a lesser extent, vegetables) characterize the diet in these regions compared to other EPIC centres, although with different orders of magnitude (**Figure 4**). Butter consumption is higher in the North-west (290%) and North-east (243%) compared to the South (207%) and South coast (167%) and tea is particularly consumed in the North-west and South and to a lesser degree elsewhere. In contrast, consumption of dairy products, milk, eggs and fresh meat (except in the North-east, where it is higher) is quite similar across regions. Consumption of alcohol is about 25–30% above the EPIC mean, except in the South where it is lower than elsewhere in France. Fruit consumption is slightly below the EPIC mean whereas vegetable intake is higher in all regions (113–132%). Fish products are an additional strong characteristic of the North-west region (140%), but consumption is around the EPIC mean in the other regions.

Although consumption is around or below the EPIC mean, great variation (≥25%) by regions is also observed for potatoes, vegetable oils, legumes, coffee, cakes and margarines, whereas consumption of processed meats, juices, fruits, sugar products, cereals and milk varies only from 8% to 15% by region. Soft drinks, margarine, milk, and vegetable oils (except in the South) are the food groups consistently less consumed in France than in the other EPIC centres.

Germany

In both German centres, the same food groups characterize diet, i.e. butter, juices, followed by processed meat and coffee, with different orders of magnitude when compared to the EPIC mean (**Figures 5 a & b**). Margarine is a further important characteristic of diet in Potsdam (formerly in East Germany) (>230%) but not in Heidelberg (<60%). To a lesser extent, potato consumption is greater in Potsdam than in Heidelberg, and *vice versa* for alcohol in women. Other food groups, with lower relative consumption, such as soft drinks and sauces show values ≥25% higher in Heidelberg than in Potsdam and inversely for fruits (men). Both centres are characterized by relatively low consumption of vegetable oils, legumes, fish products and milk.

The Netherlands

As the Utrecht cohort includes only women, no cross-centre comparison can be made for men between the two Dutch centres. The dietary pattern in the Netherlands is dominated by margarine, tea, coffee, dairy products, sugar products, potatoes, processed meats, juices (women) and soft drinks (men), with consumption 25–150% above the reference EPIC mean (**Figures 6 a & b**). Except for butter and juices and alcohol in men, which are 10–15% above the EPIC mean, all the other food groups have a relative consumption around or below the EPIC mean. Marked differences are, however, observed among women, with higher values in Utrecht than in Bilthoven for tea, milk, dairy products and fruit, and *vice versa* for soft drinks and, to a lesser degree, sugar products, margarine and coffee. Vegetable oils, legumes, fish products and, to a lesser extent, vegetables and fruits (particularly in Bilthoven) are not typical of the Dutch diet.

United Kingdom

Various beverages (tea, soft drinks, and to a lesser extent, coffee and alcohol in women), butter, margarines, milk, sauces, cakes, potatoes and sugars (men) are major characteristics of the diet of the British general population when compared to the EPIC mean (**Figures 7 a & b**). The consumption of juices and dairy products is about 10–20% above the mean whereas for the other food groups it is around or below the EPIC mean. Vegetable oils and legumes (men) and, to a lesser degree, fresh meat and fruits are food groups that are eaten in relatively smaller amounts by the British general population.

The British "health-conscious" cohort includes a heterogeneous group of ovo-lacto vegetarians, pure vegans, fish (but not meat) eaters and meat eaters. The relatively low intake of animal products and the exclusive consumption of soy products (≥1600% deviation from the EPIC mean, not reported value), compared to the other centres, are the first strongest dietary characteristics of this population. Animal products such as fresh meat, processed meat, eggs and fish products are the food groups less representative of this population (≤40%), although dairy products, milk and particularly butter (140%) are consumed in greater amounts. As in Spain and Greece, a relatively high consumption of legumes is a characteristic of their diet, particularly in women (185%), compared to the other EPIC centres, whereas vegetables and fruits are a relatively weak characteristic, but still ~30–45% above consumption in the general population. Consumption of potatoes and cereal and cereal products is near the EPIC mean intake in men and respectively 25% and 10% higher in women. The relative consumption of alcohol is around the mean in women and more than 30% below in men. Fat intake comes mainly from margarines, butter (and sauces), all high markers of the health-conscious diet; vegetable oils, although more widely consumed than in the general British population, are not characteristic of the health-conscious diet, compared to the EPIC mean. The "health-conscious" group differs from the general population for both sexes. Differences between the two group populations are large and range from ~15% (e.g. sauces) to ≥1642% (soy products).

Denmark

In Denmark, coffee, soft drinks, alcohol, margarine and tea are primary characteristics of the diet when compared to the EPIC mean, but with variations by sex and centre (**Figures 8 a & b**). Among men, consumption of sauces, sugars, potatoes and milk (only in Aarhus) and, among women (from Aarhus only) eggs, cereals and cereal products, sauces, potatoes and sugar products is about 20–30% above the mean. The Danish diet is also characterized by several other food groups with consumption around or 15% above the EPIC mean such as dairy products, fresh meats, milk, fish

(Copenhagen) and eggs (except women in Aarhus). Consumption of most of the other food groups is below the EPIC mean, particularly for vegetables and fruits (except women in Aarhus) and, to an even greater degree, for legumes and vegetable oil.

Sweden

In Sweden, margarine, sauces, coffee, potatoes, sugar products, processed meats and cakes dominate the diet in both centres and genders, but with different orders of importance by sex and centre, when compared to the EPIC mean (**Figues 9 a & b**). Milk (men), dairy products and soft drinks are further strong specific characteristics of diet in Umeå (152–181%). Although relatively less consumed, coffee, sauces and juices have greater values in Malmö than in Umeå. All the other groups have consumption around or below the sex-specific EPIC means with intakes varying across centres and genders. Consumption is particularly low for vegetable oils, legumes, butter, and to a lesser extent, fruits, vegetables and alcohol (Umeå).

Norway

The Norwegian cohort included only women. Soft drinks, coffee, margarine, processed meats, juices, sauces and, to a lesser extent, sugar products, characterize diet in both regions (**Figure 10**). Fish products (173%) and potatoes (145%) are further important markers of women's diet in the North & West region, although consumption is still relatively high (~120%) in the South & East region compared to the EPIC mean. Intake of milk, eggs, cereal and cereal products, dairy products and cakes (South & East) is within ±5–10% of the EPIC mean in both centres, whereas intake of most of the other food groups is below the EPIC mean. Legumes, vegetable oils and, to a lesser extent, butter, fruit and vegetables are not characteristic of the Norwegian diet in either centre.

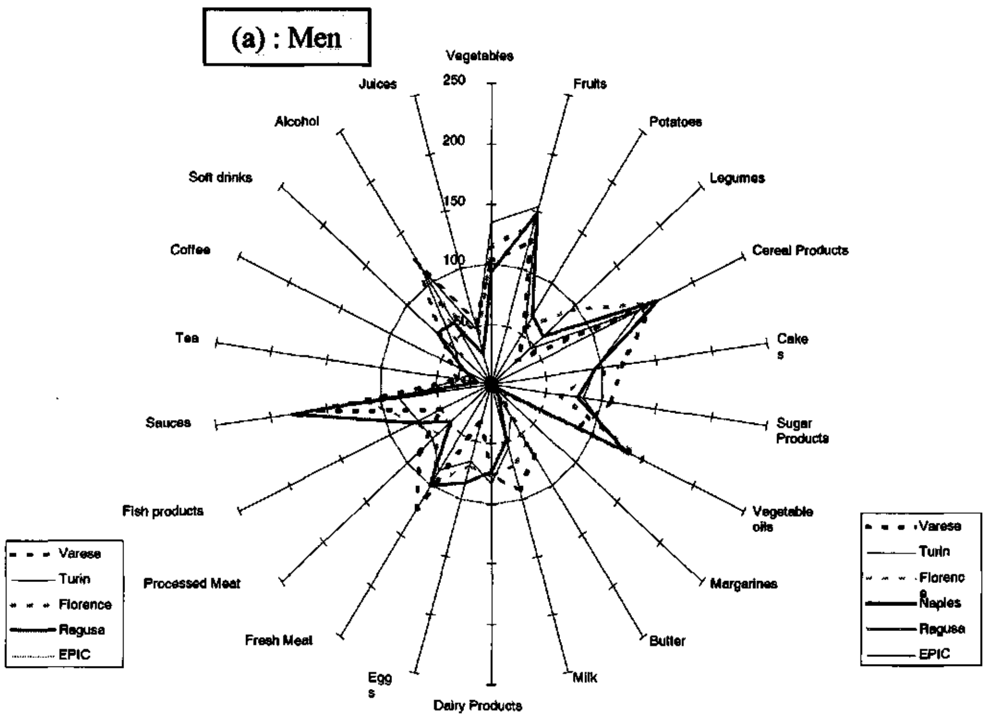
Discussion

The EPIC study was initiated in 1992 with the aim of combining several large European cohorts selected to maximize differences both in dietary exposure and disease outcome (36). The objective of this analysis was to describe and highlight the contrasts in dietary patterns observed across study centres, by comparing multiple food group consumption relative to the overall EPIC mean. The comparison of absolute intakes across centres, which is discussed in detail elsewhere in this supplement (35, 37-43), was therefore not the main concern of our study.

Our analysis has led to a first major conclusion that large differences in dietary patterns exist between the 27 Western European centres participating in EPIC. Different *types*, *numbers* and *magnitude of consumption* of food groups characterize these various dietary patterns. With few exceptions, we consistently observed that the number of food groups dominating the diet tends to increase from the Mediterranean dietary patterns observed, particularly in the Greek and Italian centres, to more or less pronounced Western dietary patterns observed elsewhere, with a strong south–north gradient within and between countries. Overall, we observed that in most of Western Europe, diets tended to be more heterogeneous than the Mediterranean diets, with a wider variety of food groups consumed; these are, however, essentially of animal origin or highly refined/processed, although exceptions and intermediate stages exist. However, the average total number of food items consumed per day is quite similar across centres (31). Further investigations on food choices, food consumption diversity and meal patterns will help to better understand the differences in dietary patterns across the EPIC centres.

Figures 1 to 10: Deviation (%) of the centre-adjusted (age, weekdays, seasons) food grop gr) means from the sex-speific "EPIC means"[a]

Figure 1. Italy



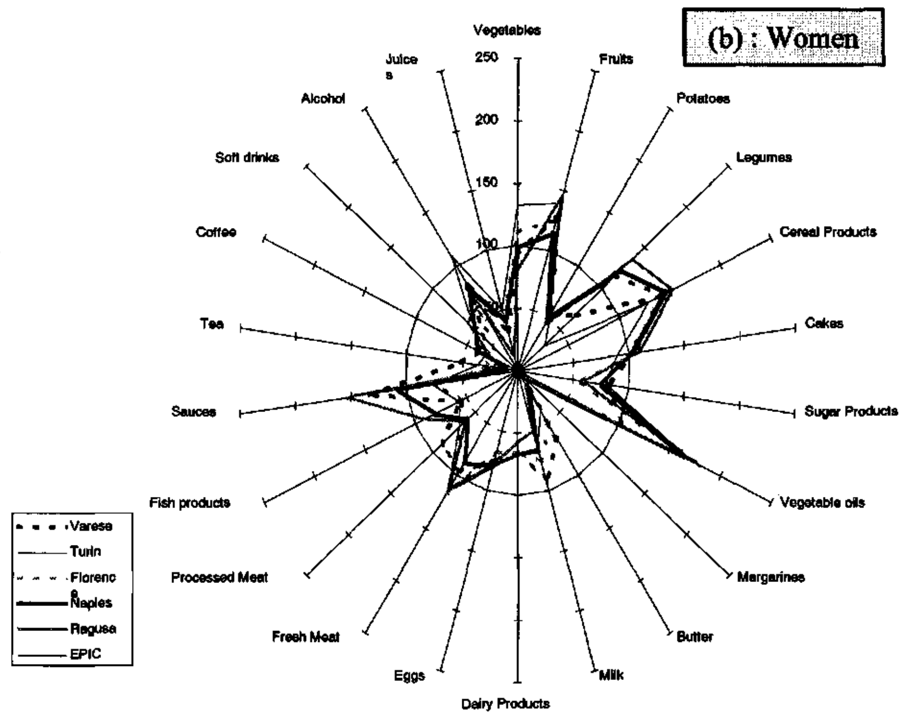
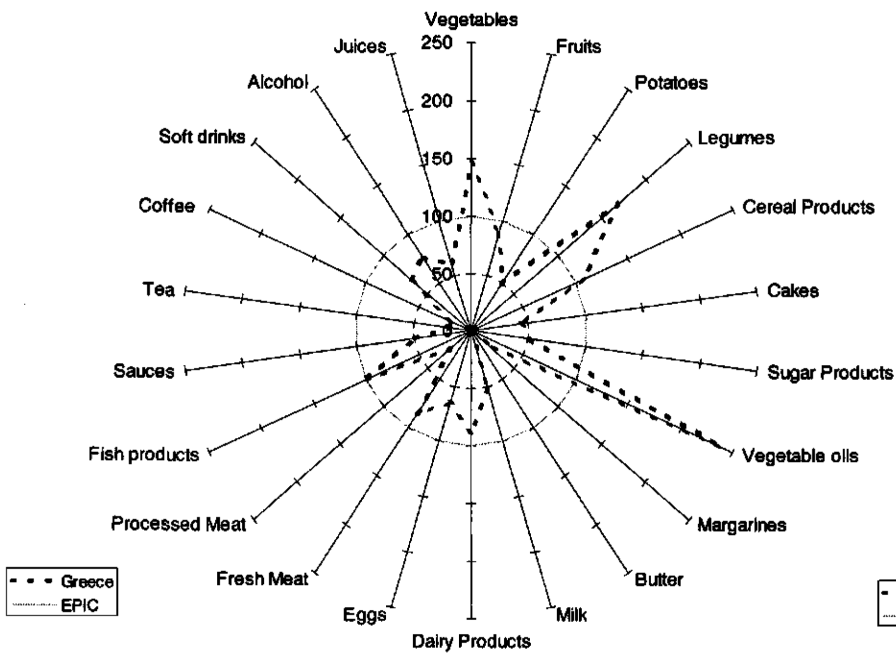


Figure 2. Greece

(a) : Men



(b) : Women

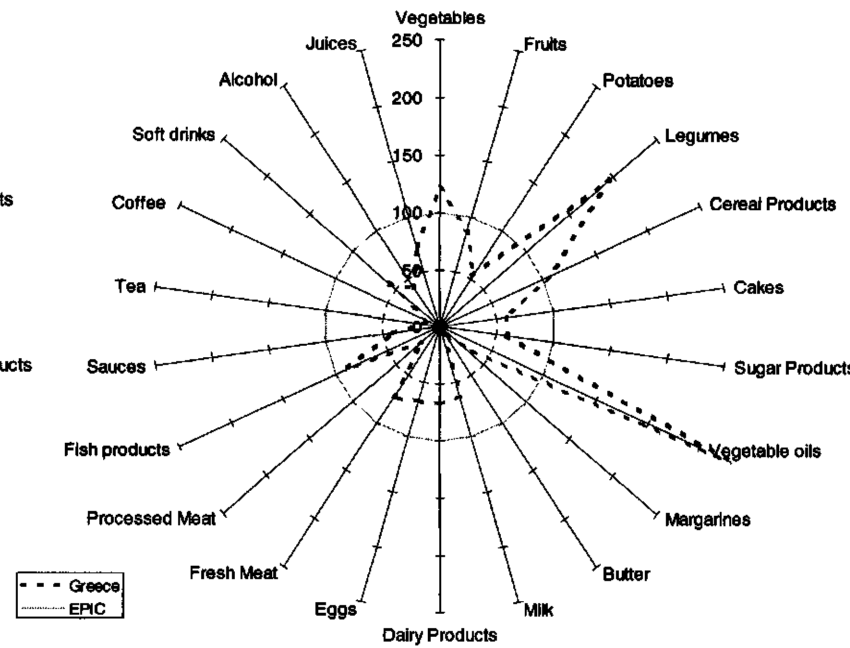


Figure 3. Spain

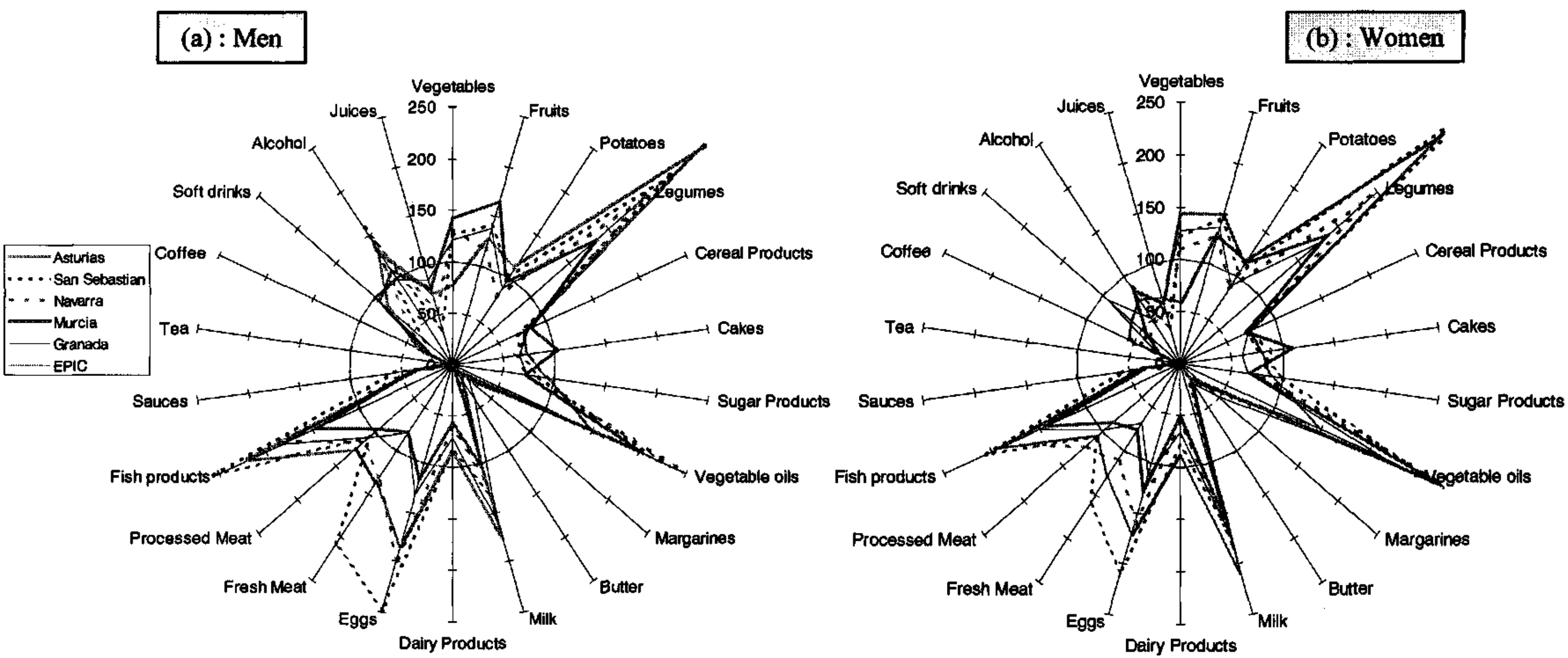


Figure 4. France

Women

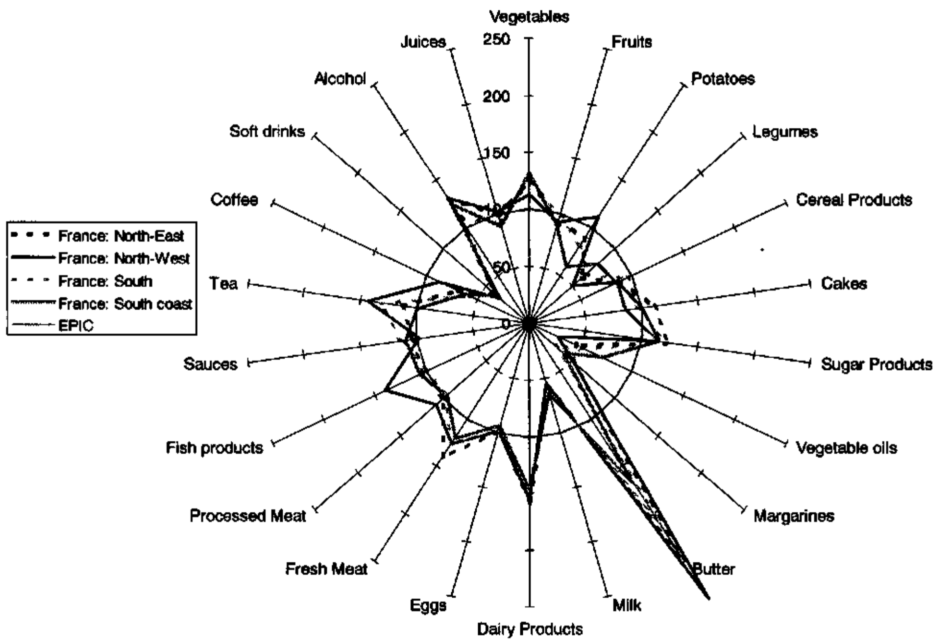
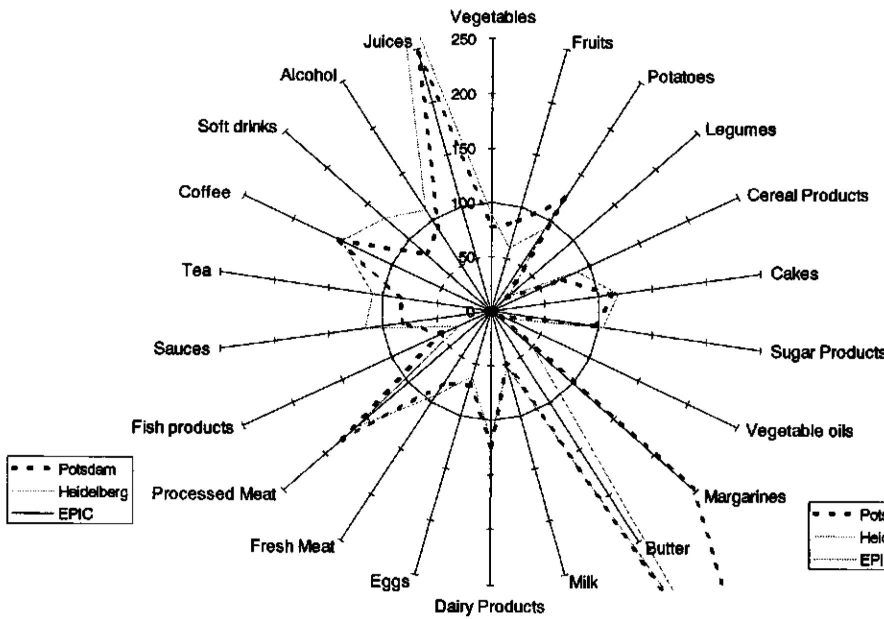


Figure 5.Germany

(a) : Men



(b) : Women

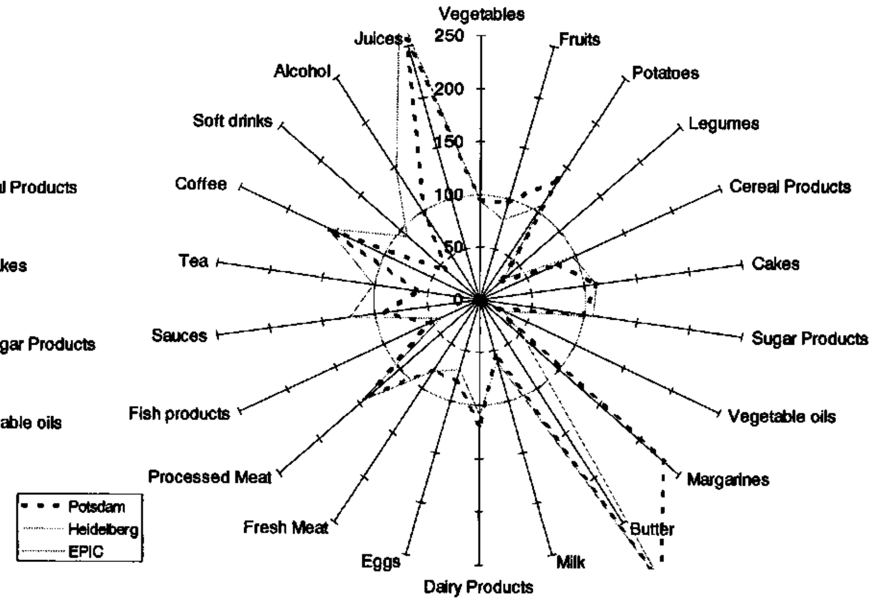
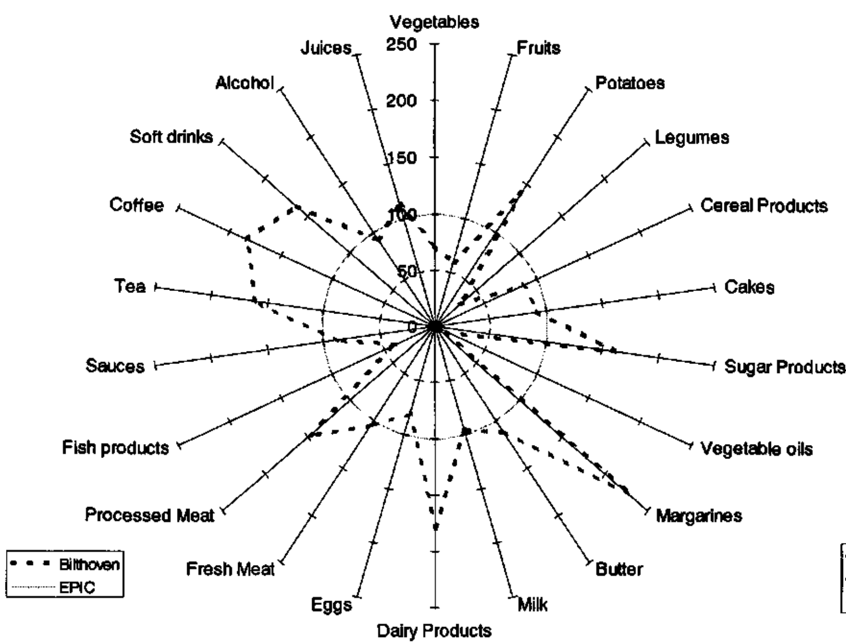


Figure 6. The Netherlands

(a) : Men



(b) : Women

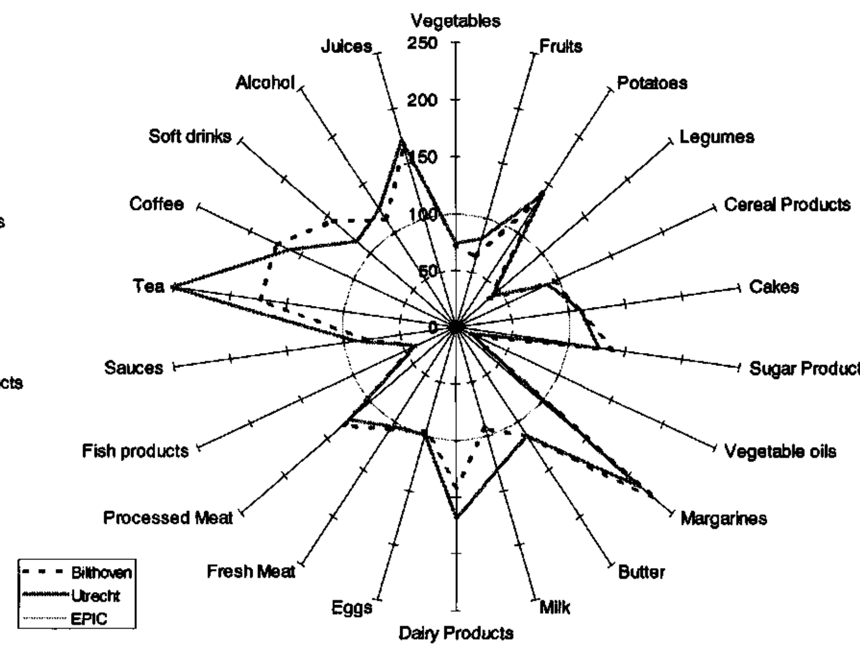
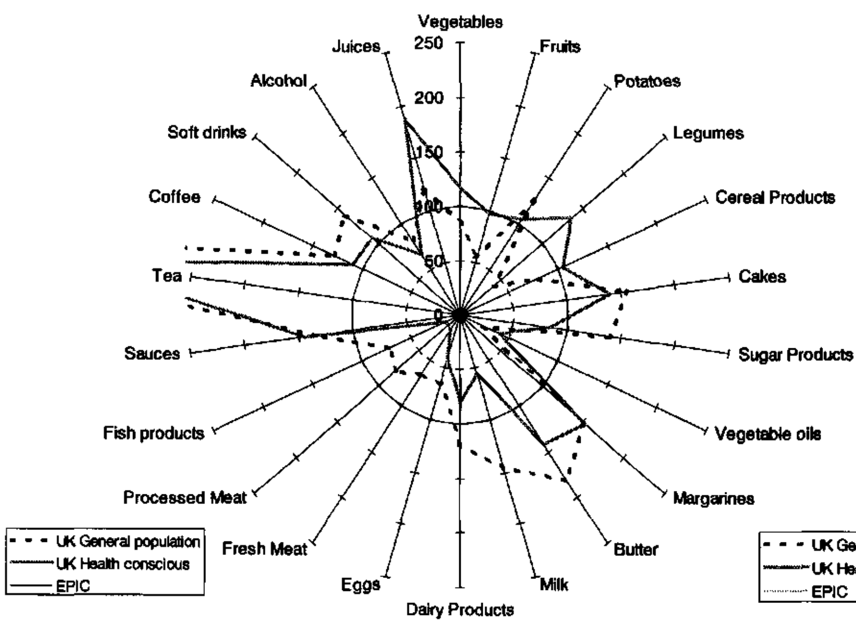


Figure 7. United Kingdom

(a) : Men



(b) : Women

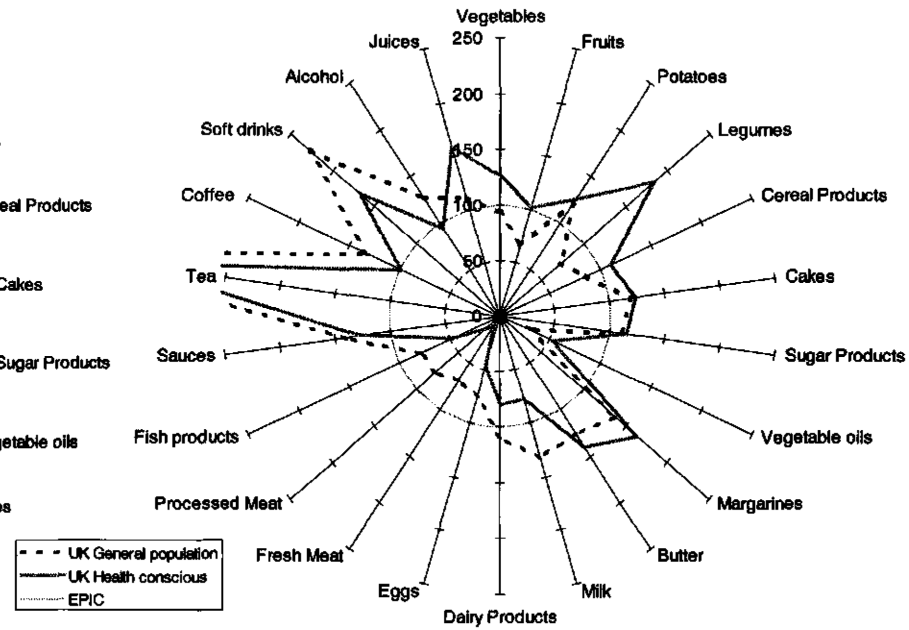


Figure 8. Denmark

(a) : Men

(b) : Women

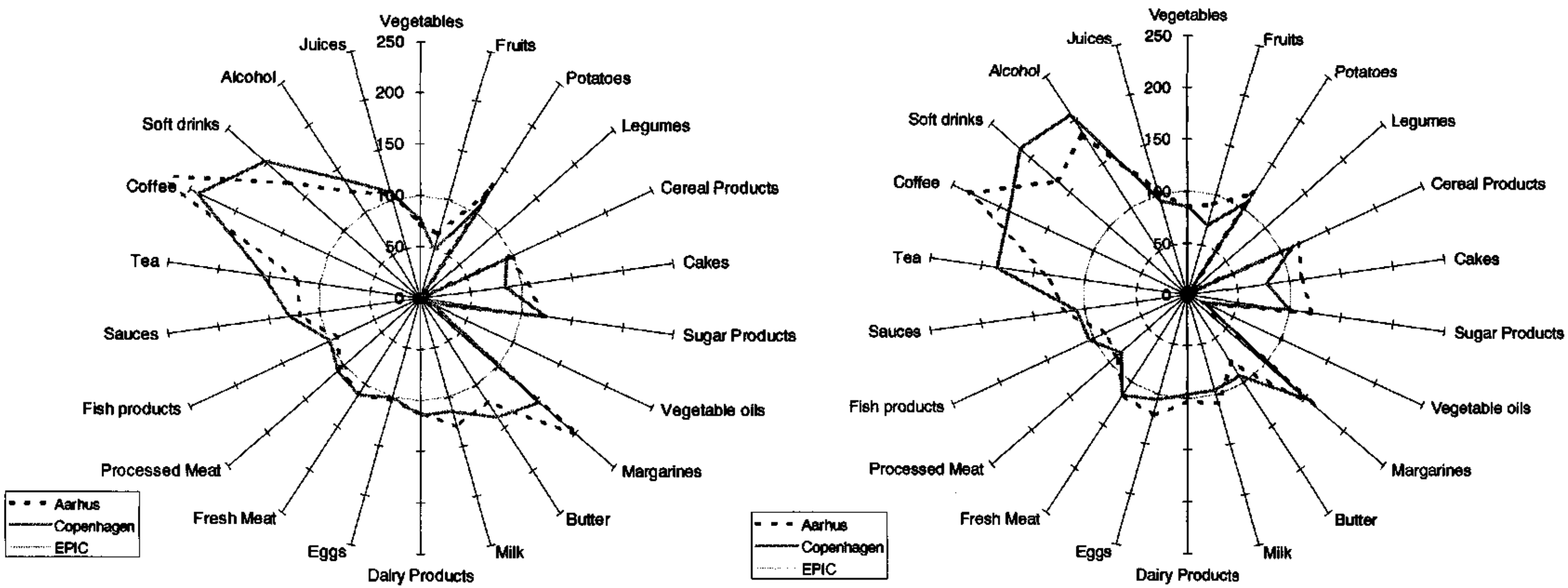
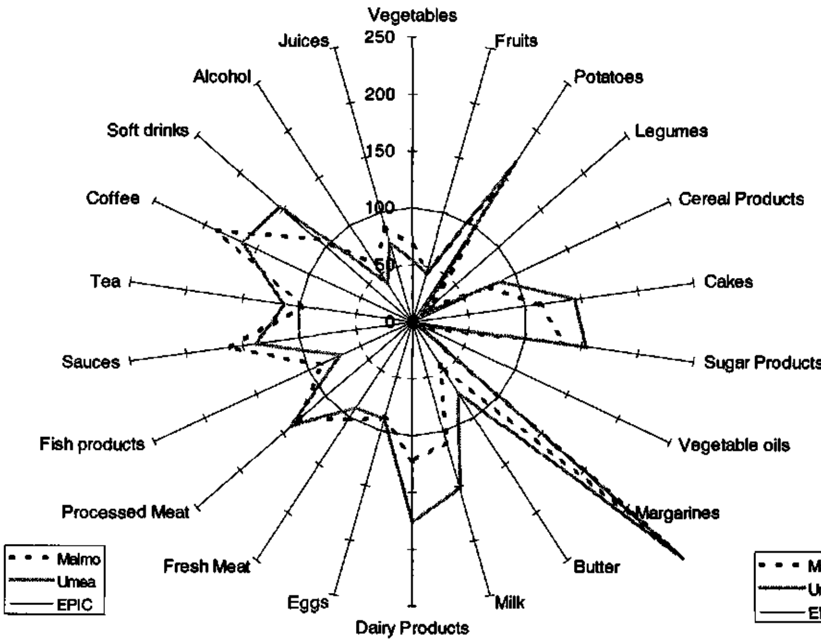


Figure 9. Sweden

(a) : Men



(b) : Women

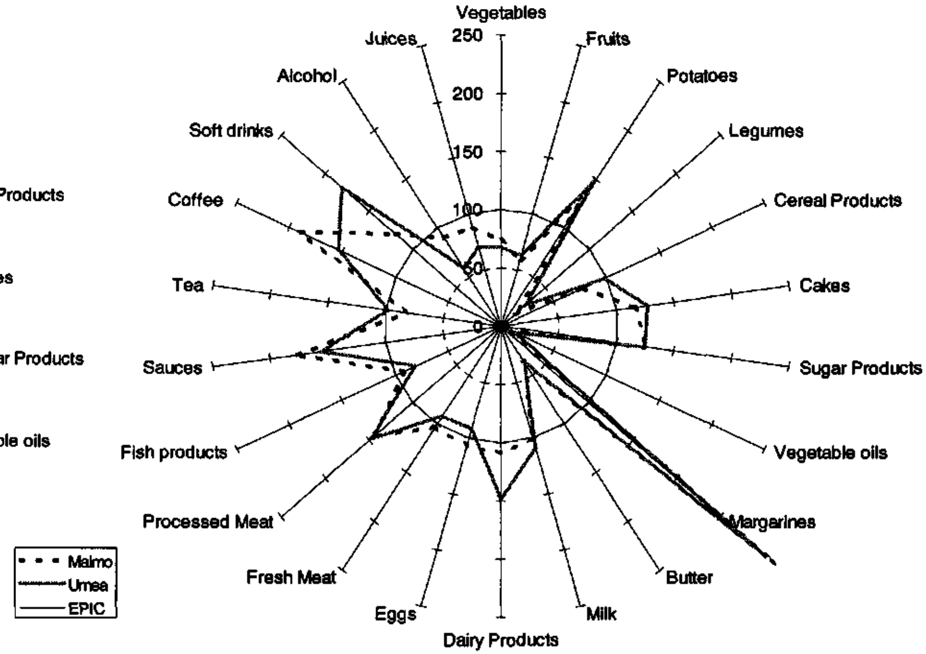
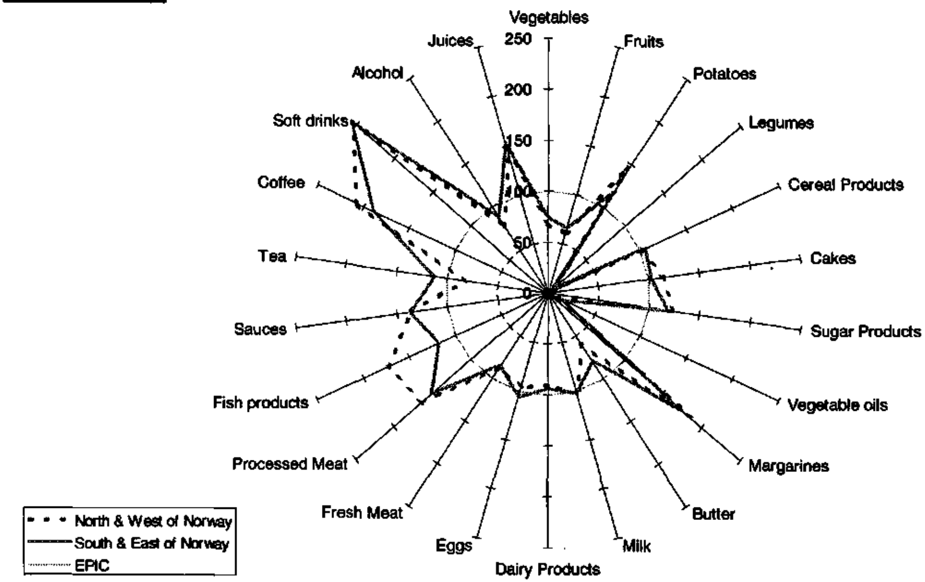


Figure 10. Norway

Women



[a]The reference circle of radius 100% indicated in each figure corresponds to the EPIC means and the spikes indicate the deviation of specific food group means from the reference EPIC means.

The analysis of the EPIC data presented here also showed that although considerable differences in diet were observed between centres, dietary patterns are markedly distinguishable from one country to another. Apart from the atypical UK "health-conscious" population, which should be considered separately, the EPIC country-specific dietary patterns can be classified broadly into three main categories.

1) Overall, Italy and Greece have different dietary patterns but both are characterized by plant foods (except potatoes), with lower consumption of animal and processed foods than the other EPIC countries.
2) France and, to a greater degree, Spain have a more heterogeneous dietary pattern. The average Spanish diet, for example, is characterized by high consumption of both plant foods (legumes, vegetable oils, fruits and vegetables) and animal products such as fresh meat, eggs, processed meat, fish, milk, and alcohol (men), but varies by centre and gender.
3) In contrast, the diet in the Nordic countries, the Netherlands, Germany and the UK general population was relatively high in potatoes and animal, processed and sweetened foods, with proportions varying across countries, indicating additional country- and centre-specific dietary patterns. Other common characteristics of these countries are that consumption of vegetables and fruit is about or below the overall EPIC mean, and that legumes and vegetable oils are not typical of their diets.

The "health-conscious" group in the UK reported an atypical diet different from the UK general and other EPIC populations. The diet is closer to Italy and Greece with respect to legumes, vegetables and fruits, but with lower consumption. In addition, soy products, which are not reported in this analysis, are consumed almost exclusively by this population group (43). However, apart from characteristics associated with vegetarian eating habits, the UK "health-conscious" group also exhibits a high consumption of tea, cakes, soft drinks, margarine and butter, as observed in the UK general population and other central and Nordic countries.

It is also interesting to note that the overall dietary patterns are similar between genders, although compared to different sex-specific EPIC means. However, the magnitude and ranking of relative food consumption across centres varied between genders for certain food groups. For example, we consistently observed a combined geographical/gender trend in alcohol consumption. In Spanish centres and Greece, the deviation from the sex-specific EPIC mean was consistently higher in men than in women for alcohol, and inversely in most of the other central and, particularly, Nordic centres.

Dietary patterns were similar when crude or adjusted means were used in the analysis, although adjustment for age, day of the week and season did affect the magnitude of the deviations from the EPIC mean (data not shown). When total energy intake was added to the model, the overall patterns were very similar in most centres and genders, particularly for women (not shown data). However, greater differences, between the values obtained with the model which included total energy intakes and the one which did not, were observed, particularly in Greece, San Sebastian and, to a lesser extent, Aarhus (women), suggesting a higher degree of under-reporting or other possible factors, discussed in greater detail elsewhere in the supplement (44).

These overall results must, however, be interpreted with caution as, even within food groups, considerable heterogeneity in consumption of products may exist across centres. For example, in our analysis we compared the average volume of coffee drunk across centres, independently of the active compounds contained and method of preparation. Further details on food items (e.g. expresso vs filter coffee), redefinition of food groups and/or composition or bio-marker data would be required in order to study, for example, the consumption of caffeine and its association with diseases. In addition, the non-inclusion of the food group "soups" in this analysis means that consumption of vegetables or legumes (e.g. in Sweden) may have been underestimated, although this is unlikely to have affected the ranking of consumption across centres. Moreover, the data presented in this paper were obtained from individuals using a highly standardized methodology and the mean 24-hour dietary recall intakes should provide, in most centres, relatively reliable estimates at the EPIC population level.

However, the EPIC study cohorts are not designed to be representative of their countries; consequently extrapolation to general populations or to other studies should be made with caution. Furthermore, although other useful dietary data (e.g. Food Balance Sheet or Household Budget Survey data) for between-country comparisons exist (45,46), they have several methodological limitations (47), which means that no direct comparison can be made with individual estimates, particularly *overall* dietary patterns. Further work is needed to compare these different sources of dietary data.

This first overall analysis of EPIC data confirms a number of general trends in dietary patterns observed elsewhere. The Mediterranean centres involved in EPIC *indirectly* show an evolution in their traditional diets towards a more Western diet with different intermediate stages in Spain, France and the northern Italian centres. During the last 30–40 years, substantial changes in diet have been observed in both northern European and Mediterranean countries, reducing the differences traditionally existing between European dietary patterns (48,49); the phenomenon is accelerating among younger generations (50-52). Over this period, Greece (53) and more markedly Italy (54) and Spain (55) have increased their consumption of animal products (i.e. meat, eggs, milk and dairy products) and decreased that of cereals while conserving some of the main features of their traditional Mediterranean diet (i.e. high consumption of fruits, vegetables, legumes, fish, olive oil), with even an increase in fruit and vegetable intakes.

Nordic and central European countries have also changed their different traditional diet to more Western diets. These changes are characterized particularly by a decrease in cereal and potato consumption and an increase in animal and processed foods (56-58). For example, as suggested by the comparison of FAO food supply data from 1960–1990 (58), they have also increased their consumption of meat in varying degrees, but less rapidly, as intake was already higher in the 1960s than in the Mediterranean countries (except France). It has been reported that diets in Nordic countries are healthier than in the past, due particularly to increased consumption of fruit and vegetables, although desirable intakes have not yet been reached (59). These healthy changes were associated with the availability of an increased variety of foodstuffs, greater awareness due to public health programmes started in the early 1950s, and more openness to changes in diet (60). Increased consumption of total vegetable oils, but not particularly of olive oil, was also observed in Nordic countries (58), but this was essentially due to a substantial increase in margarine consumption in recent decades. However, the type of margarine consumed varies by centre, from pure vegetable margarine in Germany to high intakes of mixed-fat margarine (up to 70% of fat as dairy fat), particularly in Sweden and Denmark (41).

Conclusions

This first multi-dimensional analysis of EPIC dietary data, using a common methodology, has confirmed that there are considerable differences in food group consumption and dietary patterns between the study populations. This heterogeneity should help to investigate the relationship between diet and cancer and to formulate new aetiological hypotheses related to overall dietary patterns and diseases. Put in perspective, our data add further indirect support to the observations that dietary patterns in Western societies are changing continuously and lead to the following recommendations for the future:

1) As other authors have stated (61-64), there is a need to identify and better define or redefine the different variants of dietary patterns (54,63-66) which deviate from the original Mediterranean diet as observed more than 40 years ago (67), as well as the different forms of Western diets existing in Europe, as suggested by our data and others (56,68-72).
2) These dietary changes and their determinants should be monitored over time with repeated measurements using a common methodology at the national and international level, with particular attention to younger generations.
3) As Willett (73) has pointed out, a future challenging area of nutritional research is the analysis of dietary patterns and their association with chronic diseases, as the main dietary exposure or as a covariate to control for confounding in univariate analyses considering one specific dietary component.

Acknowledgements

The work described in the paper was carried out with the financial support of the "Europe Against Cancer" Programme of the European Commission (SANCO); Ligue contre le Cancer (France); Société 3M (France); Mutuelle Générale de l'Education Nationale; Institut National de la Santé et de la Recherche Médicale (INSERM); Institut Gustave Roussy; German Cancer Aid; German Cancer Research Center; German Federal Ministry of Education and Research; Danish Cancer Society; Health Research Fund (FIS) of the Spanish Ministry of Health; the Spanish Regional Governments of Andalucia, Asturia, Basque Country, Murcia and Navarra; Cancer Research UK; Medical Research Council, UK; the Stroke Association, UK; British Heart Foundation; Department of Health, UK; Food Standards Agency, UK; the Wellcome Trust, UK; Greek Ministry of Health; Greek Ministry of Education; Italian Association for Research on Cancer; Italian National Research Council; Dutch Ministry of Public Health, Welfare and Sports; Dutch Prevention Funds; LK Research Funds; Dutch ZON (Zorg Onderzoek Nederland); World Cancer Research Fund (WCRF); Swedish Cancer Society; Swedish Scientific Council; Regional Government of Skane, Sweden; Norwegian Cancer Society; the Norwegian

Research Council. Partial support for the publication of this supplement was provided by the Centre de Recherche et d'Information Nutritionnelles (CERIN).

In addition, we wish to thank all study participants for their cooperation and all the interviewers who participated in the field work studies in each EPIC centre.

References

1. Murcott A. You are what you eat: anthropological factors influencing food choice. In: Ritson C, Gofon L, McKenzie J, eds. *The Food Consumer*. New York: John Wiley & Sons Ltd, 1986.
2. Bingham SA. The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutr. Abstracts Rev. (Series A)* 1987; 57, 705–42.
3. Archer VE. Cooking methods, carcinogens, and diet-cancer studies. *Nutr. Cancer* 1988; 11: 75–9.
4. Prentice RL. Measurement error and results from analytic epidemiology: dietary fat and breast cancer. *J. Natl. Cancer Inst.* 1996; 88: 1738–47.
5. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for random within-person measurement error. *Am. J. Epidemiol.* 1992; 136: 1400–13.
6. Willett WC. *Nutritional Epidemiology*, 2nd ed. NewYork: Oxford University Press, 1998.
7. Kipnis V, Midthune D, Freedman LS, Bingham S, Schatzkin A, Subar A, Carroll RJ. Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am. J. Epidemiol.* 2001; 153: 394–403.
8. Ursin G, Ziegler RG, Subar AF, Graubard BI, Haile RW, Hoover R. Dietary patterns associated with a low-fat diet in the National Health Examination Follow-up Study: identification of potential confounders for epidemiologic analysis. *Am. J. Epidemiol.* 1993; 137: 916–27.
9. Weinberg CR. Toward a clearer definition of confounding. Letter to the Editor. *Am. J. Epidemiol.* 1994; 140: 583–584.
10. Wu AH, Yang D, Pike MC. A meta-analysis of soyfoods and risk of stomach cancer: the problem of potential confounders. *Cancer Epidemiol. Biomarkers Prev.* 2000; 9: 1051–8.
11. Klatsky AL. Diet, alcohol, and health: a story of connections, confounders, and cofactors. *Am. J. Clin. Nutr.* 2001; 74: 279–80.
12. Peto J. Cancer epidemiology in the last century and the next decade. *Nature* 2001; 411: 390–5.
13. WCRF/AICR: *Food, Nutrition and the Prevention of Cancer: A Global Perspective*. Washington DC: World Cancer Research Fund/American Institute for Cancer Research, 1997.
14. Tarasuk V, Beaton GH. The nature and individuality of within-subject variation in energy intake. *Am. J. Clin. Nutr.* 1991; 54: 464–70.
15. Tarasuk V, Beaton GH. Day-to-day variation in energy and nutrient intake: evidence of individuality in eating behaviour? *Appetite* 1992; 18: 43–54

16. Hair JF, Anderson RE, Tatham RL, Black WC. *Multivariate Data Analysis with Readings*. New York: McMillan Publishing Co., 1992.
17. Ciampi A, Ferrari P, Slimani N, Overvad K, Ocké M, Wirfält E, Gullberg B, Schulze M. et al. Discovering dietary patterns in nutritional epidemiology: towards a strategy integrating several techniques of multivariate analysis (in preparation)
18. Huijbregts PPCW, Feskens EJM, Kromhout D. Dietary patterns and cardiovascular risk factors in elderly men: The Zutphen elderly study. *Int. J. Epidemiol.* 1995; 24: 313–20.
19. Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma KN. Eating patterns and risk of colon cancer. *Am. J. Epidemiol.* 1998; 148: 4–16.
20. Hu FB, Rimm EB, Stampfer MJ, Ascherio A, Spiegelman D, Willett WC. Prospective study of major dietary patterns and risk of coronary heart disease in men. *Am. J. Clin. Nutr.* 2000; 72: 912–21.
21. Palli D, Russo A, Decarli A. Dietary patterns, nutrient intake and gastric cancer in a high-risk area of Italy. *Cancer Causes Control* 2001; 12: 163–72.
22. Terry P, Suzuki R, Hu FB, Wolk A. A prospective study of major dietary patterns and the risk of breast cancer. *Cancer Epidemiol. Biomarkers Prev.* 2001; 10: 1281–5.
23. Martinez ME, Marshall JR, Sechrest L. Invited commentary: factor analysis and the search for objectivity. *Am. J. Epidemiol.* 1998; 148: 17–9.
24. Wirfält E, Mattisson I, Gullberg B, Berglund G. Food patterns defined by cluster analysis and their utility as dietary exposure variables: a report from the Malmö Diet and Cancer Study. *Public Health Nutr.* 2000; 3: 159–73.
25. Jacques PF, Tucker KL. Are dietary patterns useful for understanding the role of diet in chronic disease? *Am. J. Clin. Nutr.* 2001; 73: 61–7.
26. Riboli E, Hunt K, Saracci R, Hémon B on behalf of the EPIC group. The EPIC study: rationale and design. *Public Health Nutr.* 2002; (this supplement)
27. Tjønneland A, Overvad K, Haraldsdóttir J, Bang S, Ewertz M, Jensen OM. Validation of a semiquantitative food frequency questionnaire developed in Denmark. *Int. J. Epidemiol.* 1991; 20: 906–12.
28. Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, Sneyd MJ, Key TJA, Roe L, Day NE. Comparison of dietary assessment methods in nutritional epidemiology: weighted records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. *Br. J. Nutr.* 1994; 72: 619–43.
29. Riboli E, Elmståhl S, Saracci R, Gullberg B, Lindgärde F. The Malmö Food Study: validity of two dietary assessment methods for measuring nutrient intakes. *Int. J. Epidemiol.* 1997; 26 (Suppl 1): S161–71.
30. Margetts BM, Pietinen P, Riboli E, eds. EPIC: European Prospective Investigation into Cancer and Nutrition: validity Studies on Dietary Assessment Methods. *Int. J. Epidemiol.* 1997; 26 (Suppl 1): S1–189.
31. Slimani N, Kaaks K, Ferrari P, Casagrande C on behalf on the EPIC group. EPIC calibration study: rationale, design and population characteristics. *Public Health Nutr.* 2002; (this supplement)

32. Slimani N, Deharveng G, Charrondière RU, van Kappel AL, Ocké MC, Welch A, Lagiou A, van Liere M, Agudo A, Pala V, Brandstetter B, Andrén C, Stripp C, van Staveren WA, Riboli E. Structure of the standardized computerized 24-hour diet recall interview used as reference method in the 22 centres participating in the EPIC project. *Comput. Methods Programs Biomed.* 1999; 53: 251–66.
33. Slimani N, Ferrari P, Ocké M, Welch A, Boeing H, van Liere M, Pala V, Amiano P, Lagiou A, Mattisson I, Stripp C, Engeset D, Charrondière R, Buzzard M, van Staveren W, Riboli E. Standardization of the 24-hour diet recall calibration method used in the European Prospective Investigation into Cancer and Nutrition (EPIC): general concepts and preliminary results. *Eur. J. Clin. Nutr.* 2000; 54: 900–17.
34. Brustad M, Skeie G, Braaten T, Slimani N, Lund E. Comparison of telephone versus face-to-face interview in the assessment of dietary intake by the 24-hour recall EPIC-SOFT programme – the Norwegian calibration study. (EJCN, in press)
35. Sieri S, Agudo A, Kesse E, Klipstein-Grobush K, San-José B, Welch A, Krogh V, on behalf of the EPIC group. Patterns of alcohol consumption in 10 European countries participating in the EPIC project. *Public Health Nutr.* 2002; (this supplement)
36. Riboli E. Nutrition and cancer: background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann. Oncol.* 1992; 3:783–791.
37. Agudo A, Slimani N, Ocké MC, Naska A on behalf of the EPIC group. Consumption of vegetables, fruit and other plantfoods in the EPIC cohorts from 10 European countries *Public Health Nutr.* 2002; (this supplement)
38. Hjartåker A, Lagiou A, Slimani N, Lund E, on behalf of the EPIC group. Data from 35,955 24-hour dietary recalls in 10 European countries. *Public Health Nutr.* 2002; (this supplement)
39. Wirfält E, McTaggart A, Pala V, Gullberg B, on behalf of the EPIC group. Food sources of carbohydrates in an European cohort of adults. *Public Health Nutr.* 2002; (this supplement)
40. Linseisen J, Kesse E, Slimani N, on behalf of the EPIC group. Meat consumption in the EPIC cohorts: results from the 24-hour dietary recalls. *Public Health Nutr.* 2002; (this supplement)
41. Linseisen J, Bergström E, Gafá L, González CA, Thiébaut A, Trichopoulou A, Tumino R, on behalf of the EPIC group. Consumption of added fats and oils in EPIC centres across 10 European countries as assessed by 24-hour recalls. *Public Health Nutr.* 2002; (this supplement)
42. Welch A, Lund E, Dorronsoro M, Amiano P, on behalf of the EPIC group. Fish consumption in 10 European countries participating in the EPIC study (European Prospective Investigation into Cancer and Nutrition). *Public Health Nutr.* 2002; (this supplement)
43. Keinan-Boker L, Peeters PHM, Mulligan AA, Navarro C, Slimani N, on the behalf of the EPIC group. Soy product consumption in 10 European countries – The EPIC study. *Public Health Nutr.* 2002; (this supplement)

44. Ferrari P, Slimani N, on behalf of the EPIC group. Under-reporting observed in the EPIC 24-hour diet recalls. *Public Health Nutr.* 2002; (this supplement)
45. FAO. *Food Balance Sheets*. Rome: Food and Agriculture Organization, 2001. (http://apps.fao.org/lim500wrap.pl?FoodBalanceSheet&Domain = FoodBalanceSheet&Language=english)
46. University of Athens. *DafneSoft (DAta Food NEtworking)*. Athens: School of Medicine, Public Health and Nutritional Epidemiology Unit, 2001. (DafneSoft programme, www.nut.uoa.gr)
47. Serra-Majem L. Food availability and consumption at national, household and individual levels: implications for food-based dietary guidelines development. *Public Health Nutr.* 2001; 4: 673–6.
48. Hill MJ. Changing pattern of diet in Europe. *Eur. J. Cancer Prev.* 1997; 6 (Suppl 1): S11–3.
49. Rumm-Kreuter D. Comparison of the eating and cooking habits of Northern Europe and the Mediterranean countries in the past, present and future. *Int. J. Vitam. Nutr. Res.* 2001; 71: 141–8.
50. Amorim Cruz JA. Dietary habits and nutritional status in adolescents over Europe – Southern Europe. *Eur. J. Clin. Nutr.* 2000, 54 (Suppl 1): S29–35.
51. Rolland-Cachera MF, Bellisle F, Deheeger M. Nutritional status and food intake in adolescents living in Western Europe. *Eur. J. Clin. Nutr.* 2000; 54 (Suppl 1): S41–6.
52. Samuelson G. Dietary habits and nutritional status in adolescents over Europe. An overview of current studies in the Nordic countries. *Eur. J. Clin. Nutr.* 2000; 54 (Suppl 1): S21–8.
53. Simopoulos AP, Sidossis LS. What is so special about the traditional diet of Greece. *World Rev. Nutr. Diet.* 2000; 87: 24–42.
54. Ferro-Luzzi A, Branca F. Mediterranean diet, Italian-style: prototype of a healthy diet. *Am. J. Clin. Nutr.* 1995; 61 (Suppl): S1338–45.
55. Serra-Majem L, Santana-Armas JF, Salmona E. Dietary habits and nutritional status in Spain. *World Rev. Nutr. Diet.* 2000; 87: 127–59.
56. Trygg K. Food patterns in Nordic countries. *Ann. Nutr. Metab.* 1991; 35 (Suppl 1): 3–11.
57. Oshaug A. Historical perspective and problem analysis of nutrition security in Norway. Is nutrition security in Norway? A situation analysis. *Scand. J. Nutr.* 1994; 28 (Suppl): 6–29.
58. Helsing E. Trends in fat consumption in Europe and their influence on the Mediterranean diet. *Eur. J. Clin. Nutr.* 1993; 47 (Suppl 1): S4–S12.
59. Johansson L, Drevon CA, Bjørneboe G-EA. The Norwegian diet during the last hundred years in relation to coronary heart disease. *Eur. J. Clin. Nutr.* 1996; 50: 277–83.
60. De Graaf C, Van der Gaag M, Kafatos A, Lennernas M, Kearney JM. Stages of dietary changes among nationally representative samples of adults in the European union. *Eur. J. Clin. Nutr.* 1997; 51 (Suppl 2): S47–56
61. Ferro-Luzzi A, Sette S. The Mediterranean diet: an attempt to define its present and past composition. *Eur. J. Clin. Nutr.* 1989; 43: 13–29

62. Helsing E. Traditional diets and disease patterns of the Mediterranean, circa 1960. *Am. J. Clin. Nutr.* 1995; 61 (Suppl): S1329S–37.
63. Simopoulos AP. The Mediterranean diets: what is so special about the diet of Greece? The scientific evidence. *J. Nutr.* 2001; 131: S3065–73.
64. Noah A, Truswell AS. There are many Mediterranean diets. *Asia Pacific J. Clin. Nutr.* 2001; 10: 2–9.
65. Kafatos A, Verhagen H, Moschandreas J, Apostolaki I, Van Westerop JJ. Mediterranean diet of Crete: foods and nutrient content. *J. Am. Diet. Assoc.* 2000; 100: 1487–93.
66. Moreiras O, Cuadrado C. Mediterranean diet and lifestyle: Special aspects of Spain. *Int. J. Vitam. Nutr. Res.* 2001; 71: 154–8.
67. Keys A, ed. *Seven Countries: A Multivariate Analysis of Death and Coronary Heart Disease*. Cambridge, MA: Harvard University Press, 1980.
68. Hulsholf KFA, Wedel M, Löwik MRH, Kok FJ, Kistemaker C, Hermus RJJ, ten Hoor F, Ockhuizen T. Clustering of dietary variables and other lifestyle factors (Dutch Nutritional Surveillance System). *J. Epidemiol. Community Health.* 1992; 46: 417–24.
69. SENECA Investigators. Food patterns of elderly Europeans. *Eur. J. Clin. Nutr.* 1996; 50 (Suppl 2): S86–100.
70. Winkler G, Brasche S, Döring A, Heinrich J. Dietary intake of middle-aged men from an East and West German city after the German reunification: do differences still exist? *Eur. J. Clin. Nutr.* 1998, 52: 98–103.
71. Vlatier JL, Verger P. Recent national French food and nutrient intake data. *Br. J. Nutr.* 1999; 81 (Suppl 2): S57–9.
72. Pryer JA, Nichols R, Elliott P, Thakrar B, Brunner E, Marmot M. Dietary patterns among a national random sample of British adults. *J. Epidemiol. Community Health.* 2001; 55: 29–37.
73. Willett WC. Nutritional epidemiology issues in chronic disease at the turn of the century. *Epidemiol. Rev.* 2000; 22: 82–6.

Annex 1. Sex-specific EPIC means and percentage deviation of the centre-adjusted[a] food group means from the sex-specific "EPIC means"
([a] Adjusted for age, week days and seasons)

	EPIC mean (g)		ITALY									GREECE		SPAIN							
			Varese		*Turin*		*Florence*		*Ragusa*		*Naples*			*Asturias*		*San Sebastian*		*Navarra*		*Murcia*	
	M	F	M	F	M	F	M	F	M	F	F	M	F	M	F	M	F	M	F	M	F
n =			*328*	*794*	*677*	*392*	*271*	*785*	*168*	*138*	*403*	*1312*	*1374*	*386*	*324*	*490*	*244*	*444*	*271*	*243*	*304*
Vegetables	183.2	174.2	101.6	92.7	134.5	132.8	112.1	111.2	92.9	80.3	98.5	146.9	121.3	78.4	58.7	126.2	121.3	132.5	109.6	142.6	143.6
Fruits	264.0	250.9	124.5	129.6	153.4	139.9	140.0	126.1	148.2	147.0	113.8	88.7	86.9	126.4	129.0	140.2	145.3	116.8	126.7	165.0	148.3
Potatoes	87.0	60.4	54.2	54.9	55.2	52.7	57.0	50.5	67.6	45.3	48.4	49.8	54.4	111.8	117.0	101.3	102.2	78.0	85.6	95.2	113.5
Legumes	19.9	10.5	30.4	68.5	46.4	31.4	94.6	115.4	61.5	135.1	122.0	168.1	195.6	323.7	336.9	280.5	377.4	248.4	242.8	184.8	187.3
Cereals & cereal products	270.8	175.7	160.0	141.4	145.0	129.7	159.9	144.0	164.3	153.1	147.5	107.1	105.8	74.9	71.6	69.9	69.0	79.8	71.6	83.7	68.9
Cakes	54.2	51.3	118.2	105.2	94.1	89.6	117.7	100.0	94.1	110.0	101.6	43.8	58.5	72.0	79.5	66.7	76.5	61.5	87.3	102.3	108.7
Sugar/ confectionery	31.0	26.3	107.3	81.1	83.2	79.1	63.3	57.9	78.2	59.1	76.0	52.5	59.6	92.2	87.7	85.2	99.6	73.9	80.5	71.7	67.7
Vegetable oils	18.6	11.4	87.1	103.4	130.1	169.5	150.3	178.1	133.3	174.4	161.0	237.4	284.7	150.1	149.5	221.9	265.3	249.4	283.1	203.3	275.3
Margarines	14.3	7.9	4.4	4.1	4.2	6.7	2.7	3.8	2.1	7.1	6.4	18.1	27.7	20.4	33.7	10.0	31.0	8.1	22.5	13.4	27.0
Butter	4.5	4.1	71.1	64.6	32.2	36.8	60.8	50.5	11.5	13.8	18.2	0.4	6.4	1.6	19.6	7.3	12.2	0.0**	16.1	18.8	18.9
Milk	190.6	178.3	90.2	90.7	55.2	65.1	70.5	75.5	47.8	49.6	66.7	52.3	62.6	174.4	209.8	128.9	170.4	114.9	184.2	103.3	153.8
Dairy products	101.6	115.1	75.8	64.7	82.9	66.4	76.7	65.3	72.6	57.4	66.9	88.1	66.9	85.9	89.2	84.1	79.5	62.5	57.5	56.1	50.4
Eggs	19.6	14.6	34.6	73.4	67.2	81.1	70.1	67.8	85.5	79.5	79.4	65.4	67.2	185.3	169.8	246.9	208.8	198.3	159.6	115.8	127.5
Fresh meat	66.3	40.4	123.1	96.1	85.4	115.3	99.4	116.7	101.8	111.9	87.8	85.8	73.6	136.4	128.5	208.7	159.5	131.3	108.9	78.6	76.2
Processed meat	49.4	29.5	85.9	90.4	68.5	68.4	61.8	65.2	47.7	60.5	59.4	22.7	19.2	125.4	105.5	106.1	106.3	114.1	121.6	95.1	86.1
Fish & shellfish	47.4	36.1	49.7	55.4	72.4	57.4	74.3	54.5	74.6	90.9	80.8	98.2	88.3	218.9	191.1	255.2	205.8	176.1	182.8	146.2	142.9
Sauces	36.1	26.3	157.5	134.0	82.9	77.1	109.8	74.0	181.6	154.1	107.0	50.0	41.0	35.5	39.5	57.6	57.3	42.8	42.1	35.8	32.0
Tea	143.4	168.1	31.3	51.2	29.7	35.8	25.6	26.6	11.5	14.4	10.0	20.1	14.8	2.8	2.8	3.0	7.6	0.0**	1.7	3.1	2.6
Coffee	363.4	315.5	39.1	38.4	31.1	31.3	29.0	33.7	27.0	26.3	40.3	19.5	12.2	35.0	54.5	25.0	43.3	34.4	54.3	22.6	21.8
Soft drinks	92.5	64.5	59.3	40.0	42.5	50.7	34.1	27.7	64.9	60.2	51.2	68.9	57.6	84.2	63.0	37.2	43.0	55.9	39.9	94.4	44.0
Alcohol	27.5	9.7	123.7	80.6	114.0	109.3	81.6	79.3	61.2	63.0	81.4	76.3	41.0	134.8	72.2	157.6	86.6	141.3	46.8	108.3	83.0
Juices	60.8	61.3	52.9	33.6	46.1	49.7	43.5	43.1	26.4	13.5	41.0	62.0	73.7	70.4	65.0	35.5	35.3	54.5	46.8	77.5	59.4

** As the adjusted estimates were negative and the food quantities close to zero, the values reported were set to zero

Annex 1. Sex-specific EPIC means and percentage deviation of the centre-adjusted[a] food group means from the sex-specific "EPIC means" ([a] Adjusted for age, week days and seasons.) (continued)

	SPAIN		FRANCE				GERMANY				THE NETHERLANDS		
	Granada		*North-east*	*North-west*	*South*	*South coast*	*Potsdam*		*Heidelberg*		*Bilthoven*		*Utrecht*
	M	F	F	F	F	F	M	F	M	F	M	F	F
n =	*214*	*300*	*854*	*621*	*1396*	*612*	*1235*	*1063*	*1033*	*1087*	*1024*	*1086*	*1874*
Vegetables	121.5	127.3	123.3	112.9	125.5	131.9	77.6	94.4	88.7	93.6	71.3	72.9	74.1
Fruits	137.1	135.0	93.3	91.7	98.4	89.2	86.1	96.3	60.2	79.0	59.1	64.4	81.7
Potatoes	88.1	88.7	88.9	112.0	81.8	59.6	123.8	136.0	90.5	105.7	141.3	139.8	139.9
Legumes	184.0	176.6	61.1	50.7	66.7	80.5	18.9	26.8	21.5	30.5	29.9	38.0	42.7
Cereals & cereal products	80.5	71.6	88.1	85.7	95.4	86.1	72.1	82.1	87.0	96.6	86.2	94.8	88.6
Cakes	65.8	60.1	111.5	99.9	95.7	85.4	112.4	111.0	117.5	111.1	91.7	108.6	110.2
Sugar/ confectionery	73.2	66.4	122.7	115.6	121.7	115.0	96.8	98.1	103.9	107.5	161.7	140.3	127.4
Vegetable oils	200.8	218.4	45.0	27.5	53.0	71.7	10.5	18.4	14.9	25.4	11.6	17.7	13.2
Margarines	20.8	47.7	50.7	50.1	30.1	40.7	246.2	229.5	52.8	55.5	225.9	227.4	212.7
Butter	40.5	23.5	242.6	289.5	207.2	166.9	475.3	316.4	372.7	323.0	111.4	113.9	112.8
Milk	158.4	168.9	57.9	55.1	62.3	64.0	50.0	55.5	52.7	60.7	94.8	92.3	129.9
Dairy products	71.6	65.7	158.8	155.5	146.8	146.8	124.1	118.1	117.9	108.2	179.2	139.9	168.5
Eggs	131.2	111.0	98.1	98.5	97.5	93.5	69.5	80.9	62.1	67.8	80.2	94.5	98.3
Fresh meat	79.0	67.6	139.3	125.7	126.9	120.1	79.1	78.7	92.3	79.6	103.3	106.5	103.8
Processed meat	107.5	97.7	101.0	108.9	98.2	94.9	180.7	139.2	163.0	145.6	147.4	130.0	124.5
Fish & shellfish	182.3	150.3	102.9	139.9	98.2	106.2	49.4	51.2	32.9	41.3	37.9	38.3	39.1
Sauces	37.5	40.7	103.9	97.8	113.7	109.4	81.3	89.2	115.1	122.6	86.5	79.2	88.6
Tea	3.8	4.3	112.9	143.8	132.6	99.6	83.9	59.6	108.5	101.1	161.3	174.4	251.8
Coffee	32.3	25.6	68.7	87.1	58.9	61.4	154.5	156.3	150.7	157.9	183.2	173.6	161.0
Soft drinks	99.0	93.1	39.6	34.2	40.2	38.2	79.0	42.1	128.8	90.8	163.0	143.4	115.3
Alcohol	98.7	51.2	131.9	131.1	110.9	123.9	91.0	96.6	110.0	144.5	91.7	114.0	123.9
Juices	84.9	61.8	99.3	101.4	92.2	88.3	245.8	255.7	309.0	281.3	112.3	164.3	171.3

[aa] As the adjusted estimates were negative and the food quantities close to zero, the values reported were set to zero

Annex 1 Sex-specific EPIC means and percentage deviation of the centre-adjusted[a] food group means from the sex-specific "EPIC means" ([a] Adjusted for age, week days and seasons) (continued)

	UMITED KINGDOM				DEMNARK				SWEDEN				NORWAY	
	General population		*"Health conscious"*		*Aarhus*		*Copenhagen*		*Malmo*		*Umea*		*North & West*	*South & East*
	M	F	M	F	M	F	M	F	M	F	M	F	F	F
n =	*404*	*571*	*114*	*197*	*567*	*510*	*1356*	*1485*	*1421*	*1711*	*1344*	*1574*	*662*	*1136*
Vegetables	85.8	93.3	116.6	126.3	72.8	84.9	76.8	84.8	67.2	73.7	54.1	68.1	68.0	74.9
Fruits	54.4	66.5	97.5	101.6	64.4	90.8	48.9	69.4	45.3	58.8	43.9	63.1	61.5	67.0
Potatoes	125.1	119.9	105.3	124.6	131.1	117.7	113.3	108.0	144.4	137.7	166.9	150.4	145.4	119.6
Legumes	41.9	72.9	135.9	184.7	0.0**	0.0**	0.0**	0.8	20.3	16.3	9.3	27.2	4.8	7.5
Cereals & cereal products	79.2	84.9	104.0	109.5	96.8	118.5	92.0	112.2	74.0	77.8	83.7	97.3	102.8	105.4
Cakes	155.4	119.0	140.0	123.1	105.9	110.4	84.0	77.1	112.5	116.0	144.5	127.3	114.2	102.6
Sugar/ confectionery	139.4	109.7	84.2	114.9	121.1	120.7	123.5	100.5	135.1	125.0	153.7	124.7	123.1	118.4
Vegetable oils	24.3	28.1	40.5	52.3	9.1	16.6	11.1	18.0	8.9	10.1	5.5	13.3	10.5	16.2
Margarines	151.4	137.6	152.5	163.9	200.1	161.0	153.3	152.3	298.4	310.5	315.4	301.5	185.1	173.7
Butter	179.7	124.8	140.8	139.0	119.6	77.3	137.5	93.1	44.4	36.6	74.3	36.1	60.6	79.4
Milk	145.0	131.8	55.7	77.5	129.7	107.7	114.4	95.4	109.9	100.6	152.0	108.5	101.5	101.8
Dairy products	118.1	108.8	80.2	79.1	111.7	102.6	114.4	95.8	122.3	107.8	175.8	148.7	90.4	94.2
Eggs	66.3	74.5	41.3	46.1	103.0	120.5	100.5	105.3	88.2	106.4	88.4	91.2	95.4	106.5
Fresh meat	65.2	67.8	16.1	11.1	110.9	114.4	111.9	115.4	101.3	103.1	90.5	92.6	88.1	84.9
Processed meat	77.9	77.4	14.0	18.3	105.0	90.0	107.5	85.0	130.3	144.7	139.5	147.2	163.9	151.1
Fish & shellfish	71.5	78.7	14.4	46.9	91.0	88.7	98.5	104.7	87.3	91.5	69.2	80.4	173.3	118.7
Sauces	132.7	139.5	145.3	127.6	119.2	117.9	129.8	106.9	160.2	176.6	139.0	156.8	134.9	136.6
Tea	568.3	415.9	358.1	337.7	122.1	138.7	156.2	185.5	97.6	82.0	113.1	97.7	82.7	112.3
Coffee	128.7	133.1	110.1	98.4	283.1	231.9	241.6	189.4	188.6	190.0	164.6	154.1	207.5	188.9
Soft drinks	140.8	226.9	108.5	168.1	171.8	166.3	202.6	213.2	111.8	120.7	153.4	181.0	254.5	255.1
Alcohol	81.7	127.8	65.7	93.9	124.2	183.9	137.3	205.6	60.9	90.4	40.1	59.9	78.9	88.4
Juices	121.4	111.1	187.3	158.3	105.9	103.9	107.6	93.6	85.2	87.9	71.8	70.1	144.8	153.4

** As the adjusted estimates were negative and the food quantities close to zero, the values reported were set to zero.

Chapter 8

General discussion

Large prospective multi-cohort studies such as EPIC are a favourable epidemiological setting to investigate the relation between dietary intakes and diseases. The large cohort size and the heterogeneity of study populations, offer the advantage of investigating the variations in both dietary exposure and disease incidence. However, the heterogeneity of dietary habits and cultural background represent a major methodological challenge for the pooling of dietary data. The calibration approach, which shares similarities with the internal and external standards commonly used in laboratories for correcting analytical measurements, has recently been adapted to nutritional epidemiology to correct for dietary measurement errors. EPIC has been the first large multi-centre cohort study to adopt the calibration approach (1) and face the problem of its implementation (2, 3). Indeed, the calibration concept relies on statistical considerations and assumptions stated in chapter 1, which have major implications for its practical, logistical and methodological implementation. Briefly, in order to be used for collecting reference calibration measurements the selected dietary method should satisfy statistical requirements which, in practical terms, can be translated as follows: 1) The reference measurements should be highly standardised across participating centres in order to provide correct centre mean dietary estimates, 2) the reference measurements should be obtained in a representative sub-sample of the study population with high participation rates 3) a third statistical requirement for within-cohort calibration is the independence of random errors between the reference dietary calibration and individual questionnaire measurements.

The main objective of the present thesis was to translate these statistical requirements of the calibration into an epidemiological application. The third statistical requirement regarding the independence of random errors between the two dietary measurements was, however, not addressed in this thesis.

Although weighed or estimated food records were considered as possible alternatives, a single 24-hour diet recall interview was chosen as the best reference calibration method for EPIC. In the absence of an already existing standardized 24-hour diet recall for use in a European context, an *ad hoc* method was developed. Based on the rationale developed in chapters 2 and 4, it was considered that a computerized face-to-face 24-hour diet recall was the only suitable approach capable of satisfying the calibration statistical requirements summarized in chapter 1. This includes, particularly, the standardization of both the 24-hour diet recalls among the 90 interviewers involved in the 23 EPIC centres and the food composition tables to be used in large multi-centre studies (chapters 2, 3 and 4).

The design and the populations of the calibration sub-studies set up in each EPIC centre, as well as the representativeness of the calibration sub-sample compared to the whole cohort were described in chapter 5. The collected 24-hour diet recalls were then evaluated in a validation study detailed in chapter 6 before being used for describing dietary patterns in EPIC, beside the calibration exercise.

This marks the first time that a calibration sub-study using the same dietary method has been carried out in a large multi-centre study. Consequently, although much has been published over the last 20-30 years on dietary methodology (4-13), little has been reported on the use and standardization of a single dietary method in an international calibration study context. This discussion will therefore focus on an evaluation of the calibration study in EPIC and its degree of success in meeting the first and second calibration statistical requirements described previously (chapter 1), together with improvements needed in the reference dietary methodology and study design for optimizing the calibration sub-studies in the future. As the interest of an existing standardized dietary methodology for between- population comparisons goes beyond the pure methodological calibration exercise conducted in EPIC, the discussion will be extended to the use of the EPIC-SOFT computer programme in other study contexts.

Standardisation and validity of the 24-hour diet recall measurements

So far, the degree of standardization and the (relative) validity of the EPIC-SOFT 24-hour diet recalls, used as reference for between-cohort calibration, have been evaluated in three different ways: 1) in order to estimate the level of standardisation across interviewers from the same country, an analysis of variance was performed using mean total energy intake per interviewer as the dependent variable (Chapter 2); 2) estimation of the degree of under- or over-reporting in the EPIC-SOFT measurements by comparing mean total energy intake and EI:BMR ratio per gender across centres (14); 3) A validation of centre mean 24-hour diet recall nitrogen and energy intakes against mean urinary nitrogen conducted on a convenient sample from a restricted number of EPIC centres (Chapter 6).

The analysis of variance using mean total energy intake per interviewer as dependent variable, has shown that a reasonably good degree of standardisation can be achieved between interviewers, although significant variations were associated with gender- and, in some situations, an interviewer- and an EPIC-SOFT-version effect. No statistical difference between interviewers was observed in any country for men, and five out of eight countries for women, after adjustment for physical activity, the EPIC-SOFT version used and after the exclusion of two interviewers with outlying values in two gender-specific centres. The absolute differences in total energy intake between interviewers were relatively modest in both genders with percentage of interviewers within ± 5% and ± 10% of the country mean energy intake of, respectively, 71% and 98% in men and 74% and 94% in women.

Although it is not possible to correct for all measurement errors, as discussed in Chapters 1 and 2, the aim of the standardisation of a dietary method is to minimize possible biases (systematic over- or under-reporting) and design the method so that any residual distortion will apply with equal magnitude and in the same direction across study populations. This effect of standardisation in the 24-hour diet recall measurements was evaluated by comparing the magnitude of under- and over-reporting across centres (14). Standardized 24-hour diet recalls are also expected to improve the precision of dietary measurements and decrease random errors in the reference method. Statistically, this should reduce the standard error of the lambda calibration coefficient applied to de-attenuate the relative risk estimates. Current on-going research indicates that the calibration factors estimated within EPIC have indeed a very small standard error, but this goes beyond the scope of the present thesis.

In the analysis conducted by Ferrari et al. (14) total energy intake and the ratio of total energy intake (EI) on the predicted basal metabolic rate (BMR) and its confidence limits were used for evaluating the relative validity of the EPIC 24-hour diet recall measurements at the population level as proposed by Goldberg et al. (15) and more recently by Black (16). Indeed, in addition to giving an estimate of measurement errors of the overall diet, total energy and the ratio of total energy intake on BMR are expected to be relatively comparable across centres, if correctly adjusted for possible confounders. In the analysis conducted by Ferrari et al., the mean values of EI and EI:BMR, crude and adjusted for age, BMI, height, day of the week and physical activity, before and after subject exclusions according to the cut-off points of Goldberg et al. and of Black, were therefore compared.

A reference Physical Activity Level (PAL = EI:BMR = 1.55) was used for comparison at the population level, considering the moderate physical activity observed in the EPIC cohorts (17). The confidence limits were estimated according to Goldberg's et al. equation (15) using constant variables on physiological variations in both BMR, PAL and energy intake. These values were recently updated using metabolic and doubly labelled water measurements (16). This calculation also takes into account the number of repeated measurements collected per individual and the size of the study sample. It was therefore also possible to estimate, at the individual level, the "*high over- and under-reporters*" in total energy intake defined as the subjects below or above the calculated PAL lower and upper limits, respectively 0.88 and 2.72 (when only one measurement is obtained per subject). It is important to note that although this empirical method is highly specific it is not sensitive enough, as demonstrated by Black (18). In the absence of doubly labelled water measurements, the Goldberg cut-off points are, however, considered useful for approximating mis-reporting in nutritional studies, particularly if information on physical activity is available for the study subjects, as is the case in EPIC (19).

When compared at the population level, the percentage of high under-reporters (i.e. individual EI:BMR values below 0.88), the total EI and the EI:BMR ratio show a systematic tendency to under-estimate energy intake in all EPIC centres by an order of magnitude greater in women than in men. Whereas the percentage of high over-reporters is marginal in both genders (≤ 3% in most centres), the percentage of high under-reporters ranges from 3% to 15% in men and 5% to 19% in women, when Greece is excluded. Greece and, to a lesser extent, the UK (particularly the health conscious group), as well as Granada (women) have consistently higher under-reporting when it is expressed as a percentage of high under-reporters (20% in men and 33% in women in Greece), EI and EI:BMR ratio.

Overall under-reporting, expressed as EI or EI:BMR, was generally heterogeneously distributed among countries but homogeneously distributed among centres from the same countries, in both genders. When expressed as mean adjusted for age, height, BMI, day of the week, and physical activity, without subject exclusions, total EI and EI:BMR ranged from 2359±878 to 2998±852 Kcal and from 1.34 (1.31-1.36) to 1.73 (1.69-1.78) in men and from 1749±622 to 2052±620 Kcal and from 1.27 (1.23-1.30) to 1.51 (1.46-1.57) in women, respectively. The lowest values were observed for Greece in both genders with EI:BMR at 1.31 (1.28-1.34) in men and 1.16 (1.13-1.18) in women. Compared to crude values, adjustment for possible confounders changed the magnitude but did not meaningfully alter the ranking of centres. The exclusion of high over- and under-reporters, in addition to adjustment, further reduced the heterogeneity of EI and EI:BMR between centres, particularly when the number of outlier values was high, such as for Greece, the UK health conscious group and Granada (women). In men for example, the average centre EI ranged from ~ 2500 Kcal to ~ 2800 Kcal in most centres. After exclusion and adjustment, the range of mean total EI in women (1900 to 2000 Kcal) was more markedly reduced than in men, in most centres. In both genders, we still observed consistently slightly lower values for Greece and UK (particularly the health conscious group) and Granada (women), despite a relatively higher level of exclusion. In contrast, higher values were reported in San Sebastian due to higher physical activity levels as compared to other centres, particularly in men. The highest variation in professional and leisure time activities observed in men explains the gender difference in EI range across centres (17). Although the observed mean EI:BMR (or EI) were, in most cases, below the 1.55 reference value, the difference between centres was relatively small in both genders, particularly after adjustment and exclusion (≤ 10% in both genders for both EI and EI:BMR, except women in Greece and the UK). Overall, these differences represent 4% in men and 7.4% in women.

The same general observations were reported in chapter 6 where the centre mean urinary nitrogen quantities served as the reference measurement for validating mean dietary nitrogen and total energy intakes estimated from 24-hour diet recall and baseline dietary questionnaire (n=1103 men and women). Despite different degrees of mis-reporting observed when the analysis was stratified by centre and gender, the partial Pearson's correlation, adjusted for gender, was high for the two dietary measurements and for nitrogen and total energy intake with values between 0.70 to 0.90. Furthermore, the β coefficient of the regression lines representing the power of dietary nitrogen estimates to predict urinary nitrogen were not statistically significantly different from 1, suggesting a similar order of under-reporting across centres. Surprisingly, despite the lack of standardisation, the dietary questionnaires provide as good a ranking as 24-hour diet recalls, particularly in men. However, in contrast to 24-hour diet recalls, the measurement errors do not only differ in magnitude across centres but also in direction (over- and under-reporting), particularly in women. This may cause problems of misclassification when pooling different individual questionnaire data without correction for different systematic over- or under-reporting.

These overall results suggest that, although the centre mean 24-hour diet recalls contain errors, their direction and magnitude is quite comparable across centres. These measurements could therefore still be used for relative calibration of baseline dietary questionnaires using a reference method without an absolute validity but with a comparable relationship with the true latent variable. This demonstrates the possible benefits of *a priori* standardization and *a posteriori* calibration of dietary measurements across centres. However, centres with outlier dietary intake values (e.g. Greece), identified by Ferrari et al. (14) and confirmed by the validation study (Chapter 6), would need particular attention when calibrating. It is interesting to note that the centres with the highest levels of under-reporting (i.e. Greece, UK health conscious and women in Granada) also reported a relatively lower participation rate, suggesting that under-reporting may be, at least in part, attributable to subject (selection) bias. Furthermore, it has been shown that individuals' characteristics such as BMI, age and special diet were (strongly) associated with high under-reporting and could contain a systematic bias component, known as person- or group- specific bias (20) which should be taken into account when calibrating (14).

Representativeness of the calibration sub-samples

In order to obtain a representative sample of the entire cohort, a high participation rate is needed from the subjects invited to participate in the calibration sub-studies. 24-hour diet recall is usually associated with a high participation rate, as confirmed in this study. Overall about 70% of the study centres reported a participation rate between 75% and 93%. As reported in chapter 5, this response rate was consistently higher when subjects were recruited just after the baseline examination than when invited to come back later. Lower response rates (≤ 60%) were observed in centres where (representative) samples of the general populations (UK general population, Norway) or atypical population groups (UK health-conscious group) were recruited, or because of logistical problems in approaching subjects as suggested by high passive non response (30% in Greece).

In addition to the participation rate, we checked whether calibration participants differ from the rest of the cohort regarding characteristics potentially related to dietary intakes such as anthropometry, physical activity, level of education and smoking. The results suggest that the calibration samples, after taking into account the sampling scheme which was stratified by age, are reasonably representative of their respective cohorts in most EPIC centres. This evidence was further supported by the comparison of centre mean dietary intakes of 16 main food groups obtained from the baseline dietary measurement between the calibration sample and the rest of the cohort. Overall, for 89% of the centre-sex-food group combination, the difference between the two groups was less than ± 10% (69% had a difference within ± 5%). However, 59% of the differences above 10% were attributable to only 4 of the 24 centres considered (UK health conscious, Ragusa, Granada and Umea). The relatively low participation rate and limited representativeness of the UK health conscious sub-cohorts (i.e. vegans, vegetarians, fish eaters, meat eaters) may explain these observations, but further investigations are needed to see how to deal with these data when calibrating.

General recommendations for optimizing calibration study designs

From different aspects of the EPIC calibration developed in this thesis it is possible to make broad recommendations for optimising calibration study design in the future. Some of them concern the 24-hour diet recall method (1 to 4) whereas the others deal with the calibration setting:

1) A computerized 24-hour diet recall programme, compatible with the time and cost constraints of large multi-centre studies, can facilitate the standardisation of dietary measurements across interviewers and respondents within and between cohorts.
2) Ideally the 24-hour diet recall should be performed by surprise when the subject is invited for the baseline examination, as indicated by the successful procedures used in several EPIC centres.
3) Alternative methods should however be considered if the 24-hour diet recall interview cannot be performed immediately after the baseline examination. One possible option is a telephone interview using EPIC-SOFT along the line of the study successfully conducted in Norway (21). Interviewing subjects with a 48-hour time interval (e.g. interviews concerning Saturdays performed on Mondays), as experienced in EPIC, is another possible alternative. A sufficient number of dietary recalls should however be collected with both 24 and 48 hours elapsed time in order to estimate the time interval's effect on the dietary intakes.
4) The analysis of EPIC 24-hour diet recalls confirmed that sampling for calibration purposes should ensure good coverage of the different seasons and particularly days of the week (working days vs Friday and week-ends). However, the EPIC field work revealed logistical and practical difficulties in interviewing subjects during non-working days which led to an limited coverage of Fridays and Saturdays in certain centres. The telephone interviews could also be used to improve coverage of weekdays and seasons.

5) Other aspects of the sampling procedure such as the number of repeated measurements to correct better for within-subject random errors or the use of fully random, instead of stratified, sampling should be considered on an individual basis.
6) In order to improve the evaluation, interpretation and use of the calibration measurements, additional information needs to be collected. A number of co-variates or possible confounders such as age, gender, anthropometry, smoking, physical activity, level of education, special day or special diet of the subject the day of the interview, should be systematically collected in order to proceed with appropriate stratifications and adjustments when analyzing the data.
7) Ideally, quantitative biological markers such as urinary nitrogen and/or doubly labelled water should be collected both as an additional independent reference calibration measurement and for validating the reference dietary calibration measurements. Whether such measurements should be collected from the whole calibration sample, a sub-sample or from another comparable independent group should be considered case by case. Ideally, these additional biological measurements should be done without altering the response rate.
8) Finally, the field work should be evaluated regularly and intermediate statistical analyses performed to estimate the degree of standardisation obtained across interviewers and the overall quality of the data collected so far. This practice can enable the investigators to identify possible systematic problems related to interviewers or respondents and help to find the appropriate solutions. However, the interpretation of results should be careful before proceeding to any change in the interview procedures or logistics. For example, a systematic difference in mean energy intake from one interviewer compared to others does not necessarily mean that the data are incorrect. A possible confounding effect, particularly in studies involving large heterogeneous populations, can explain the difference, as discussed in the example of one Spanish centre in Chapter 2.

Future perspectives of using the EPIC-SOFT programme in other study designs

The relatively recent increasing interest in multi-centre studies on lifestyle and chronic diseases, particularly in Europe, re-inforces the need to develop procedures and tools to collect dietary measurements in a standardized way across countries. In contrast to Food Frequency Questionnaires (or dietary histories), dietary records and 24-hour diet recalls have the advantage of being easier to standardise across different study populations (6, 7, 22). Their open-ended structure can be adapted to different dietary habits and level of details of foods consumed and leave more *a posteriori* flexibility to reclassify individual food items according to different (etiological) classification systems in studies investigating diet-disease associations (22, 23).

As detailed in Chapters 2 and 3, the EPIC-SOFT programme has been conceptually developed with an intrinsically flexible structure which can be easily changed and adapted to different end-users' needs and developed in further national versions in addition to those already available for the 10 EPIC countries (i.e. France, Italy, Spain, Germany, The Netherlands, Greece, UK, Sweden, Denmark, Norway). An additional version, in French only, exists for a non-EPIC study conducted in Geneva (Switzerland). A programme such as EPIC-SOFT could therefore be of potential interest in different methodological, clinical or epidemiological studies within or outside the EPIC context and at the national (regional) or international level.

1) EPIC-SOFT as a possible reference method in calibration or validation studies

Traditionally, the use of the 24-hour diet recall in nutritional research was restricted to validation and more recently to calibration study designs involving relatively small population samples (12, 24). These methodological sub-studies are usually conducted to validate dietary questionnaires designed for specific research objectives and within single relatively homogeneous populations, using different dietary methodologies. Although the EPIC-SOFT programme was initially developed for international uses, the software could also be of interest as an alternative reference validation or calibration method in studies conducted in countries for which a version of the programme exists. EPIC-SOFT has, for example, already been used as a reference validation method in various studies in Germany (25, 26) and in Spain and Norway (unpublished data). In these validation studies, repeated EPIC-SOFT measurements per subject are collected in order to estimate long-term usual individual intakes.

The use of a common dietary tool makes it possible to compare, in the future, the results of different validation and calibration studies across Europe.

2) EPIC-SOFT as a possible reference method for measuring mean and usual dietary intake distribution of populations

Improvements in comparability and exchange of dietary information and health indicators are a keys elements for future development and harmonization of public health in Europe. A number of European initiatives, evaluating public health and risk assessment at population level, are therefore increasingly concerned about methodological issues related to estimation of dietary exposures at the European level (27, 28).

As an example, a recent concerted action project "*European Food Consumption Survey Method*" (EFCOSUM) supported by the EU programme on Health Monitoring was set up with the aim of proposing a harmonized methodology for monitoring national nutritional surveys in Europe, using experience from existing initiatives (29-32) and previous concerted actions (FLAIR EUROFOODS-ENFANTS, COST-ACTION 99). A final report on the main conclusions of this methodological working group was published in 2002 (28). Although different dietary methods are used in current national monitoring surveys (33), 24-hour diet recall was proposed as the optimal method for inter-cultural comparisons across Europe (34). Two non-consecutive 24-hour diet recalls were proposed to be collected from a representative sample in each country, in order to estimate the distribution of individual intakes and correct for random errors due to within-subject day-to-day variations (35, 36). These 24-hour diet recalls will be used both as the main method, to estimate individual intakes, or alternatively for calibration in on-going national studies where the adhesion to the reference methodology could be delayed. EPIC-SOFT was recommended as the first choice instrument for collecting standardized 24-hour diet recall measurements across countries (37, 38). Alternative softwares and/or manual methods were, however, not excluded provided that they optimize their comparability with EPIC-SOFT. In the absence of a reference European food composition table, it was also recommended to make use of the on-going European Nutrient DataBase project (ENDB) coordinated by IARC (39). From the general theoretical concept of standardizing food composition tables at the European level developed in chapter 4, the ENDB project has progressed well during the last 2 years (39). The national food composition tables, used as the main source of information to compile the EPIC food composition matrixes discussed in Chapter 4, were documented in terms of origin of nutrient values (e.g. analytical, borrowed or calculated values) as well as definition and chemical analytical methods used. This information was loaded in a specifically designed software (FTI) according to a common format in order to be evaluated and standardized before being used for the actual compilation. This will concern first a restricted number of nutrients (energy, macro-nutrients, dietary fibres) and will, subsequently, be extended to a larger list of food components (e.g. fatty acids and fractions, vitamins, minerals). The lack of chemical data and comparable food description systems, food component nomenclature and definition, description of analytical and calculation methods are major obstacles for the development of strictly comparable European food composition tables (39, 40).

Although EPIC-SOFT has several advantages for use in pan-European monitoring surveys, additional work is needed to adapt it to new specific requirements (41). EPIC-SOFT was initially developed for adult study populations but further work would be needed to adapt it to children and teenagers or different ethnic European groups. Furthermore, additional programmes to make the maintenance of EPIC-SOFT databases more independent of EPIC logistic need to be developed and the standardized food composition databases currently being developed need to be implemented (39). Consequently, the maintenance of the EPIC-SOFT food and nutrient databases, to adapt the marker and individual food habit changes over the time, should also be addressed. Furthermore, new EPIC-SOFT versions for European countries not participating in EPIC should be developed (e.g Portugal, Belgium, Ireland, Iceland or East European countries).

The preparation of new EPIC-SOFT versions for countries that joined the EPIC project when the software was already developed (i.e. Sweden, Denmark and Norway) benefited from existing versions and the full concept of standardisation. The operational versions were therefore completed much faster (2 x 6 months of a dietician's time) than the initial ones. However, it may be necessary to allow for more time, particularly if additional country-specific photographs need to be prepared and added to the EPIC-SOFT picture book and related databases.

A recent analysis was carried out to estimate whether an existing EPIC-SOFT version (i.e. Sweden) could be useful to develop a new one (i.e. Finland) (41). Food and recipe lists and portion pictures available in the Swedish EPIC-SOFT version (n= 2311 items and 140 photos) and in the Finnish 24-hour diet recall computerized programme used in the national FINDIET 1997 study (n= 1373 items and 126 photos) were compared. About 95% of the foods and mixed recipes in the Swedish lists were found to be consumed in Finland, although the Swedish EPIC-SOFT version contained 50% more items than the FINDIET lists. 62% of the Finnish photos have an equivalent in the EPIC-SOFT picture book but the range and sizes of the photographs were larger in the latter; particularly for vegetables. The reason for this is that the same picture book was used across EPIC countries, so the selection of pictures was adapted to cover the large range of food consumption existing across EPIC centres, particularly for vegetables (42).

This first experience could be extended to other countries. It is anticipated that the existing versions available for Southern countries could be used for Portugal, the Dutch version for Belgium (a French version already exists), The British version for Ireland, and the Nordic versions for Finland and Iceland, thus making EPIC-SOFT available for a wider range of potential users across Europe.

3) EPIC-SOFT as a possible reference method to estimate long-term individual intakes in large epidemiological studies

Although in the past it was always considered that the collection of multiple dietary records or recalls was inappropriate for large epidemiological studies (22), recent arguments suggest that this statement should be reconsidered, as well as the new balance of advantages and disavantages of this approach compared to FFQs in large epidemiological studies. At that stage, our purpose is not to make a definitive demonstration but rather to raise questions on a methodological issue which may have substantial implications for the design of future nutritional studies, particularly when conducted at the international level. Therefore, a series of arguments supporting the need to re-consider open-ended dietary methods as competitive possible alternatives to FFQs in epidemiological studies will be enumerated and discussed.

24-hour diet recalls or records can theoretically be used to estimate individual usual long-term diet if a sufficient number of measurements are collected to adequately cover within-subject variations in dietary consumption. This number depends on day-to-day variations in the food (group) or nutrient of interest and the precision of the measurements required (43-45). A minimum of 3 to 10 days is required for estimating usual individual energy and macro-nutrient intakes whereas more than 20 days are needed for nutrients with large day-to-day variations such as cholesterol and vitamins C or A (6, 22). Weekday and seasonal variations, which also contribute to the consumption variance, should be taken into account in the sampling procedures. As a practical consequence, these statistical requirements increase the complexity and cost of study design and the burden for both the participants and the investigators. As an example, data handling of open-ended dietary measurements after collection, particularly with the traditional method on paper which predominated in 70s and 80s, represented more than two thirds of the investigator's time (46).

These major disadvantages explain historically the exclusion of the 24-hour dietary recall as an option for large epidemiological studies and the emergence of different versions of FFQs as a cheap method for measuring long-term individual intake in epidemiological studies (8). It is also important to remember that until the end of the 80s the dominant epidemiological study designs were essentially case-control studies. In contrast to prospective studies, individual exposures measured in both cases and controls can only be retrospective using dietary questionnaires on long-term past intakes, such as FFQs or dietary history questionnaires.

The appropriateness of FFQs to assess and compare usual dietary intake has recently been questioned by different authors (47-49). Such questionnaires rely on subjects' long-term memory and on their capacity to conceptualize and summarize the complex notion of average frequency of consumption and food portions. Furthermore, food frequency questionnaires are country-specific and based on a relatively short closed list of food items, frequency categories and standard portion sizes. The increase in the number of foodstuffs available on the market and the dietary changes observed in western populations (50-51), particularly in younger generations (52-54), lead to question whether traditional closed dietary questionnaires are still appropriate to capture the full complexity of dietary patterns existing within and between countries. Recent validation studies using biological markers as reference measurement confirm that, although FFQs may be acceptable for ranking individuals with varying degrees of misclassification, open-ended methods give closer results to the estimated true of both individual and population dietary intakes (10, 12, 19, 20, 25, 55, 56).

The problems associated with FFQ can be even greater when used as a common method in multi-centre or multi-ethnic studies. For example, the closed nature of the FFQ means that it is necessary to include in the same list both the food items contributing to *and* discriminating between individual food consumption within and between populations, in order to estimate accurate and comparable individual intakes and to correct for possible confounders. Hankin et al. (57) developed and validated a dietary questionnaire for different ethnic groups in Hawaii and Los Angeles (58). In contrast to international studies, this was facilitated because all subjects were from the same country so the same foods consumed in one ethnic group were possibly also eaten in small amounts by other groups. However, despite standardisation, Stram et al. reported statistically significant differences in the slope estimates in calibration equation between sub-cohorts suggesting that the questionnaire performed differently across ethnic groups (59).

These disadvantages associated with FFQs should favour further the consideration of open-ended methods such as dietary records and recalls as a possible alternative for estimating long-term individual intakes, particularly in large epidemiological studies. The studies conducted as part of this thesis have shown that 24-hour diet recalls can be standardised reasonably well across populations (chapter 2, 3 and 4) and provide reliable and detailed dietary estimates for between-population comparisons (chapter 6 and 7).

An alternative approach to repeated 24-hour diet recalls was successfully used in the EPIC-UK branch, where a validated seven-day diary combined with a short semi-quantitative FFQ and a self-reported 24-hour diet recall were collected from each individual at baseline (10). A comprehensive data entry system (DINER), for entering the dietary information collected on paper, has been developed and tested for between-coder standardisation (23). In the UK context, this method gave a high response (~ 90%), obtained even after three repeated dietary collections at 0, 18 and 48 months. This method worked well in the UK but would not be easily applicable in other settings, particularly large multi-centre (multi-ethnic) studies. Indeed, as discussed in Chapter 2 and 6 and elsewhere (14, 60-62), the respondents' individual bias is probably the major source of dietary measurement errors and the most difficult to control for. The cost and time needed for the data entry and the subsequent delay in obtaining a full dataset for analysis, is another important limiting factor in using such methods in large international epidemiological studies. Indeed, despite a huge successful effort to automate and standardize the data entry procedures, it still takes about 2 hours to complete the coding and data entry for one diary. Subsequently 12% (n=5979) only of the total diaries (n=50 465) collected in the EPIC Cambridge cohort, on three occasions between 1993 to 2000, are actually available for statistical analysis in 2001 (23).

The number of supporting arguments for considering 24-hour recall (or dietary records) as a more competitive alternative method for estimating usual intakes in large epidemiological studies is greater than in the past. Its actual practical feasibility should therefore be investigated further in the future. The availability of a computerized face-to-face 24-hour diet recall and a possible alternative approach by telephone should help to develop a better cost-effective study design and increase the participation rate. Furthermore, the minimum number of days required per individual should be investigated for the specific purposes of large epidemiological studies. The combination of a restricted number of 24-hour diet recalls compatible with logistical constraints and a short questionnaire on foods with the greatest day-to-day variability in intake could also be considered as an alternative method.

Conclusion

EPIC was the first project to implement a very large calibration study based on a common standardized dietary method in the context of a multi-centre cohort study. Despite methodological and practical difficulties, the standardisation and setting up of the calibration worked well, and the calibration sub-sample can reasonably be considered to be representative of the EPIC cohort in most centres. The standardization across interviewers and the validation of 24-hour diet recalls for between-cohort calibration using urinary nitrogen measurements as reference have shown very encouraging results which are supported by other empirical analyses on under- and over-reporting of energy intake conducted on the EPIC calibration data. However, measurement errors, particularly conscious or unconscious behaviour of respondents and/or interviewers, cannot be prevented entirely and the problems of measurement error, its structures and determinants need to be considered when applying the calibration. In certain centres, we identified more serious problems of mis-reporting, inadequate representativeness of the calibration sub-sample compared to the entire cohort or imperfect coverage of the days of the week and seasons, which should be taken into account when using the calibration. This first experience should lead to further improvements in the reference calibration method and to the setting up of nested calibration sub-studies in the future. The international standardisation of the 24-hour diet recall also showed that this method is particularly appropriate for describing in details food consumption and dietary patterns across heterogeneous populations.

Progress in any branch of science has frequently been associated with improvement or new development of methodological concepts or tools. The field of nutritional epidemiology, focusing on the association between diet and diseases, is particularly concerned with methodological issues related both to study designs and methods used to collect and correct for errors in dietary measurements. This applies also, however, to a number of other domains of research, including public health, industry and economics, where diet is of central interest.

The standardisation and calibration of dietary measurements at the international level opens a promising area of (methodological) research and analyses on diet by changing both the scale and the size of the investigation. Currently, a convergence and concordance in *requests* (i.e. standardized methodologies for international use), *method* (e.g. 24-hour recall emerging as reference validation or calibration or individual measurements in international studies) and *time* (promoted by EU initiatives and programmes) from different domains of dietary applications can be observed. A unique historical opportunity exists today to go beyond the standardisation of 24-hour diet recall across countries and extend it, as far as possible, to standardisation across dietary disciplines. From a European Community point of view, inter-disciplinary interactions, including food industries, would be the best cost-effective way to define and standardize common dietary methodologies, including food chemical component databases, across state members and to ensure their maintenance and improvement over time at the European level.

References

1. Riboli E. Nutrition and cancer: background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). Ann Oncol. 1992; 3:783–791.
2. Kaaks R, Plummer M, Riboli E, Estève J, van Staveren WA. Adjustment for bias due to errors in exposure assessments in multi-centre cohort studies on diet and cancer: a calibration approach. Am J Clin Nutr. 1994;59:S245-50.
3. Plummer M, Clayton D, Kaaks R. Calibration in multi-centre cohort studies. Int J Epidemiol. 1994;23:419-26.
4. Block G: A review of validations of dietary assessment methods. Am J Epidemiol. 1982; 115, 492-505.
5. Pietinen P: An overview of the dietary methods used in epidemiologic studies on nutrition and cancer. In Diet and Human Carcinogenesis, eds. J Joossens, M Hill & J Geboers. Amsterdam: Elsevier.1985: 235-244.

6. Bingham SA The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. Nutrition Abstracts and Reviews (series A). 1987; 57: 705-742.
7. Cameron ME & van Staveren WA. Manual on Methodology for Food Consumption Studies. Oxford: Oxford Medical Publications. 1987.
8. Willett WC. Nutritional Epidemiology, 1st ed. NewYork: Oxford University Press. 1990.
9. Black AE, Goldberg GR, Jebb SA, Livingstone MBE, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys (1991). Eur J Clin Nutr. 1991;45:583-599.
10. Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, Sneyd MI, Key TJA, Roe I, Day NE. Comparison of dietary assessment methods in nutritional epidemiology: weighted records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. Br J Nutr. 1994; 72: 619–43.
11. Black AE, Bingham SA, Johansson G, Coward WA. Validation of dietary intakes of protein and energy against 24 hour urinary N and DLW energy expenditure in middle-aged women, retired men and post-obese subjects: comparisons with validation against presumed energy requirements. Eur J Clin Nutr. 1997;51:405-13.
12. Margetts, B.M., Pietinen, P., Riboli, E. EPIC European Prospective Investigation into Cancer and Nutrition. Validation studies on dietary assessment methods. Int J of Epidemiol. 1997; 26 (Suppl. 1): 1-189.
13. Willett WC. Nutritional Epidemiology, 2nd ed. NewYork: Oxford University Press. 1998.
14. Ferrari P, Slimani N, Ciampi A, Trichopoulos D, Naska A, Lauria C, Veglia F, Bueno-de-Mesquita HB, Ocké MC, Brustad M, Braaten T, Tormo MJ, Amiano P, Mattisson I, Johansson I, Welch A, Davey G, Overvad K, Tjønneland A, Clavel-Chapelon F, Thiebaut A, Linseisen J, Boeing H, Hemon B, Riboli E. Evaluation of mis-reporting in the 24-hour diet recalls in EPIC. Public Health Nutr. 2002; (in press).
15. Goldberg GR, Black AE, Jebb SA et al. Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. Eur J Clin Nutr. 1991;45:569-81.
16. Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy inake: basal metabolic rate. A practical guide to its calculation, use and limitations. Int J Obes Relat Metab Disord. 2000; 24:1119-30.
17. Haftenberger M, Schuit AJ, Tormo MJ et al. Physical activity of subjects aged 50 to 64 years involved in the European Prospective Investigation into Cancer and Nutrition (EPIC). Publ Health Nutr. 2002 (in press).
18. Black AE. The sensitivity and specificity of the Goldberg cut-off for EI:BMR for identifying diet reports of poor validity. Eur J Clin Nutr. 2000;54:394-404.

19. Black AE, Bingham SA, Johansson G, Coward WA. Validation of dietary intakes of protein and energy against 24 hour urinary N and DLW energy expenditure in middle-aged women, retired men and post-obese subjects: comparisons with validation against presumed energy requirements. Eur J Clin Nutr. 1997;51:405-13.
20. Kipnis V, Midthune D, Freedman LS, Bingham S, Schatzkin A, Subar A, Carroll RJ. Empirical evidence of correlated biases in dietary assessment instruments and its implications. Am J Epidemiol. 2001; 153: 394–403.
21. Brustad M, Skeie G, Braaten T, Slimani N, Lund E. Comparison of telephone versus face-to-face interview in the assessment of dietary intake by the 24-hour recall EPIC-SOFT programme – the Norwegian calibration study. Eur J Clin Nutr.(in press).
22. Buzzard M. 24-hour dietary recall and food record methods. ed. W Willett, Monogr. Epidemiology Biostatist. 30 50-73. New York, Oxford: Oxford University Press. 1998.
23. Welch AA, McTaggart A, Mulligan AA, Luben R, Walker N, Khaw KT, Day NE, Bingham SA. DINER (Data Into Nutrient for Epidemiological Research) - a new data-entry programme for nutritional analysis in the EPIC-Norfolk cohort and the 7-day diary method. Publ Health Nutr. 2001; 146:1017-22.
24. Thompson FE, Moler JE, Freedman LS, Clifford CK, Stables GJ, Willett WC. Register of dietary assessment calibration-validation studies: a status report. Am J Clin Nutr. 1997; 65: S1142–7.
25. Kroke A, Klipstein-Grobusch K, Voss S, Moseneder J, Thielecke F, Noack R, Boeing H. Validation of a self-administrated food-frequency questionnaire administrated in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with a doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. Am J Clin Nutr. 1999; 70(4):439-47.
26. Mensink GB, Haftenberger M, Thamm M. Validity of DISHES 98, a computerized dietary history interview: energy and macro-nutrient intakes. Eur J Clin Nutr. 2001, 55:406-17.
27. Barlow S, Dybing E, Edler L, Eisenbrand G, Kroes R, van den Brandt P (eds). Food Savety in Europe (FOSIE): Risk Assessment of Chemicals in Food and Diet Food Chem Toxicol. 2002; 40: 140-427.
28. Löwik MRH, Brussaard JH (eds). EFCOSUM: European Food Consumption Survey Method. Eur J Clin Nutr. 2002; 56 (suppl 2): S1-96.
29. Bothig S. WHO MONICA Project: objectives and design. Int J Epidemiol. 1989; 18 (3 suppl 1): S29-37.
30. De Groot LCPGM, Hauvast JGAJ, van Staveren WA. Nutrition and health of elderly people in Europe: the EURONUT-SENECA study. Nutr Rev. 1992; 50: 185-94.
31. Lagiou P, Trichopoulou A and the DAFNE contributors. The DAFNE initiative: the methodology for assessing dietary patterns across Europe using household budget survey data. Public Health Nutr. 2001; 4 :1135-1142.

32. Riboli E, Hunt K, Slimani N, Ferrari P, Norat T, Fahey M, Charrondière R, Hémon B, Casagrande B, Vignat J, Overvad K, Tjønneland A, Clavel-Chapelon F, Wahrendorf J, Boeing H, Trichopoulos D, Trichopoulou A, Vineis P, Palli D, Bueno-de-Mesquita HB, Peeters PHM, Lund E, Engeset D, González CA, Barricarte A, Berglund G, Hallmans G, Day N, Key T, Kaaks R, Saracci R. EPIC: study population and data collection. Public Health Nutr. 2002; (in press).
33. Verger et al. Comparability of existing data. Improvements of intake assessment using currently available data based on individual consumption surveys. Eur J Clin Nutr. 2002; 56: S18-24.
34. Biró G, Hulshof KFAM, Ovesen L, Amorim Cruz JA, for the EFCOSUM group. Selection of methodology to assess food intake. Eur J Clin Nutr. 2002; 56, (suppl 2): S25-32.
35. Volatier JL, Turrini A, Welten D for the EFCOSUM group. Some statistical aspects of food intake assessment. Eur J Clin Nutr. 2002; 56, (suppl 2): S46-52.
36. Hoffmann K, Boeing H, Dufour A, Volatier JL, Telman J, Virtanen M , Becker W, De Henauw for the EFCOSUM group. Estimating the distribution of usual dietary intake by short-term measurements. Eur J Clin Nutr. 2002, 56 (Suppl 2): S53-62.
37. De Henauw, Brants HAM, Becker W, Kaic-Rak A, Ruprich J, Sekula W, Mensink GBM, Koenig JS for the EFCOSUM group. Operationalisation of food consumption surveys in Europe : Recommendations from the EFCOSUM (European Food Consumption Surveys Methods) project. Eur J Clin Nutr. (2002); 56 (suppl 2), S75-88.
38. Brussaard JH, Löwik MRH, Steingrimdottír L, Møller A, Kearney J, De Henauw S, Becker W for the EFCOSUM group. A European food consumption survey method - conclusions and recommendations. Eur J Clin Nutr. (2002); 56 (suppl 2), S89-94.
39. Charrondière UR, Vignat J, Møller A, A, Ireland J, Becker W, Church S, Farran A, Holden J, Klemm C, Linardou A, Mueller D, Salvini S, Serra L, Skeie G, van Staveren W, Unwin I, Westenbrink S, Slimani N, Riboli E.. The European Nutrient Database (ENDB) for Nutritional Epidemiology. J Food Comp Anal. 2002 (in press).
40. Deharveng G, Charrondière R, Slimani N, Riboli E & Southgate DAT: Comparison of food composition tables available in the nine European countries participating in EPIC. Eur J Clin Nutr. 1999; 53, 60-79.
41. Slimani N & Valsta L for the EFCOSUM group. Perspectives of using the EPIC-SOFT programme in the context of pan-European nutritional monitoring surveys: methodological and practical implications. Eur J Clin Nutr. 2002; 56 (suppl 2), S63-74.
42. Agudo A, Slimani N, Ocké MC, Naska A on behalf of the EPIC group. Consumption of vegetables, fruit and other plantfoods in the EPIC cohorts from 10 European countries Public Health Nutr. 2002; (in press).

43. Beaton GH, Milner BA, Corey P, McGuire V, Cousins M, Stewart E, de Ramos M, Hewitt D, Grambsch PV, Kassim N & Little Ja. Sources of variance in 24-hour dietary recall data : Implications for nutrition study design and interpretation. Am. J. Clin. Nutr. 1979; 32, 2546-2559.
44. Beaton GH, Milner J, McGuire V, Feather TE, Little JA. Source of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr. 1983;37:986-95.
45. Nelson M, Black AE, Morris JA, Cole TJ. Between- and within-subject variation in nutrient intake from infancy to old age: Estimating the number of days required to rank dietary intakes with desired precision. Am J Clin Nutr. 1989; 50:155-67.
46. Black AE. The logistic of dietary surveys. Hum Nutr Appl Nutr. 1982; 32:85-94.
47. Briefel RR, Flegal KM, Winn DM, Loria CM, Johnson CL, Sempos CT: Assessing the nation's diet: Limitations of the food frequency questionnaires. J am diet assoc. 1992; 92:959-962.
48. Kushi LH. Gaps in epidemiological research methods: design considerations for studies that use food-frequency questionnairesl. Am J Clin Nutr. 1994; 59 (suppl): 180S-184S.
49. Liu K. Statistical issues related to semiquantitative food frequency questionnaires. Am J Clin Nutr. 1994;59 (Suppl), 262S-265S.
50. Hill MJ. Changing pattern of diet in Europe. Eur J Cancer Prev. 1997; 6 (Suppl 1): S11–3.
51. Rumm-Kreuter D. Comparison of the eating and cooking habits of Northern Europe and the Mediterranean countries in the past, present and future. Int J Vitam Nutr Res. 2001; 71: 141–8.
52. Amorim Cruz JA. Dietary habits and nutritional status in adolescents over Europe – Southern Europe. Eur J Clin Nutr. 2000, 54 (Suppl 1): S29–35.
53. Rolland-Cachera MF, Bellisle F, Deheeger M. Nutritional status and food intake in adolescents living in Western Europe. Eur J Clin Nutr. 2000; 54 (Suppl 1): S41–6.
54. Samuelson G. Dietary habits and nutritional status in adolescents over Europe. An overview of current studies in the Nordic countries. Eur J Clin Nutr. 2000; 54 (Suppl 1): S21–8.
55. Bingham SA, Cassidy A, Cole T et al. Validation of weighed records and other methods of dietary assessment using the 24 h urine nitrogen technique and other biological markers Br J Nutr 1995;73:531-50.
56. Sawaya AL, Tucker K, Tsay R, Willett W, Saltzman E, Dallal GE, Roberts SB. Evaluation of four methods for deter energy intake in young and older women: comparison with doubly labelled water measurements of total energy expertiture. Am J Clin nutr. 1996; 63: 491-499.
57. Hankin JH & Wilkens LR. Development and validation of dietary assessment methods for culturally diverse populations. Am J Clin Nutr. 1994; 59 (Suppl):198S-200S.

58. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, Stram DO, Monroe KR, Earle ME, Nagamine FS. A multiethnic cohort in Hawaii and LosAngeles: baseline characteristics. Am J Epidemiol. 2000, 15: 358-70.
59. Stram DO, Hankin JH, Lynne R, Wilkens LR, Pike MC, Monroe KR, Park S, Henderson BE, Nomura AMY, Earle ME, Nagamine FS, Kolonel LN. Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. Am J Epidemiol. 2000; 151: 358–70.
60. Klesges RC, Eck LH, Ray JW. Who underreports dietary intake in a dietary recall ?. evidence from the second national health and nutrition examination surveys. Journal of Consulting and Clinical Psychology. 1995; 63: 438-444
61. Pryer JA, Vrijheid M, Nichols R, Kiggins M, Elliott P. Who are the "low energy reporters in the dietary and nutritional surveys og British adults ?. Int J Epidemiol, 26: 146-154
62. Black AE, Cole TJ. Biased over- or under-reporting is characteristic of individuals whether over time or by different assessment methods. J Am Diet Assoc, 2001 101 (1):70-80.

Summary

Multi-centre (or ethnic) cohort studies provide an ideal setting to investigate the association between long-term dietary exposure and diseases by overcoming most of the methodological problems traditionally associated with case-control studies. These studies are designed to increase the statistical power to detect an association between diet and diseases by including large study populations varying both in the type of dietary patterns and in cancer incidence rates, increasing the heterogeneity of both dietary exposure and outcome diseases. The use of different dietary assessment methods across study populations for capturing local dietary patterns, however, increases the likelihood of introducing systematic errors varying in nature and magnitude across study populations and distorting the estimation and interpretation of the relationship between diet and disease when all data are pooled together. Calibration was proposed as a method to correct, at the population level, for systematic over- or under-estimation of mean dietary intakes in each study population (*between-cohort calibration*) and, at the individual level, for the attenuation of the relative risk measuring the relationship between diet and disease due to random errors in dietary measurements (*within-cohort calibration*). In practice, this calibration approach consists in using in addition to the dietary method used to collect individual long-term dietary intake at baseline a second highly standardized dietary method in a representative sample of the entire cohort as common reference across study populations. Apart from the quantitative biological markers (i.e. urinary nitrogen and doubly labelled water measurements), the best reference calibration methods are considered to be 24-hour diet recall, food weighed or recorded methods.

Despite increasing interest in the concept of calibration in dietary surveys there is still little experience in the use and standardisation of a common reference dietary method, particularly in the context of large nutritional multi-centre studies. This thesis deals with translating the theoretical statistical theory of calibration into an epidemiological application in the context of a large multi-centre cohort study on diet and cancer, the EPIC project. The European Investigation into Cancer and Nutrition (EPIC) involves half a million subjects from 23 centres in 10 Western European countries (Denmark, France, Germany, Greece, Italy, Norway, Spain, The Netherlands, Sweden and United Kingdom). Information on individual long-term dietary intakes was collected at baseline from each participants using country-specific dietary questionnaires and a single 24-hour diet recall, obtained from a representative sample of the entire cohort, as the reference calibration method.

In the absence of an already existing reference standardized 24-hour diet recall for use in an international European context, an *ad hoc* method was developed. It was considered that the development of a computerized face-to-face 24-hour diet recall was the only suitable approach to ensure the highest possible level of standardization of the 24-hour diet recalls among the 90 interviewers involved in the 23 EPIC centres. In the first part of the thesis, we described the theoretical concept on the standardisation and the main sources of errors associated with 24-hour diet recalls and the rationale and specific approaches developed to prevent and/or minimize each of them in an international context. These possible sources of measurement error are numerous and varied but they can schematically be associated to three main sources: *the interviewer* (who asks the questions), *the respondent* (who answers the questions) and *the dietary method* (24-HDR interview). Although the respondent's (and interviewer's) memory remains the central inherent problem, particularly with regard to food reporting, identification and quantification, other sources of errors (or systematic differences) may also occur at the different stages of the interview procedure. Type and number of quantification methods used, food coding and classification systems, methods of data calculation of foods or mixed recipes, food composition tables, statistical methods of analysis and overall quality controls are the other possible sources of errors discussed. Chapter 2 also gives a first insight into the level of standardisation of the 24-hour diet recalls across interviewers, which will be developed further in chapter 6. An analysis of variance using a random effect model, where mean total energy per interviewer was used as the dependent variable and the « interviewer » variable was considered as a random effect, show that 24-hour diet recalls measurements were reasonably well-standardized across the interviewers from the same country, although a gender and, sometimes, an interviewer effect was observed. No statistical difference between interviewers was observed in any country for men and, in five out of eight countries for women. This after adjustment for physical activity, the computer programme version used to collected the dietary information and the exclusion of one interviewer in Germany (for men) and one in Denmark (for women). The percentage of interviewers within ± 5% and ± 10% of the country mean energy intake represents, respectively, 71% and 98% in men and 74% and 94% in women, whereas the difference between mean centre energy intake (i.e. without stratifying by interviewers) and country mean energy intakes is within ± 5% for 82% of the centres for men and 89% for women, without exclusion of any extreme energy intake values.

The analysis of the possible sources of error (or systematic differences) in the 24-hour diet recall interview, and the theoretical procedures to prevent or minimize them detailed in chapter 2, served as background to develop an ad hoc computer program (EPIC-SOFT) used as reference calibration across the EPIC centres. The structure, content and main functions of the EPIC-SOFT program is described in chapter 3. EPIC-SOFT has the same structure and interface, translated into 10 languages, in all the EPIC countries. Common rules were pre-entered into the system to describe, quantify and probe about 1500-2000 foods and 150-350 mixed recipes depending on the country. During the interview, the interviewers are guided similarly across countries by windows automatically prompted on the screen on how to search, describe, quantify and check each reported item selected from the pre-defined food or recipe lists. The level of detail for describing food is, for example, standardized by means of automatic prompts on the screen with a series of questions (called « facets ») and their pre-defined possible answers (called « descriptors ») common across countries. The same quantification methods adapted to the countries (e.g. household measurements, standard units) were used to quantify similar foods, including a common picture book containing a set of photos of 140 foods or recipes. Different (semi-) automatic quality controls on the information reported were included at the different stages of the interview procedure in order to prevent or correct errors or inconsistencies, while the subject is still present. Furthermore, the same food classification and format were used to classify and export dietary data and facilitate their comparison, exchange and pooled analysis. The information finally stored in EPIC-SOFT is not restricted to final food quantities only. The food consumption occasion (meal), place, hour of consumption of each individual item was systematically stored, as well as the food- (or recipe-) specific facets/descriptors and methods used to describe, quantify or probe them, and individual information on the interviewer and the respondent. This information was crucial for the (semi-) automatic maintenance and re-calculation of the data using additional specifically design programs, and permit more refined analyses on the interviewers, cooking methods or meal patterns across centres, for example.

In order to study the relationship between diseases and nutrient (or other food component) intakes, standardized food composition tables are required. Although national food composition tables, varying in content and quality, exist across Europe, no common reference standardized food composition table for use at the international level exists so far. Chapter 4 presents an *ad hoc* approach to standardize the EPIC food composition databases across countries, after speficying their needs in the context of large multi-centre nutritional studies such as EPIC. The problem of the comparability and standardisation of food composition tables has been addressed by different international and regional initiatives (EUROFOODS, COST 99, NORFOODS or INFOODS), which provide guidelines and tools to improve the harmonization of analytical laboratory methods, definition of mode of expression of foods and nutrients. However, despite all these efforts, large differences still exist between national food composition and nutrient databases. This is to a large extent due to the difficulties and cost of compiling comparable nutrient values. Another possible reason is that so far attempts at standardization between national food composition tables have always focused on harmonizing of the mode of expression, unit and chemical method of analyzing nutrients (*horizontal axis of the matrices*). Complete standardisation of existing food composition tables has, however, always come up against the difficulty of standardizing the food lists (vertical axis of the matrices) which vary enormously both in number and detail across food composition tables. Furthermore these lists do not always correspond to the type and level of information actually reported by the study subjects. In order to overcome these two major methodological problems, chapter 4 proposes to re-build completely the food composition table matrices using as vertical axis the standardized food items as collected from the 36,900 subjects by means of EPIC-SOFT in the 10 EPIC countries. The standardization of the horizontal axis of the matrices involved selecting the common nutrients and their mode of expression, unit and chemical method of analysis using the guidelines prepared by regional and international expert groups. The implications of this approach on future priorities and development of national food composition tables are also discussed in depth. This general concept is currently used to develop standardized food composition tables for EPIC and other multi-centre European studies in an on-going project (ENDB project). However this goes beyond the subject of this thesis.

In the second part of this thesis, chapter 5 presents the design and population of the calibration sub-studies set up in each EPIC centre. Furthermore, to estimate the representativeness of the calibration sub-sample compared to the whole cohort, a series of subjects' characteristics known to influence dietary intakes (e.g. anthropometry, smoking status, education level and physical activity) are compared between the two population groups. In order to obtain a representative sample of the entire EPIC population, a high participation rate must be achieved from the subjects invited to participate in the calibration study. 70% of study centres reported a participation rate above 75%. This rate was consistently higher when the subjects were recruited immediately after the baseline examination than when invited to came back at a later date. The response rate was also comparately lower (≤ 60%) in centres where (representative) samples (UK general population and Norway) or atypical population group (« health conscious » group in Oxford) were recruited. In most centres, after age adjustement, no differences were observed between the calibration and whole cohort sample for height, weight, BMI and smoking status. Greater differences were observed between level of education and physical activity, although the absolute differences were in most cases modest. Furthermore, the comparison of the mean consumption of 16 food groups estimated from the centre-specific baseline dietary questionnaires confirmed that the calibration sub-sample and the entire EPIC cohort are fairly comparable. Overall 89% of the centre-sex-food group combinations show a mean difference lower than 10% (69% had a difference within ± 5%). 59% of the differences above 10% were concentrated in 4 out of 24 redefined EPIC centres. The time needed to perform the interviews with EPIC-SOFT (~30-35 mn) was quite comparable across centres and compatible with the cost and logistic constraints of large nutritional studies. However, the problem of interviewing subjects during the week-end and equally through the year, explain the imperfect distribution of day of the week and seasons observed in certain centres. In conclusion, it is shown that such a study design works relatively well in practice and can provide useful additional information for analyzing dietary data in a multi-centre study context.

In chapter 6 the actual validity of the 24-hour diet recalls as a reference measurement for between-cohort calibration is estimated in a study comparing mean dietary intakes against an independent quantitative biological marker, urinary nitrogen. The aim was to estimate whether the mean 24-hour diet recall measurements rank centres correctly when urinary nitrogen is used as reference. In the absence of doubly labelled water measurements in our study, urinary nitrogen was also considered as an indirect reference of total energy intake, assuming a high physiological correlation between nitrogen and total energy intakes. Furthermore, in order to estimate the usefulness of the calibration in EPIC, mean nitrogen (and total energy intake) obtained from non-calibrated EPIC baseline questionnaires were also compared to mean urinary nitrogen. The study involved 1103 middle aged men and women recruited in 12 EPIC centres. Despite different degrees of mis-reporting observed when the analysis was stratified by centre and gender, the partial Pearson's correlation adjusted for gender between center mean urinary and dietary nitrogen (and total energy) was high with both dietary methods (N=22). For nitrogen, the correlation was higher with 24-hour diet recalls (0.84; 0.90 after exclusion of two outliers) than dietary questionnaires (0.72). Inversely, the correlation was higher with DQ (0.82) than with 24-hour diet recall (0.69 ; 0.77 after exclusion) when energy was considered. The β coefficient of the regression line of centre means for urinary nitrogen on both dietary nitrogen estimates (or energy intakes) were not statistically significant different from 1, suggesting that overall the systematic bias across centres is of comparable magnitude across centres. Group level correlations were systematically higher in men than in women for both dietary measurements and nitrogen and total energy intakes. Different possible explanations, not mutually exclusive, may explain this gender difference. Two centres (Greece and Varese) with outlier values markedly affect the correlation in women; the range of variation of energy intakes in particular, is higher in men than women, and the random measurement errors may affect the correlation more in women who may also more likely under-report their diet, consciously or inconsciously than men, as suggested by other EPIC analyses.

Overall the analysis gives good comparable results using 24-hour diet recall or dietary questionnaires for ranking centres according to mean nitrogen or total energy intakes. However, in contrast to 24-hour diet recalls, errors in dietary questionnaires, when compared to mean urinary nitrogen, are not only different in magnitude but also in direction (over- or under-reporting) across centres. This may have implications on the dietary measurement error structure and support the potential benefit of between-cohort calibration. Furthermore, this analysis did not include all EPIC centres, particularly those who joined the EPIC project afterwards and used different dietary questionnaires than those used in the initial EPIC centres and may have depreciated the observed correlations.

Apart from their pure methodological interest for calibration purposes, the 24-hour diet recall data, collected from a large representative sample of the entire cohort (N=36,900) and using the same dietary methodology, are of great interest for investigating the dietary consumptions and patterns existing across EPIC centres. EPIC was initially set up with the aim of combining large European cohorts selected to maximize differences in dietary exposures (and outcome diseases). In chapter 7 we describe and highlight the contrasts in dietary patterns observed across 27 re-defined study centres, using a multi-dimensional graphic representation of the consumption of 22 main (sub-) food groups relative to their sex-specific overall EPIC mean. A standard multiple linear regression model was used to adjust for age, day of the week and season in order to correct for different distributions of these variables observed across centres. Although wide differences were observed across centres, the countries participating in EPIC are characterized by specific dietary patterns. Overall, Italy and Greece have a dietary pattern characterized by plant foods (except potatoes) and a lower consumption of animal and processed foods than the other EPIC countries. France and particularly Spain have more heterogeneous dietary patterns, with relatively high consumption of both plant foods and animal products. Apart from characteristics specific to vegetarian groups, the UK "health-conscious" group shares with the UK general population a relatively high consumption of tea, sauces, cakes, soft drinks (women), margarine and butter. In contrast, the diet in the Nordic countries, the Netherlands, Germany and the UK general population is relatively high in potatoes and animal, processed and sweetened/refined foods, with proportions varying across countries/centres. In these countries, consumption of vegetables and fruit is similar to, or below, the overall EPIC mean, and is low for legumes and vegetable oils. Overall, dietary patterns were similar for men and women, although there were large gender differences for certain food groups. It was concluded that there are considerable differences in food group consumption and dietary patterns among the EPIC study populations. This large heterogeneity should be an advantage when investigating the relationship between diet and cancer and formulating new aetiological hypotheses related to dietary patterns and disease.

This was the first time that calibration sub-studies using the same dietary methodology has been set up in a large multi-centre European study. Despite inherent methodological and practical problems, the calibration worked well, and the calibration sub-sample can reasonably be considered to be representative of the entire EPIC cohort in most centres. The standardization across interviewers and the validation of 24-hour diet recalls for between-cohort calibration show encouraging results supported by other empirical analyses on under- and over- reporting conducted on the EPIC calibration data and reported elsewhere. However, measurement errors, particularly conscious or unconscious behaviour of respondents and/or interviewers, cannot be prevented entirely and the problems of measurements error, its structures and its determinants need to considered when applying the calibration. In certain centres, we identified more serious problems of mis-reporting, inadequate representativeness of the calibration sub-sample compared to the entire cohort, or imperfect coverage of the days of the week or seasons which should also be taken into account when using the calibration. This first experience should also lead to further improvements in the reference calibration method and setting up nested calibration sub-studies in the future and opens perspectives of using programs such as EPIC-SOFT in other study contexts.

Samenvatting

Meerdere centra omvattende (multi-center of ethnische) cohort studies vormen een ideale setting om het verband tussen langdurige blootstelling aan bepaalde voeding en ziektes te onderzoeken, daar zij het hoofd bieden aan de meeste methodologische problemen die van oudsher worden geassocieerd met case-control studies. Deze studies zijn ontworpen om het onderscheidingsvermogen bij het opsporen van het verband tussen voeding en ziektes te vergroten door het opnemen van uitgebreide studie populaties die variëren zowel in type voedingpatroon als in kankerindicentie, waardoor de heterogeniteit van zowel de voedingsblootstelling als de ziekte- en uitkomstmaten wordt vergroot. Het gebruik van verschillende voedselconsumptiemethodes om plaatselijke voedingpatronen vast te leggen over de studie populaties, verhoogt echter de kans op het introduceren van systematische fouten die in aard en omvang kunnen verschillen en die de vaststelling en interpretatie van het verband tussen voeding en ziekte kunnen verstoren wanneer alle gegevens worden samengevoegd. Kalibratie wordt gezien als een potentiële methode om bij het meten van het verband tusen voeding en ziekte op populatie niveau te corrigeren voor over- of onderschatting van gemiddelde voedselinname per deelnemende populatie *(inter cohort kalibratie)* en, op individueel niveau, voor de attenuatie die bij het schatten van het relatieve risico optreedt ten gevolge van willekeurige fouten in voedsel metingen *(intra cohort kalibratie)*. In de praktijk bestaat deze kalibratie-benadering uit het toepassen van een tweede strict-gestandaardiseerde voedselconsumptiemethode onder een representatief deel van het gehele cohort geldend als referentiekader voor alle studie populaties naast de voedselconsumptiemethode ter vaststelling van van de individuele langdurige voedselinname bij aanvang van de studie (baseline). Naast kwantitatieve biologische indicatoren (zoals urine stikstof en dubbel gelabelde water metingen), worden als de beste referentie kalibratie methodes beschouwd de 24uurs navraag naar de voedselconsumptie en de opschrijfmethoden al dan niet met weging van voedingsmiddelen.

Ondanks toegenomen interesse voor het concept van kalibratie in voedingsstudies is er nog steeds weinig ervaring met het gebruik en de standaardisatie van een gemeenschappelijke voedselconsumptiemethode, vooral in de context van grote multi-center voedingsstudies. Deze dissertatie is gericht op de vertaling van de theoretische statistische kalibratie theorie in een epidemiologische toepassing gebaseerd op een uitgebreide multi-center cohort studie naar voeding en kanker, het EPIC project. Deze Europese studie (European Investigation into Cancer and Nutrition) omvat een half miljoen personen afkomstig uit 23 centra gelegen in 10 Westeuropese landen (Denemarken, Frankrijk, Duitsland, Griekenland, Italië, Noorwegen, Spanje, Nederland, Zweden en het Verenigd Koninkrijk).

Informatie over individuele voedselinname over een langere periode werd bij aanvang van de studie verzameld van elke deelnemer waarbij land-specifieke voedingsvragenlijsten werden gebruikt. Daarnaast werd, als referentie calibratie methode, eenmalig navraag gedaan naar de voedselconsumptie over de voorbije 24 uur, die verkregen werd bij een representatief deel van het gehele cohort.

Bij afwezigheid van een reeds bestaande gestandaardiseerde referentie 24-uurs navraag methode voor gebruik in een internationale Europese context, werd een ad hoc methode ontwikkeld. Hierbij bleek de ontwikkeling van een gecomputeriseerde *face to face* 24-uurs voedingsnavraag de enige geschikte benadering om het hoogst mogelijke niveau van standaardisatie van de 24-uurs navraagmethode onder de 90 interviewers in de 23 EPIC centra te waarborgen. In het eerste deel van deze dissertatie is het theoretische concept van de standaardisatie beschreven en de voornaamste foutbronnen bij de 24-uurs navraagmethode alsmede de rationale en specifieke benaderingen die zijn ontworpen om ieder van deze in een internationale context te voorkomen en/of te minimaliseren. Deze mogelijke bronnen van meetfouten zijn talrijk en sterk uiteenlopend, maar zij kunnen schematisch worden teruggebracht tot drie hoofdbronnen: *de interviewer* (degene die de vragen stelt), *de respondent* (degene die de vragen beantwoordt) en de *onderzoeksmethode* (navraag naar de voedselconsumptie over de afgelopen 24 uur). Alhoewel het geheugen van de respondent (en de interviewer) – vooral m.b.t. het rapporteren, identificeren en kwantificeren van voedingsmiddelen – het centrale inherente probleem ofwel sleutel blijft, kunnen andere foutbronnen (of systematische verschillen) ook voorkomen in verschillende fase van de navraagmethode. Andere mogelijke foutbronnen die hier bediscussieerd worden zijn het type en aantal gebruikte kwantificatiemethoden, voedselcoderings- en classificatiesystemen, gegevensberekeningsmethodes voor samengestelde voedingsmiddelen of recepten, voedingsmiddelentabellen, statistische analyse methodes en algehele kwaliteitscontroles.
Hoofdstuk 2 geeft een eerste inzicht in het niveau van standaardisatie van de 24-uurs navraagmethode onder interviewers, dat verder uitgewerkt wordt in hoofdstuk 6. Een analyse van variaties d.m.v. een *random effect* model, waarbij de gemiddelde totale energie per interviewer werd gebruikt als de afhankelijke variabele en de "interviewer variabele" werd beschouwd als *random effect*, laat zien dat 24-uurs navraagmethode redelijk goed gestandaardiseerd verliep onder interviewers van hetzelfde land, hoewel een effect naar geslacht en soms naar interviewer optrad. In geen van de landen werd een statistisch verschil tussen interviewers gevonden voor de mannen, en in vijf van de acht landen voor de vrouwen, hierbij rekening houdend met

lichamelijke activiteit,de EPIC SOFT versie en de uitsluiting van één interviewer in Duitsland (bij de mannen) en één in Denemarken (bij de vrouwen). Het percentage interviewers dat waarden rapporteerde binnen ± 5% en ±10% van de gemiddelde energie inneming van het land vertegenwoordigde respectievelijk 71% en 98% bij de mannen en 74% en 94% bij de vrouwen, terwijl het verschil tussen de gemiddelde energie inneming per centrum (d.w.z. zonder stratificatie naar interviewer) en per land binnen ±5% ligt voor 82% van de centra voor de mannen en 89% voor de vrouwen, zonder hierbij extreme waarden uit te sluiten.

De analyse van de mogelijke foutbronnen (of systematische verschillen) voor de 24 uurs navraagmethode, en de theoretische benaderingen om deze te voorkomen of minimaliseren, zoals uitgewerkt in hoofdstuk 2, hebben als achtergrond gediend bij de ontwikkeling van een ad hoc computerprogamma (EPIC-SOFT) dat als referentie kalibratie programma werd gebruikt in de EPIC centra. De structuur, inhoud en belangrijkste functies van het EPIC SOFT programma zijn beschreven in hoofdstuk 3. In alle EPIC landen heeft EPIC SOFT dezelfde structuur en interface, vertaald in 10 talen. Vooraf werden algemene regels in het systeem ingevoerd om circa 1500-2000 voedingsmiddelen en 150-350 verschillende recepten, afhankelijk van het land, te beschrijven, kwantificeren en na te vorsen. Gedurende het interview worden interviewers in de verschillende landen op gelijke wijze door op het scherm verschijnende 'windows'geleid, waarbij wordt aangegeven hoe ze elk gerapporteerd item uit een vooraf gedefinieerde lijst voedingsmiddelen of recepten gezocht, geschreven, gekwantificeerd en gecontroleerd moet worden. De mate van detail in het beschrijven van voedsel is door middel van standaardisatie met behulp van automatische prompts op het scherm bestaande een serie vragen ("facets" genaamd) en vooraf gedefinieerde mogelijke antwoorden ("descriptors" genaamd) vergelijkbaar gemaakt tussen landen. Dezelfde aan het land aangepaste kwantificatie methodes (bijv. huishoudmaten, standaard verpakkingen) werden gebruikt om hetzelfde voedingsmiddel te kwantificeren, gebruikmakend van een fotoboek met foto's van 140 soorten voedsel en recepten. Verschillende (semi-) automatische kwaliteitscontroles t.a.v. de gerapporteerde informatie werden ingebouwd bij verschillende fases van de interview procedure om vergissingen of inconsistenties te voorkomen of te corrigeren terwijl de persoon nog aanwezig was. Verder werd dezelfde voedsel klassificatie en format gebruikt om voedingsgegevens te klassificeren en exporteren en de vergelijking, uitwisseling en analyse van samengevoegde gegevens te vergemakkelijken. De informatie die in EPIC SOFT werd opgeslagen bleef uiteindelijk niet beperkt tot de hoeveelheid van het geconsumeerde voedsel. Het moment van voedselconsumptie (maaltijd), en de plaats en tijd van consumptie van elk individueel item werd systematisch opgeslagen, alsmede de voedsel(of recept-)specifieke facets/desciptors en de methodes die werden gebruikt om ze te beschrijven, kwantificeren of na te vorsen, als ook individuele informatie over de interviewer en de respondent. Deze informatie was cruciaal voor het (semi)automatisch bijhouden en opnieuw calculeren van de gegevens met behulp van extra specifiek ontworpen programma's, en het uitvoeren van specifieke analyses bijvoorbeeld gericht op interviewers, bereidingswijzen of maaltijdpatronen over de centra.

Om de relatie tussen ziektes en de inneming van nutrienten (of andere voedselbestanddelen) te bestuderen, zijn gestandaardiseerde voedingsmiddelentabellen nodig. Alhoewel er in Europa nationale voedingsmiddelentabellen, variërend qua inhoud en kwaliteit, bestaan, bestaat er tot dusverre geen algemeen referentie voedingsmiddelenstabel voor gebruik op internationaal niveau. Hoofdstuk 4 geeft een ad hoc benadering om de EPIC voedingsmiddelen databases te standaardiseren over de verschillende landen, nadat de behoeften in de context van grote multicenter voedingsstudies zoals EPIC gespecificeerd zijn. Het probleem van de vergelijkbaarheid en standaardisatie van voedingsmiddelentabellen is aangepakt door verschillende internationale en regionale initiatieven (EUROFOODS, COST 99, NORFOODS of INFOODS), die richtlijnen en instrumenten aanreiken ter harmonisatie van analytische laboratorium methodes en ter verbetering van het definiëren van voedingsmiddelen en voedingsstoffen. Ondanks al deze inspanningen, echter, bestaan er nog steeds grote verschillen tussen nationale voedingsmiddelen en voedingsstoffen databases. Dit is onder meer te wijten aan de moeilijkheden en kosten die gepaard gaan met het compileren van vergelijkbare nutriëntenwaarden. Een andere mogelijke reden is dat tot dusverre pogingen tot standaardisatie van nationale voedingsmiddelentabellen altijd gericht waren op harmonisatie van de uitdrukkingswijze, de uniteenheden en de chemische methode van het analyseren van voedingsstoffen *(horizontale as van de matrices).* Het probleem van volledige standaardisatie van bestaande voedselsamenstellingstabellen is echter altijd het standaardiseren van de voedsellijsten (verticale as van de matrices), die zowel in aantal als in detaillering enorm kunnen variëren tussen de verschillende tabellen. Bovendien corresponderen deze lijsten niet altijd met het type en niveau van informatie dat door de studie personen wordt gerapporteerd. Om deze twee grote methodologische problemen te overwinnen, wordt in hoofdstuk 4 voorgesteld de matrices van de voedingsmiddelentabel compleet opnieuw op te bouwen, door in de verticale as de gestandaardiseerde voedingsmiddelen, zoals die verzameld zijn van de 36.900 personen d.m.v. EPIC SOFT in 10 EPIC landen, weer te geven. De standaardisatie van de horizontale as van de matrices impliceert de selectie van de overeenkomstige voedingsstoffen en hun uitdrukkingswijze, eenheid en chemische analysemethode, met gebruikmaking van de richtlijnen zoals die door de regionale en internationale expert-commissies zijn voorbereid. De implicaties van deze benadering voor toekomstige prioriteiten en de ontwikkeling van nationale voedingsmiddelentabellen worden ook uitgebreid besproken. Dit algemene concept wordt nu gebruikt om gestandaardiseerde voedingsmiddelentabellen te gebruiken voor EPIC en andere multi-center Europese studies in een lopend project (ENDB project). Dit laatste onderwerp valt echter buiten het bereik van deze dissertatie.

In het tweede deel van deze dissertatie, wordt in hoofdstuk 5 het ontwerp en de populatie van de kalibratie sub-studies in elk EPIC centrum besproken. Bovendien wordt een aantal karakteristieken, waarvan bekend is dat ze invloed hebben op de voedselconsumptie (bijv. antropometrie, roken, ontwikkelingsniveau en lichamelijke activiteit), tussen de kalibratie sub-groep en het gehele cohort vergeleken, om zo de representativiteit van de kalibratie sub-groep te kunnen inschatten. Om een representatief deel van de gehele EPIC populatie te verkrijgen, moet er een hoge deelnamebereidheid zijn onder de personen die worden uitgenodigd deel te nemen aan de kalibratie studie. Zeventig procent van de studiecentra rapporteerde een deelnamepercentage van meer dan 75%. Dit percentage was stelselmatig hoger wanneer de personen meteen na het baseline onderzoek werden gerecruteerd dan wanneer ze later werden uitgenodigd terug te komen. Het responspercentage was ook verhoudingsgewijs lager (≤ 60%) in centra waar (representatieve) sub-groepen (UK algemene populatie en Noorwegen) of atypische populatie groepen ("gezondheidsbewuste groep" in Oxford) werden gerecruteerd. In de meeste centra werden er, na correctie voor leeftijd, geen verschillen opgemerkt tussen de kalibratie- en de hele cohort groep v.w.b. lengte, gewicht, BMI en rokersstatus. Grotere verschillen werden opgemerkt tussen educatieniveau en lichamelijke activiteit, alhoewel absolute verschillen in de meeste gevallen erg klein waren. Bovendien bevestigde vergelijking van de gemiddelde consumptie van 16 voedingsmiddelen groepen uit de centrum-specifieke baseline (voedsel)vragenlijsten dat de kalibratie sub-groep en het gehele EPIC cohort vrij goed vergelijkbaar zijn. Over het algemeen geeft 89% van de centrum-geslacht-voedingsmiddelen groep combinaties een gemiddeld verschil van minder dan 10% te zien (69% had een verschil van ± 5% of minder). Negenenvijftig procent van de verschillen boven de 10% was geconcentreerd in 4 van de 24 opnieuw gedefinieerde EPIC centra. De tijd die nodig was om de interviews met EPIC-SOFT af te nemen (~30-35 min.) was vergelijkbaar in alle centra en compatibel met de kosten en logistieke beperkingen van grote voedingsstudies. Het probleem om personen in het weekend en gelijkmatig verdeeld over het jaar te interviewen, verklaart de ongelijke verdeling over dagen van de week en seizoenen, zoals die in sommige centra werd opgemerkt. Concluderend, zien we dat een dergelijke studie-opzet vrij goed werkt in de praktijk en nuttige extra informatie kan verschaffen voor de analyse van voedingsgegevens in een multi-center studie context.

In hoofdstuk 6 wordt een beschouwing gegeven over de validiteit van de 24 uurs navraagmethode als referentie methode voor inter-cohort kalibratie, op basis van een studie waarin gemiddelde voedselinneming wordt vergeleken met een onafhankelijke kwantitatieve biologische merker, urine stikstof. Het doel was om vast te stellen of de gemiddelde 24uurs navraag metingen de centra correct rangschikken als urine stikstof wordt gebruikt als referentie. Bij afwezigheid van dubbel gelabelde water metingen in onze studie, werd urine stikstof ook overwogen als een indirecte referentie voor de totale energie inname, waarbij men uitging van een hoge fysiologische correlatie tussen stikstof en de totale energie inneming. Bovendien, om het nut van kalibratie in EPIC vast te stellen, werd de gemiddelde stikstof (en totale energie) inneming verkregen uit niet-gekalibreerde EPIC baseline vragenlijsten ook vergeleken met de gemiddelde urine stikstof. De studie omvatte 1103 mannen en vrouwen van middelbare leeftijd, die in 12 EPIC centra waren gerecruteerd. Ondanks verschillende mate van misrapportage die werd opgemerkt toen de analyse werd uitgevoerd gestratifeceerd naar centrum en geslacht, was de partiele Pearson's correlatie gecorrigeerd voor geslacht hoog bij beide methodes, tussen centrum-gemiddeld urine en voedingsstikstof (en totale energie) (N=22).
Voor stikstof was de correlatie hoger met de 24uurs navraagmethode (0.84; 0.90 na uitsluiting van twee extremen) dan met de voedselvragenlijsten (0.72). Omgekeerd, was voor energie de correlatie hoger met de voedselvragenlijsenDQ (0.82) dan met de 24uurs navraagmethode (0.69; 0.77 na uitsluiting van extremen).

De β coëfficienten van de regressielijn van de gemiddelde N-gehaltes in de urine versus beide schattingen van N uit de voeding (of de energie inneming) weken statistisch niet significant af van 1, wat suggereert dat over het algemeen de systematische afwijking over alle centra van vergelijkbare grootte is. Correlaties op groepsniveau waren systematisch hoger voor mannen dan voor vrouwen voor schattingen van de N- en energie inneming op basis van beide voedselconsumptie methoden. Er zijn verschillende mogelijke verklaringen, die elkaar wederzijds niet uitsluiten, voor dit verschil tussen de geslachten. Twee centra (Griekenland en Varese) met extreme waarden hebben een duidelijk effect op de correlatie bij de vrouwen; met name de variatie in energie inneming is hoger bij de mannen dan de vrouwen en de willekeurige meetfouten kunnen een groter effect hebben op de correlatie bij de vrouwen, die ook eerder dan mannen (bewust of onbewust) geneigd zijn hun voeding te 'onderrapporteren', zoals uit andere EPIC analyses blijkt.

Over het algemeen geeft de analyse goed vergelijkbare resultaten bij het gebruik van de 24uurs navraagmethode of voedselvragenlijsten voor het indelen van centra naar gemiddelde stikstof of totale energie inneming. In tegenstelling tot de 24uurs navraagmethode echter, zijn fouten over de centra heen in de voedselvragenlijsten, vergeleken met gemiddelde urine stikstof, niet alleen verschillend qua grootte maar ook qua richting (over- en onderrapportage). Dit zou gevolgen kunnen hebben voor de foutstructuur van de voedselconsumptiemethoden en het mogelijke nut van inter-cohort kalibratie ondersteunen. Bovendien, omvatte deze analyse niet alleen the originele EPIC centra, maar ook de centra die zich in een later stadium aan het EPIC project hebben geassocieerd, en die andere voedselvragenlijsten gebruikten dan de eerste EPIC centra, en dit zou de hoogte van de geobserveerde correlaties verminderd zou kunnen hebben.

Naast het zuiver methodologische belang voor kalibratie doeleinden, zijn de 24uurs navraag gegevens, die zijn verzameld onder een grote representatieve groep van het gehele cohort (N=36,900) gebruikmakend van dezelfde voedselconsumptiemethodologie, van groot belang voor het onderzoek naar voedselconsumptie en voedingspatronen in de verschillende EPIC centra. EPIC werd oorspronkelijk opgezet met het doel grote Europese cohorten te combineren, geselecteerd om verschillen in bepaalde voedingsblootstellingen en daaruitvolgende ziektes te maximeren. In hoofdstuk 7 worden de verschillen in voedingspatronen beschreven en benadrukt, zoals die in de 27 opnieuw gedefinieerde centra werden geobserveerd, door middel van een multi-dimensionale grafische weergave van de consumptie van 22 hoofd (sub-) voedsel groepen gerelateerd aan het algemene geslacht-specifieke EPIC gemiddelde. Een standaard multiple lineair regressie model werd gebruikt om te corrigeren voor verschillen in leeftijd, dag van de week en seizoen tussen de verschillende centra Alhoewel er grote verschillen werden gezien over de centra heen, worden de landen die deelnemen aan EPIC gekarakteriseerd door specifieke voedingspatronen. Over het algemeen hebben Italië en Griekenland een voedingspatroon dat wordt gekarakteriseerd door plantaardig voedsel (behalve aardappels) en een lagere consumptie van dierlijk en industrieel bewerkt voedsel dan de andere EPIC landen. Frankrijk en vooral Spanje hebben een meer heterogeen voedingspatroon, met een relatief hoge consumptie van zowel plantaardig voedsel als dierlijke producten. Afgezien van karakteristieken die specifiek horen bij vegetarische groeperingen, heeft de "gezondheidsbewuste"groep uit het Verenigd Koninkrijk gemeen met de bevolking van het UK in het algemeen dat zij relatief grote hoeveelheden thee, sauzen, taart, frisdrank (vrouwen), margarine en boter consumeren. Daarentegen bevat de voeding in de Scandinavische landen, Nederland, Duitsland en het VK in het algemeen relatief veel aardappels en dierlijk, industrieel bewerkt en gezoet/geraffineerd voedsel, waarbij de verhoudingen per land en centrum variëren. In deze landen is de consumptie van groenten en fruit gelijk aan of minder dan het algemene EPIC gemiddelde, en is die laag v.w.b. peulvruchten(?) en plantaardige oliën. Over het algemeen waren voedingspatronen gelijk voor mannen en vrouwen, hoewel er voor bepaalde voedsel groepen grote verschillen waren tussen de geslachten. In conclusie, blijken er aanzienlijke verschillen zijn in consumptie van bepaalde voedingsmiddelengroepen en voedingspatronen onder de verschillende EPIC studie populaties. Deze grote heterogeniteit is in principe een voordeel bij het onderzoek naar het verband tussen voeding en kanker en bij het formuleren van nieuwe etiologische hypotheses gerelateerd aan voedingspatronen en ziekte.

Dit was de eerste keer dat kalibratie sub-studies met gebruikmaking van dezelfde voedingsconsumptie methodologie zijn opgezet in een grote multi-center Europese studie. Ondanks inherente methodologische en praktische problemen, werkte de kalibratie goed en voor de meeste centra kan de kalibratie sub-groep vrij goed beschouwd worden als representatief voor het gehele EPIC cohort. De standaardisatie voor interviewer en de validiteit van de 24uurs navraagmethode voor inter-cohort kalibratie geven bemoedigende resultaten te zien, die worden ondersteund door empirische analyses gericht op onder- en overrapportage, uitgevoerd op de EPIC kalibratie gegevens en elders gerapporteerd. Meetfouten, met name bewust of onbewust gedrag van respondenten en/of interviewers, kunnen echter niet geheel voorkomen worden en de problemen m.b.t. meetfouten, hun structuur en determinanten moeten in aanmerking genomen worden bij de toepassing van de kalibratie. In sommige centra kwamen we ernstigere mis-rapportage problemen tegen, zoals onvoldoende representativiteit van de kalibratie sub-groep in vergelijking met het gehele cohort of onvoldoende dekking van alle dagen van de week of seizoenen, hetgeen ook in aanmerking genomen moet worden bij het gebruik van de kalibratie. Deze eerste ervaring zou ook moeten leiden tot verdere verbeteringen in de referentie kalibratie methode en het opzetten van op elkaar aansluitende kalibratie sub-studies in de toekomst en opent perspectieven om programma's zoals EPIC SOFT in andere studie contexten te gebruiken.

Dankwoord

The setting up of the EPIC nested calibration studies has been a great scientific and human experience. In addition to its large and complex study design, it is characterized first by the involvement of many scientists and technicians from different areas of expertise and from all the EPIC centres. Their essential contribution made it possible to achieve the work described in this thesis and maintain, over the years, a productive interaction in a good and constructive spirit. I therefore would like, first, to thank them all. I would like also to thank my colleagues at IARC particularly Corinne Casagrande, who has been involved in the study from the beginning, and Bertrand Hémon, Johanne Amoyel, Anne-Linda van kappel, Geneviève Deharveng, Ruth Charrondière, Jérôme Vignat and Pietro Ferrari who contributed, at one moment or another, to this project. I would like also to thank Sarah Somerville for her editorial assistance and Lisette de Groot, Karin Timmermans-Groenen and Rudolf Kaaks for their precious assistance in the translation of the Dutch summary.

Elio Riboli was the scientist who first dreamt about EPIC before, applying his enthusiasm and abilities to mobilize others to aim for a common objective, making it true. I would like to thank him again for giving me the opportunity to complete this thesis within my activities on the EPIC project and for acting as my supervisor in Lyon. I would also like to thank, Wija van Staveren, my second supervisor who, from the beginning has always been available and supportive, both from a scientific and a human point of view.

My colleague Rudolf Kaaks who, in his PhD thesis, developed the idea of using a calibration approach in EPIC, gave me the rationale to start mine. I hope have convinced him that the "philosophical" discussions on the calibration issues we had some years ago, as reported in the acknowledgements of his PhD thesis, were not in vain.

Finally, I want to thank my husband (Michel) and my three sons Michaël, Alexandre and Hugo who, despite their young age, understood that their mother sometimes needed time and understanding during stressful periods in order to complete a work which had an inexplicable importance for her. I would like also to thank my parents and my parents-in-law for their discrete and continuous support. This booklet is dedicated to all of them.

Affiliations of co-authors

Dr Antonio Agudo
Department of Epidemiology
Catalan Institute of Oncology (ICO)
Autovía de Castelldefels, Km 2,7
E-08907 Hospitalet de Llobregat
(Barcelona)
Spain

Dr Pilar Amiano
Dirección de Salud de Guipuzcoa
Av. de Navarra 4
E-20013 San Sebastian
Spain

Ms Caren Andrén
The Malmö Diet and Cancer Study
Dept. of Medicine, Lund University
Medical Research Center, Entrance 59
UMAS
205 02 Malmö
Sweden

Dr Paul Appleby
Cancer Research UK
Epidemiology Unit
Gibson Building, Radcliffe Infirmary
Oxford OX2 6HE
UK

Dr Christina Bamia
Dept. of Hygiene and Epidemiology
University of Athens Medical School
M. Asias 75
GR-Athens 11527
Greece

Dr A. Barcos
Service of Surveillance &
Epidemiological
Control
Institute of Public Health of Navarra
C/Leyre 15
20013 San Sebastian
Spain

Dr Göran Berglund
The Malmö Diet and Cancer Study
Dept. of Medicine, Lund University
Medical Research Center, Entrance 59
UMAS
205 02 Malmö
Sweden

Ms Manuela Bellegotti
Servizio di Epidemiologia
Istituto Nazionale dei Tumori
Via Venezian 1
I-20133 Milan,
Italy

Dr Erik Bergström
Epidemiology
Dept. of Public Health & Clinical
Medicine
University of Umeå
S-901 87 Umeå
Sweden

Dr Sheila Bingham
Deputy Director
MRC - Dunn Human Nutrition Unit
Hills Road
GB-Cambridge CB2 2DH
UK

Dr Heiner Boeing
Deutsches Institut für
Ernährungsforschung
Arthur Scheunert Allee 114 -116
D-14558 Bergholz-Rehbrücke,
Germany

Ms Birgit Brandstetter
Deutsches Institut für
Ernährungsforschung
Arthur Scheunert Allee 114 -116
D-14558 Bergholz-Rehbrücke,
Germany

Ms Magritt Brustad
Institute of Community Medicine
University of Tromsø
N- 9037 Tromsø
Norway

Dr Bas Bueno de Mesquita
Dept. of Epidemiology,
National Institute of Public Health
and the Environmental (RIVM)
P.O. Box 1
NL-3720 BA Bilthoven
Netherlands

Dr Marilyn Buzzard
Dept. of Preventive Medicine &
Community
Health
Medical College of Virgina
Virginia Commonwealth University
Box 980212
Richmond, VA 23298-0212
USA

Ms Corinne Casagrande
Unit of Nutrition and Cancer
International Agency for Research on
Cancer
150 cours Albert-Thomas
F69372 Lyon cedex 08
France

Ms Ruth U. Charrondière
(Formerly with the International Agency for Research on Cancer, Lyon, France)
Nutrition Assessment and Evaluation
FAO
Viale delle Terme di Caracalla
I-00100 Rome,
Italy

Dr Y. Chloptsios
Dept. of Hygiene and Epidemiology
University of Athens Medical School
M. Asias 75
GR-Athens 11527
Greece

Dr Françoise Clavel
INSERM, U. 521
Institut Gustave Roussy
Rue Camille Desmoulins
F-94805 Villejuif Cedex
France

Dr Gwyneth Davey
(Formerly with Cancer Research UK, Oxford)
Unit of Nutrition and Cancer
International Agency for Research on
Cancer
150 cours Albert-Thomas
F69372 Lyon cedex 08
France

Dr Nicholas E. Day
Institute of Public Health
Strangeways Research Laboratory
Wort's Causeway
GB-Cambridge CB1 8RN,
UK

Ms Geneviève Deharveng
Formerly with the International Agency for Research on Cancer
2, allée des Châtaigniers
Vieux Boisseuil
87220 Boisseuil
France

Ms Dagrun Engeset
Institute of Community Medicine
University of Tromsø
9037 Tromsø
Norway

Mr Michael Fahey
Unit of Nutrition and Cancer
International Agency for Research on Cancer
150 cours Albert-Thomas
F69372 Lyon cedex 08
France

Dr Pietro Ferrari
Unit of Nutrition and Cancer
International Agency for Research on Cancer
150 cours Albert-Thomas
F69372 Lyon cedex 08
France

Dr Carlos A. González
Department of Epidemiology
Catalan Institute of Oncology (ICO)
Autovía de Castelldefels, Km 2,7
E-08907 Hospitalet de Llobregat (Barcelona)
Spain

Dr Ingegerd Johansson
Dept. of Nutritional Research
University of Umeå
901 87 Umeå
Sweden

Dr Rudolf Kaaks
Leader, Hormones and Cancer Group
International Agency for Research on Cancer
150 cours Albert-Thomas
F69372 Lyon cedex 08
France

Dr Emmanuelle Kesse
INSERM, U. 521
Institut Gustave Roussy
Rue Camille Desmoulins
F-94805 Villejuif Cedex
France

Dr Timothy J.A. Key
Cancer Epidemiology Unit (ICRF)
Gibson Building
The Radcliffe Infirmary
GB-Oxford OX2 6HE
UK

Dr Anja Kroke
Deutsches Institut für Ernährungsforschung
Arthur Scheunert Allee 114 -116
D-14558 Bergholz-Rehbrücke,
Germany

Dr Areti Lagiou
Dept. of Hygiene and Epidemiology
University of Athens Medical School
M. Asias 75
GR-Athens 11527
Greece

Dr Nerea Larrañaga
Dirección de Salud de Guipuzcoa
Av. de Navarra 4
E-20013 San Sebastian
Spain

Ms C. Lauria
Registro Tumori Ragusa
Lega Italiana per la Lotta contro i Tumori –
Sezione di Ragusa
Via Dante 109
97100 Ragusa (Sicily)
Italy

Dr Jakob Linseisen
(Formerly with the German Cancer Research Center, Heidelberg)
Life Science Center Weihenstephan
TU München
Div. of Human Nutrition & Cancer Prevention
Alte Akademie 16
85350 Freising-Weihenstephan
Germany

Dr Gerhard Lotze
Division of Clinical Epidemiology
Deutsches Krebsforschungszentrum
Im Neuenheimer Feld 280
D-69120 Heidelberg 1,
Germany

Dr Eiliv Lund
Institute of Community Medicine
University of Tromsø
9037 Tromsø
Norway

Dr Irene Mattison
The Malmö Diet and Cancer Study
Dept. of Medicine, Lund University
Medical Research Center, Entrance 59
UMAS
205 02 Malmö
Sweden

Dr Anthony B. Miller
Division of Clinical Epidemiology
Deutsches Krebsforschungszentrum
Im Neuenheimer Feld 280
D-69120 Heidelberg 1,
Germany

Ms Angela Mulligan
Strangeways Research Laboratory
Institute of Public Health
University of Cambridge
Wort's Causeway
GB-Cambridge CB1 8RN
UK

Dr Marga Ocké
Dept. of Epidemiology
National Institute of Public Health and the Environment (RIVM)
P.O. Box 1
NL-3720 BA Bilthoven
Netherlands

Dr Kim Overvad
University of Aarhus
Institute of Epidemiology & Social Medicine
Vennelyst Boulevard 6
8000 Aarhus C
Denmark

Dr Valeria Pala
Servizio di Epidemiologia
Istituto Nazionale dei Tumori
Via Venezian 1
I-20133 Milan,
Italy

Dr Domenico Palli
Molecular & Nutritional Epidemiology Branch
Centro per lo Studio e la Prevenzione Oncologica (CSPO)
Via di San Salvi 12
I-50135 Florence
Italy

Dr Salvatore Panico
Dipartimento di Medicina Clinica e Sperimentale
Università di Napoli, Federico II
Via S. Pansini 5
80131 Naples
Italy

Dr Petra Peeters
Associate Professor of Epidemiology
Julius Center for Health Sciences & Primary Care
P.O. Box 85500
NL-3508 TA Utrecht
Netherlands

Dr Elio Riboli
Chief, Unit of Nutrition and Cancer
International Agency for Research on Cancer
150 cours Albert-Thomas
F69372 Lyon cedex 08
France

Ms Shirley Runswick
MRC - Dunn Human Nutrition Unit
Hills Road
GB-Cambridge CB2 2DH
UK

Dr Rodolfo Saracci
Unit of Nutrition and Cancer
International Agency for Research on Cancer
150 cours Albert-Thomas
F69372 Lyon cedex 08
France

Dr Mathias B. Schulze
Deutsches Institut für Ernährungsforschung
Arthur Scheunert Allee 114 -116
D-14558 Bergholz-Rehbrücke,
Germany

Dr Sabina Sieri
Servizio di Epidemiologia
Istituto Nazionale dei Tumori
Via Venezian 1
I-20133 Milan,
Italy

Ms Connie Stripp
Institute of Cancer Epidemiology
Danish Cancer Society
Strandboulevarden 49
Box 839
2100 Copenhagen Ø
Denmark

Dr Anne Tjønneland
Institute of Cancer Epidemiology
Danish Cancer Society
Strandboulevarden 49
Box 839
2100 Copenhagen Ø
Denmark

Dr Antonia Trichopoulou
Dept. of Hygiene and Epidemiology
University of Athens Medical School
M. Asias 75
GR-Athens 11527
Greece

Dr Dimitrios Trichopoulos
Dept. of Epidemiology
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115, USA
and
Dept. of Hygiene and Epidemiology
University of Athens Medical School
M. Asias 75
GR-Athens 11527
Greece

Dr Rosario Tumino
Registro Tumori Ragusa
Lega Italiana per la Lotta contro i
Tumori –
Sezione di Ragusa
Via Dante 109
97100 Ragusa (Sicily)
Italy

Dr Marti van Liere
(Formerly with the Institut Gustave Roussy, Villejuif, France)
Dept. of Health Care & Disease Control
Royal Tropical Institute (KIT)
Maurits kade 63
PO Box 95001
1090 HA Amsterdam
Netherlands

Dr Anne Linda van Kappel
(Formerly with the International Agency for Research on Cancer, Lyon, France)
Cryo-Bio System
29, rue Tronchet
75008 Paris
France

Dr Wija van Staveren
Professor in Nutrition & Gerontology
Dept. of Human Nutrition
Wageningen Agricultural University
PO Box 8129
6700 EV Wageningen
The Netherlands

Dr Fabrizio Veglia
Institute for Scientific Interchange
Foundation
Fondazione I.S.I.
Viale Settimio Severo 65
10133 Turin
Italy

Mrs Ailsa Welch
Institute of Public Health
Strangeways Research Laboratory
Wort's Causeway
GB-Cambridge CB1 8RN,
UK

Dr Anna Winkvist
Epidemiology
Dept. of Public Health & Clinical
Medicine
University of Umeå
S-901 87 Umeå
Sweden

Dr Elisabet Wirfält
Dept. of Medicine, Surgery &
Orthopaedics
Lund University
University Hospital in Malmö
20502 Malmö
Sweden

Curriculum Vitae

Nadia Slimani was born on 24 July 1962 in Toulon, France. After completing her secondary education at the French Lycée Louis Pasteur in Oran, Algeria (main subjects: physics, chemistry and biology) she obtained a degree as a clinical dietician at the Lycée Limayrac in Toulouse, France in 1985. In 1987 she completed a Masters degree in cellular biology and animal physiology at the University Blaise Pascal, Clermont-Ferrand, France. In 1988 she completed a postgraduate specialization in nutrition in developing countries (Diplôme d'Etude Supérieure Spécialisée) at the University of Languedoc-Roussillon, Montpellier, France. She then spent a 6-month training period, partly in Algeria and partly at the Agronomic Institute of Montpellier on "Weaning foods and weaning methods used in Algeria". Between 1989 and 1990 she obtained a special training award at the International Agency for Research on Cancer (IARC) to carry out methodological studies on dietary questionnaires and to prepare a review of past dietary consumption in collaboration with Drs Christine Friedenreich and Elio Riboli. Since 1990 she has worked as a nutritionist in the Unit of Nutrition and Cancer at IARC where has been involved in the development of the European Prospective Investigation into Cancer and Nutrition (EPIC). She was in charge of the international coordination of the dietary component of the EPIC study, and particularly the development of the calibration methodology which is the theme of her PhD thesis.

Printing: Y. Slimani

The research described in this thesis was carried out at the International Agency for Research on Cancer (Lyon France) as part of the EPIC project, supported by a grant from the "Europe Against Cancer" Programme of the European commission (SANCO)