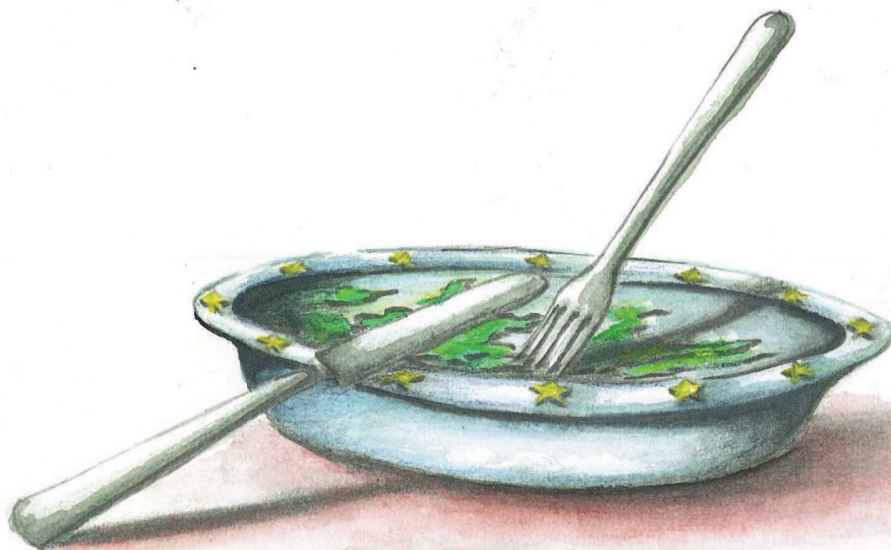# Evaluation of the two non-consecutive 24-h recall instrument

for pan-European food consumption surveys

Sandra P. Crispim

Evaluation of the two non-consecutive **24-h recall** instrument for pan-European food consumption surveys

Sandra P. Crispim

Thesis committee

Thesis supervisor
Prof. dr. ir. P. van't Veer
Professor of Nutrition and Epidemiology
Wageningen University, the Netherlands

Thesis co-supervisors
Dr. J.H.M. de Vries
Assistant professor, Division of Human Nutrition
Wageningen University, the Netherlands

Dr. ir. A. Geelen
Assistant professor, Division of Human Nutrition
Wageningen University, the Netherlands

Other members
Dr. ir. J.J.M. Castenmiller
Food and Consumer Product Safety Authority (VWA), The Hague, The Netherlands

Prof. dr. ir. FA van Eeuwijk
Wageningen University, Wageningen, The Netherlands

Dr. ir. C. van Rossum
National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

Prof. dr. L.A. Moreno
University of Zaragoza, Zaragoza, Spain

Evaluation of the two non-consecutive **24-h recall** instrument for pan-European food consumption surveys

Sandra P. Crispim

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 17 June 2011
at 11 a.m. in the Aula.

*To my beloved
parents Dilma and Joaquim
and husband Omar,
who have always being there for me*

ABSTRACT

Background: The comparability of food consumption data originating from national nutritional surveys in Europe is currently hampered because of different methodologies used. Therefore, experts in the European Food Consumption Survey Method (EFCOSUM) consortium proposed to use two non-consecutive 24-h recalls for standardised dietary monitoring in European countries.

Aim: Within the European Food Consumption Validation (EFCOVAL) consortium, this thesis aimed to evaluate the data collected with two non-consecutive 24-h recalls using EPIC-Soft for comparisons of dietary intake in adults between countries in future pan-European surveys.

Methods: To evaluate the bias in protein and potassium intake as well as the ranking of individuals according to their fish and fruit & vegetable intake collected with two non-consecutive 24-h recalls, we developed a validation study within EFCOVAL. The study included biomarker data of 600 subjects from five European centres in Belgium, the Czech Republic, France, the Netherlands, and Norway. To gain further insight into the determinants of the accuracy of the method by using multilevel analysis, we combined EFCOVAL data from one day with similar data from twelve other centres participating in the European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study. Then, we used the EFCOVAL data for assessing the impact of different modes of administration (telephone vs. face-to-face), recall days (1st vs. 2nd) and days of the week (weekdays vs. weekend) on the bias in protein and potassium intake. Finally, data from the Netherlands was used to explore the usefulness of collecting individual dietary data with the 24-h recalls for estimating dietary exposure to flavouring substances.

Results: On average, men and women underreported protein intake by 8% in the EFCOVAL study. Underreporting of potassium intake was 7% in men and 4% in women. The coefficient of variation of bias in observed protein and potassium intake between centres ranged from 4 to 7%. The prevalence of subjects with adequate protein and potassium intake according to the observed data at the lower and upper end of the usual intake distribution agreed fairly well (<10% difference) with the prevalence according to the excretion data. The results of the multilevel analysis indicated that the bias in observed protein intake for both genders and in potassium intake for women did not vary across centres and to a certain extent varied in potassium intake for men (coefficient of variation=9.5%). One of the factors mostly influencing the different performance of the method across European populations

was BMI. Furthermore, two standardised 24-h recalls and a food propensity questionnaire appeared to be appropriate to rank individuals according to their fish and fruit & vegetable intake in a comparable manner between the five European centres. Moreover, we observed that in some centres protein intake reported by face-to-face interviews at the study site was less accurate than by telephone interviews, and that second 24-h recall assessments were less accurate than first recalls. In addition, in one out of five centres, protein intake estimated during weekends and potassium intake estimated during weekdays were less accurate than during other days of the week. Finally, the collection of detailed food consumption data at the individual level may be necessary to assess the dietary exposure to flavourings and adaptations of the databases used in EPIC-Soft software can provide more detailed information on the dietary exposure to the flavouring raspberry ketone than non-modified databases.

Conclusion: Two non-consecutive 24-h recalls using EPIC-Soft provides sufficiently valid and suitable data for comparing dietary intake across European populations.

TABLE OF CONTENTS

Chapter 1

Introduction

nutritionalmonitoringsurveys
dietaryintakemethodsmeandi
stributionindividualdetailedd
ata24hordsp
opulatemeth
odologamper
edeurmptio
nsurveymethodepicefcosume
picsoftefsaeumenusefcovalhet
erogenouspopulationscomput
erisedchemicalscountryspecifi
cmeasurementerrorsystemati
crandomusualintakesbmivalid
itybiomarkerdesignflavouring

Background

Nutritional monitoring surveys have become an important topic on the public health agenda worldwide and are often part of health surveillance systems[1]. The aims of these surveys include recognizing nutritional problems of populations, tackling health goals, evaluating intervention programmes, and guiding the development of national and international policies[2; 3]. A main component of nutritional monitoring surveys is the assessment of dietary intake of populations and often the information needed from these surveys are the mean and distribution of intake. For instance, information about the group mean may be sufficient when comparing dietary intake of subgroups of a population. On the other hand, the distribution of usual intake is necessary to evaluate the inadequacy of dietary intake or to assess the proportion of the population exceeding the upper safe levels of specific components, as for example of chemical substances[4].

To monitor dietary intake of populations, aggregated data from national food production, as estimated using food balance sheets, or from household budget surveys can be used. These data are frequently used for assessing dietary exposure to chemicals. In addition, by using methods such as 24-h recalls and dietary records, dietary intake of populations can be estimated through assessment at an individual level. Individual data are necessary to provide information on the distribution of dietary intake among populations and is, therefore, the preferred option for this purpose[5].

Currently, nutritional surveys in Europe collecting dietary data of adult populations at the individual level use different methodologies (Table 1). While many countries are employing the 24-h recall method (e.g., Belgium and Poland), some apply dietary records (e.g., UK and Italy) or a food frequency questionnaire - FFQ (e.g., Norway), and others do not currently collect data at the individual level (e.g., Greece and Luxembourg). The level of detail and the quality of the data collected in those European surveys also differ greatly across countries[6; 7]. Because of the use of diverse methodologies, it is currently not possible to be certain about differences in dietary intake across European populations, such as that Poland is the leading country in fruit consumption (282 g/day) and that the UK consumes the least (95 g/day); as was presented in the last European Nutrition and Health Report[8]. Instead, it is very likely that estimated dietary intakes of countries are not comparable; thus hampering the evaluation of nutritional risks and development of policies under the World Health Organization (WHO) or the EU commission framework[9; 10].

Table 1 – Overview of European countries with the applied dietary assessment method, the number of replicates, and the year of performance of their most recent food consumption survey among adults*.

Country	Dietary method	Number of replicates	Year
Austria	24-h recall	1	2005-2006
Belgium	24-h recall	2 (non-consecutive)	2004-2005
Bulgaria	24-h recall	1	2004
Cyprus	No information	-	-
the Czech Republic	24-h recall	2 (non-consecutive)	2003-2004
Denmark	Pre-coded food diary with open fields	7 (consecutive)	2000-2002
Estonia	24-h recall	1	1997
Finland	48-h recall and dietary record	2† and 3 (consecutive)	2007
France	Dietary record	7 (consecutive)	2006-2007
Germany	24-h recall	2 (non-consecutive)	2005-2006
Greece	No national dietary survey	-	-
Hungary	Dietary record	3 (consecutive)	2003-2004
Iceland	24-h recall	1	2002
Ireland	Dietary record	7 (consecutive)	1997-1999
Italy	Dietary record	3 (consecutive)	2005-2006
Latvia	24-h recall	2 (unknown)	2008
Lithuania	24-h recall	unknown	2007
the Netherlands	24-h recall	2 (non-consecutive)	2003
Luxembourg	No national dietary survey	-	-
Malta	No national dietary survey	-	-
Norway	FFQ	-	1997
Poland	24-h recall	1	2000
Portugal	No national dietary survey	-	-
Romania	No information	-	-
Slovakia	24-h recall and FFQ	1 and 1	2006
Spain	Dietary record	3 (unknown)	2000
Slovenia	24-h recall	1	2007-2008
Sweden	Dietary record	7 (consecutive)	1997-1998
United Kingdom	Dietary record	4 (consecutive)	2008

* Table adapted from EFSA (2009)[11] and Le Donne et al. (2011)[12].
† 2 days estimated from one 48-h recall.

In this context, the European Food Consumption Survey Method (EFCOSUM) consortium worked towards the development of a method for a pan-European food consumption survey that could provide internationally comparable data on a set of policy-relevant nutritional indicators among European member states[13]. EFCOSUM was a consortium built on existing experience from European initiatives, such as the Data Food Networking (DAFNE), the European Prospective Investigation into Cancer and Nutrition (EPIC) project, and the European Cooperation in the field of Scientific and Technical Research (COST Action 99). The main EFCOSUM conclusions included the recommendation to use two non-consecutive 24-h recalls using EPIC-Soft software as the preferred method to assess dietary intake of adults in future pan-European food consumption surveys[14].

In February 2010, members of the European Food Safety Authority's (EFSA) Advisory Forum signed a declaration supporting the establishment of a pan-European food consumption survey - the EU-menu survey. This followed a number of initiatives within EFSA, including the recommendations of the 'Expert Group on Food Consumption Data' (EGFCD), which is an EFSA network with representatives from each EU Member state. The EGFCD expert group also recommended two non-consecutive 24-h recalls using EPIC-Soft software as the most appropriate method to be applied within adults in the EU-menu survey[11].

In the meantime, under the footsteps of EFCOSUM recommendations, the European Food Consumption Validation (EFCOVAL) consortium worked towards the further development and validation of the 24-h recall methodology using EPIC-Soft software for assessing foods, nutrients, and potentially hazardous substances in future pan-European food consumption surveys[15]. The present thesis was conducted in the framework of the EFCOVAL consortium.

Assessing dietary intake with 24-h recalls in food consumption surveys

The 24-h recall method consists of describing and quantifying the intake of foods and drinks during the period prior to the interview, which can be the previous 24-h or the preceding day, and thus provides information on actual dietary intake[2; 16; 17]. A major advantage of using 24-h recalls in nutritional monitoring surveys is that they are useful to compare heterogeneous populations with different ethnicity and literacy, especially when compared to FFQs[5]. This method also poses minimal burden on the respondent and does not affect the actual food intake as distinct from dietary records[11].

Additionally, the EGFCD expert group in EFSA recognised the 24-h recalls as the most cost effective method to be implemented within a pan-European food consumption survey[11].

Traditionally, 24-h recalls have been applied by trained interviewers via face-to-face or telephone interviews[18] but recent developments include web-based 24-h recalls that are self-administered[19; 20]. Furthermore, 24-h recalls may be obtained using computer software that prompts the interviewer to collect detailed description and quantification of foods, including for example structured questions about brand names and cooking methods[21]. Such details are of interest especially for the assessment of chemicals in the diet. Some successful examples of these computerised 24-h recalls are the Automated Multiple Pass Method (AMPM) used in nutritional monitoring surveys in the USA and Canada as well as EPIC-Soft software developed by the International Agency for Research on Cancer (IARC). Both methods use a cognitive approach to enhance complete and accurate food recalls[22-24].

EPIC-Soft, the chosen method to be applied within the EU-menu survey, was designed to standardize procedures of 24-h recalls within and between European populations[22]. For instance, the quantity of the food as finally consumed (e.g., raw or cooked; with or without inedible part) is calculated automatically by the software in whatever way the food is reported and by using country-specific conversion factors when necessary[22]. Up to now, EPIC-Soft is the only available software that was designed to provide standardised individual food consumption data in different European countries[25].

Measurement errors affecting 24-h recalls in future food consumption surveys

Even if the same method will be used to compare dietary intake of European populations in future, uncertainty remains about the quality of the data to be compared. As any other dietary assessment method, 24-h recalls have drawbacks. The accuracy of 24-h recalls is highly dependent upon participant's memory and the communication skills of both the participant and the interviewer[2; 11]. In addition, 24-h recall estimates may be influenced by characteristics of the subjects. For example, individuals with a high BMI appear to underreport their food intake consciously or unconsciously[26; 27]. Furthermore, it is not known whether sources of errors in the 24-h recall estimates are different across European populations because of their diverse heterogeneity in dietary patterns or differences in socioeconomic levels. Consequently,

several specific sources of error may exist in 24-h recalls and may differ between European countries during future food consumption surveys.

Errors are generally categorized in two types: systematic and random. Systematic errors reduce the accuracy of dietary intake resulting in an under- or overestimation of the mean intake of individuals[18], which may result in bias in the estimation at the population level. Random errors, on the other hand, lead to imprecise 24-h recall estimates[1]. They include day-to-day variation in dietary intake, and random errors in response or quantification. Random errors occur, for example, when a small number of 24-h recalls per subject is used. Unlike systematic errors, they do not influence the mean intake of populations because on average random errors cancel out[18]. Nevertheless, random errors contribute to the observed total variation. Consequently, observed intake distributions of populations become wider than the true usual intake distributions and result in a biased estimate of the prevalence of the population above or below a certain cut-off point.

Random errors can be reduced by increasing the number of 24-h recalls per subject. Also, when at least two days of 24-h recalls are used, a correction for random errors is possible. With the use of statistical methods, such as the Multiple-Source-Method (MSM)[28], the Iowa State University Foods (ISUF)[29] and the National Cancer Institute (NCI) method[30], the usual intake distributions can be modelled by removing the within-person random errors in the observed dietary intake[31]. The assessment of systematic errors requires a validation or evaluation study, as will be discussed in the next section.

Evaluating 24-h recalls for use in future food consumption surveys

Because systematic errors can be present in the assessment of 24-h recalls, it is still uncertain whether dietary intake in future surveys will be comparable across countries. For example, one can question whether a mean protein intake of 100 g/day in one country is indeed higher than an intake of 80 g/day in another country; although the same dietary assessment method was used. A validation or evaluation study is helpful to determine whether an observed difference in intake between countries is a true difference or not. The *validity* of any method, that is, that the method measures what it is intended to measure, can only be assessed by comparing it with an independent method of unquestionable accuracy[32]. Unfortunately, such unquestionable method does not exist because the collection of dietary data is always influenced by sources of

errors. Nevertheless, a useful approach to evaluate dietary methods involves the use of *biomarkers of intake.*

Biomarkers are not strictly considered to be measures of the true intake, but they provide reference estimates of dietary intake with errors that are unlikely to be correlated with the errors of self-reports of dietary intake[33]. In particular, *recovery-based biomarkers* have a precisely known quantitative relation to absolute daily intake and provide a reference estimate of the bias size in dietary intake[34]. However, only a few recovery biomarkers are available to assess nutrient intake. Examples are urinary nitrogen and potassium to assess the intake of protein and potassium, respectively[35; 36]. Another type of biomarker, known as *concentration biomarkers*, can also be used to evaluate dietary assessment methods. These markers include blood nutrient concentrations, such as serum carotenoids[37-39] and fatty acids in phospholipids[40-42]. Biomarkers determined as concentrations are the result of complex metabolic processes and are not as directly related to dietary intake as recovery biomarkers[43]. Concentration markers are especially useful to evaluate the ranking of individuals according to their intake by analysing their association with the observed dietary intake, but not for estimating the bias in the assessment of intake.

Rationale and outline of the thesis

EFCOVAL study

Since this thesis was conducted within the framework of EFCOVAL, a brief description of the study is given. EFCOVAL was carried out within the EU 6th framework Program and addressed several aspects that are related to the implementation of the 24-h recall using EPIC-Soft as the instrument for future pan-European food consumption surveys.

The three main objectives of the EFCOVAL consortium were[15]:

1) To upgrade, adapt, and validate the 24-h recall method using EPIC-Soft.

2) To expand the applicability of the upgraded software program to younger age groups and for use in chemical exposure assessment.

3) To improve the methodology and statistical methods that translate short-term dietary intake information to usual intake estimates.

Within the first objective of EFCOVAL, a validation study was set-up in five centres from Belgium, the Czech Republic, France, the Netherlands and Norway. The results presented in this thesis are mainly based on this validation study. This study offered a novel opportunity to compare the quality of the assessment of dietary intake data between European populations. Especially, the usual intake distributions of European populations could be modelled and compared since two days of dietary intake and recovery biomarker data were collected.

Aim of the thesis

The overall aim of this thesis was to evaluate the data obtained with the duplicate 24-h recall method using EPIC-Soft for comparisons of dietary intake in adults between countries in future pan-European surveys.

In view of the validation of the method, Chapters 2 to 5 are part of objective 1 of the EFCOVAL consortium. Chapter 2 discusses the validity of two non-consecutive days of the 24-h recall to compare protein and potassium intake between five European centres using recovery biomarkers, i.e., urinary nitrogen and potassium. Chapter 3 gives further insight into the determinants of the accuracy of the method to estimate protein and potassium intake by using a linear multilevel analysis. This was also a recovery-based biomarker evaluation of dietary intake, in which EFCOVAL data from one day was pooled with similar data from nine other European centres participating in the EPIC calibration study. Furthermore, the assessment of usual fish and fruit & vegetable intake collected with the 24-h recall method was evaluated using concentration biomarkers, respectively fatty acids in phospholipids and serum carotenoids (Chapter 4). This chapter, which focused on the assessment of ranking of individuals according to their food group intake, aimed to give additional evidence about the quality of food group intake data collected with the 24-h recalls. Furthermore, since different aspects of the design of 24-h recall assessments may be used across countries in a future pan-European food consumption survey, the impact of different modes of administration (face-to-face vs. telephone interviews), recall days (1st vs. 2nd), days of the week (weekday vs. weekend) and interview days (1 or 2 days later) on the bias in protein and potassium intake was evaluated in Chapter 5. Chapter 6 of this thesis was part of the second objective in EFCOVAL. In this chapter, an explorative study was developed to investigate the usefulness of dietary data collected with the 24-h recalls for estimating chemical exposure assessment. For

this, the flavouring substances were the chemical category used as an example. Finally, Chapter 7 discusses the main findings and conclusions of this thesis in the perspective of a future pan-European food consumption survey.

References

1. Sahyoun NR (2003) Nutritional Surveillance: Developed Countries. In *Encyclopedia of Food Sciences and Nutrition*, pp. 6000 [BF Caballero, P.; Trugo, L., editor].
2. Thompson FE & Byers T (1994) Dietary assessment resource manual. *J Nutr* 124, 2245S-2317S.
3. Byers T (1998) Nutrition monitoring and surveillance. In *Nutritional epidemiology*, 2nd ed., pp. 347-356 [W Willett, editor]. New York etc.: Oxford University Press.
4. Sabry JH (1988) Purposes of food consumption studies. In *Manual on methodology for food consumption studies*, pp. 25-31 [ME Cameron and WA Van Staveren, editors]. Oxford: Oxford University Press.
5. Biro G, Hulshof KF, Ovesen L *et al.* (2002) Selection of methodology to assess food intake. *Eur J Clin Nutr* 56 Suppl 2, S25-32.
6. Verger P, Ireland J, Moller A *et al.* (2002) Improvement of comparability of dietary intake assessment using currently available individual food consumption surveys. *Eur J Clin Nutr* 56 Suppl 2, S18-24.
7. Ireland J, van Erp-Baart AM, Charrondiere UR *et al.* (2002) Selection of a food classification system and a food composition database for future food consumption surveys. *Eur J Clin Nutr* 56 Suppl 2, S33-45.
8. Elmadfa I (2009) European Nutrition and Health Report. *Forum of Nutrition*, available at http://www.univie.ac.at/enhr/downloads/enhrii_book.pdf
9. WHO (2007) *The challenge of obesity in the WHO European Region and the strategies for response*. Denmark: World Health Organization.
10. EFSA (2011) Use of the EFSA Comprehensive European Food Consumption Database in Exposure Assessment. *The EFSA Journal* 9, 1-34.
11. EFSA (2009) General principles for the collection of national food consumption data in the view of a pan-European dietary survey. *The EFSA Journal* 7, 1-51.
12. Le Donne C, Piccinelli R, Sette S *et al.* (2011) Overview of existing European food consumption databases: critical aspects in relation to their use for the assessment of dietary exposure to additives, flavourings and residues of food contact materials. *Int J Food Sci Nutr* 62, 121-132.
13. Brussaard JH, Johansson L & Kearney J (2002) Rationale and methods of the EFCOSUM project. *Eur J Clin Nutr* 56 Suppl 2, S4-7.
14. Brussaard JH, Lowik MR, Steingrimsdottir L *et al.* (2002) A European food consumption survey method--conclusions and recommendations. *Eur J Clin Nutr* 56 Suppl 2, S89-94.
15. de Boer E (2011) Rationale and methods of the EFCOVAL project. *Eur J Clin Nutr (in press)*.
16. Burke BS & Stuart HC (1938) A method of diet analysis: Application in research and pediatric practice. *The Journal of Pediatrics* 12, 493-503.
17. Wiehl DG (1942) Diets of a Group of Aircraft Workers in Southern California. *Milbank Memorial Fund Quartely* 20, 329-356.
18. Gibson RS (2005) *Principles of nutritional assessment*. 2nd ed. New York etc.: Oxford University Press.
19. Arab L, Wesseling-Perry K, Jardack P *et al.* (2010) Eight self-administered 24-hour dietary recalls using the Internet are feasible in African Americans and Whites: the energetics study. *J Am Diet Assoc* 110, 857-864.
20. Subar AF, Crafts J, Zimmerman TP *et al.* (2010) Assessment of the accuracy of portion size reports using computer-based food photographs aids in the development of an automated self-administered 24-hour recall. *J Am Diet Assoc* 110, 55-64.
21. Buzzard M (1998) 24-Hour dietary recall and food record methods. In *Nutritional epidemiology*, 2nd ed., pp. 50-73 [W Willett, editor]. New York etc.: Oxford University Press.
22. Slimani N, Deharveng G, Charrondiere RU *et al.* (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Methods Programs Biomed* 58, 251-266.
23. Slimani N, Ferrari P, Ocke M *et al.* (2000) 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* 54, 900-917.
24. Moshfegh AJ, Rhodes DG, Baer DJ *et al.* (2008) The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. *Am J Clin Nutr* 88, 324-332.
25. Slimani N, Casagrande C, Deharveng G *et al.* (2011) The standardized computerized 24-hour dietary recall method EPIC-Soft adapted for pan-European dietary monitoring. *Eur J Clin Nutr (in press)*.

26. Muhlheim LS, Allison DB, Heshka S *et al.* (1998) Do unsuccessful dieters intentionally underreport food intake? *Int J Eat Disord* 24, 259-266.
27. Lissner L, Troiano RP, Midthune D *et al.* (2007) OPEN about obesity: recovery biomarkers, dietary reporting errors and BMI. *Int J Obes (Lond)* 31, 956-961.
28. German Institute of Human Nutrition (2009) The Multiple Source Method (MSM). https://nugo.dife.de/msm/
29. Nusser SM, Carriquiry AL, Dodd KW *et al.* (1996) A Semiparametric Transformation Approach to Estimating Usual Daily Intake Distributions. *J Amer Statistical Assoc* 91, 1440-1449.
30. Tooze JA, Midthune D, Dodd KW *et al.* (2006) A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc* 106, 1575-1587.
31. EFSA (2010) Statistical modelling of usual intake. Scientific report. EFSA-Q-2009-00841. 1-23.
32. van Staveren WA & Burema J (1985) Food consumption surveys: frustations and expectations. *Naringsforskning* 29, 38-42.
33. Ocke MC & Kaaks RJ (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, 1240S-1245S.
34. Kaaks RJ (1997) Biochemical markers as additional measurements in studies of the accuracy of dietary questionnaire measurements: conceptual issues. *Am J Clin Nutr* 65, 1232S-1239S.
35. Bingham SA (2003) Urine nitrogen as a biomarker for the validation of dietary protein intake. *J Nutr* 133 Suppl 3, 921S-924S.
36. Tasevska N, Runswick SA & Bingham SA (2006) Urinary potassium is as reliable as urinary nitrogen for use as a recovery biomarker in dietary studies of free living individuals. *J Nutr* 136, 1334-1340.
37. Macdonald HM, Hardcastle AC, Duthie GG *et al.* (2009) Changes in vitamin biomarkers during a 2-year intervention trial involving increased fruit and vegetable consumption by free-living volunteers. *Br J Nutr* 102, 1477-1486.
38. Polsinelli ML, Rock CL, Henderson SA *et al.* (1998) Plasma carotenoids as biomarkers of fruit and vegetable servings in women. *J Am Diet Assoc* 98, 194-196.
39. Yeum KJ, Booth SL, Sadowski JA *et al.* (1996) Human plasma carotenoid response to the ingestion of controlled diets high in fruits and vegetables. *Am J Clin Nutr* 64, 594-602.
40. Katan MB, Deslypere JP, van Birgelen AP *et al.* (1997) Kinetics of the incorporation of dietary fatty acids into serum cholesteryl esters, erythrocyte membranes, and adipose tissue: an 18-month controlled study. *J Lipid Res* 38, 2012-2022.
41. Leaf DA, Connor WE, Barstad L *et al.* (1995) Incorporation of dietary n-3 fatty acids into the fatty acids of human adipose tissue and plasma lipid classes. *Am J Clin Nutr* 62, 68-73.
42. Zock PL, Mensink RP, Harryvan J *et al.* (1997) Fatty acids in serum cholesteryl esters as quantitative biomarkers of dietary intake in humans. *Am J Epidemiol* 145, 1114-1122.
43. Freedman LS, Kipnis V, Schatzkin A *et al.* (2010) Can we use biomarkers in combination with self-reports to strengthen the analysis of nutritional epidemiologic studies? *Epidemiol Perspect Innov* 7, 2.

Chapter 2

Two non-consecutive 24-h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres: results from the EFCOVAL study

Sandra P. Crispim, Jeanne H. M. de Vries, Anouk Geelen,
Olga W. Souverein, Paul J. M. Hulshof, Lionel Lafay,
Anne-Sophie Rousseau, Inger T. L. Lillegaard, Lene F. Andersen,
Inge Huybrechts, Willem De Keyzer, Jiri Ruprich, Marcela Dofkova,
Marga C. Ocke, Evelien de Boer, Nadia Slimani and Pieter van't Veer

Abstract

The use of two non-consecutive 24-h recalls using EPIC-Soft for standardised dietary monitoring in European countries has previously been proposed in the European Food Consumption Survey Method consortium. Whether this methodology is sufficiently valid to assess nutrient intake in a comparable way, among populations with different food patterns in Europe, is the subject of study in the European Food Consumption Validation consortium. The objective of the study was to compare the validity of usual protein and potassium intake estimated from two non-consecutive standardised 24-h recalls using EPIC-Soft between five selected centres in Europe. A total of 600 adults, aged 45–65 years, were recruited in Belgium, the Czech Republic, France, the Netherlands and Norway. From each participant, two 24-h recalls and two 24-h urines were collected. The mean and distribution of usual protein and potassium intake, as well as the ranking of intake, were compared with protein and potassium excretions within and between centres. Underestimation of protein (range 2–13 %) and potassium (range 4–17 %) intake was seen in all centres, except in the Czech Republic. We found a fair agreement between prevalences estimated based on the intake and excretion data at the lower end of the usual intake distribution (< 10 % difference), but larger differences at other points. Protein and potassium intake was moderately correlated with excretion within the centres (ranges = 0.39–0.67 and 0.37-0.69, respectively). These were comparable across centres. In conclusion, two standardised 24-h recalls (EPIC-Soft) appear to be sufficiently valid for assessing and comparing the mean and distribution of protein and potassium intake across five centres in Europe as well as for ranking individuals.

Introduction

National food consumption surveys aim to provide information on the mean and distribution of food and nutrient intakes of the population and related subgroups, in order to develop and evaluate nutrition policies. In addition, national food consumption surveys are essential to provide data for risk assessment work, as conducted by the European Food Safety Authority – EFSA[1]. In Europe, food consumption data originating from national surveys are not always comparable because they differ in a number of aspects, such as the choice of the dietary assessment method and the reference period of the data collection[2-4]. Furthermore, some countries do not have national food consumption surveys in place[4].

The European Food Consumption Survey Method (EFCOSUM) consortium has acknowledged the need for policy-relevant dietary indicators that are comparable among European countries, which could contribute to the establishment of a Community Health Monitoring System[5]. They recommended two non-consecutive days of 24-h recall using EPIC-Soft software (Lyon, Rhone Alpes, France) as the preferred method to assess the dietary intake in future pan-European monitoring surveys in adults. In addition, they specified total fat, saturated fatty acids and ethanol as the components of most relevance in this assessment[6-8].

The 24-h recall is a commonly used dietary assessment method in food consumption surveys in Europe[4] and is also being used in surveys in the USA[9], Canada[10], Australia[11], and New Zealand[12]. A major advantage of using 24-h recalls in (inter)national surveys is that the method is useful for comparison of heterogeneous populations with different ethnicity and literacy[6]. In addition, a computerised version of 24-h recalls seems to be the best means of standardising and controlling for sources of error attributable to 24-h recall interviews[6; 13]. Nevertheless, computerised 24-h recalls need to be tailor-made to every included country and/or study, e.g., by adaptations of the food and recipe list. Therefore, whether this methodology performs in a comparable way across countries with different food consumption patterns in Europe deserves further exploration, as validity of the 24-h recall depends on both the characteristics of the method and the study population.

Biological markers offer an important opportunity to evaluate the dietary assessment methods since errors are likely to be truly independent between the measurements of biomarker and dietary intake[14]. Urinary nitrogen and potassium are two of the few available recovery biomarkers to assess the nutrient intakes[15; 16]. With the use of these two biomarkers, a single 24-h recall using EPIC-Soft has been previously validated for assessing the group mean intakes of protein of twelve centres in six countries within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort study[17]. Yet, the accuracy of this methodology needs to be determined when aiming at estimating usual dietary intake among different European populations by collecting two independent 24-h recalls. Hence, following the path of EFCOSUM, the European Food Consumption Validation (EFCOVAL) consortium aimed to further develop and validate a European food consumption method using EPIC-Soft software for assessing the food and nutrient intakes within European countries, and for comparisons between them. In the present paper, we aim to compare the validity of usual protein and potassium intake estimated from two non-consecutive standardised 24-h recalls using EPIC-Soft between five selected centres in Europe. This was done

by addressing the bias present in the estimation of each centre's mean and distribution of intake, as well as the ranking of individuals within and between centres according to their intake.

Subjects and Methods

Subjects

Data were collected in five European countries: Belgium, the Czech Republic, France (Southern part), the Netherlands, and Norway. These countries were selected to represent a large variety in food patterns across Europe. Data were collected in the South of France to include the characteristics of the Mediterranean diet. A food pattern from Central/Eastern Europe was represented by the Czech Republic, from the Scandinavian countries by Norway and from the western part of Europe by Belgium and the Netherlands. Another reason for their selection was their experience in performing nutrition monitoring surveys. The present study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving human subjects were approved by ethical committees in each centre involved in the data collection.

We recruited subjects by convenience sampling through advertisements (newspaper and websites), mailing lists, among others. Recruitment of institutionalised subjects was not allowed, nor included more than one member of a household. Subjects were informed about the study through information meetings at the institutions/universities in the Czech Republic, France, and the Netherlands, and by phone, letter and personally in Belgium and Norway. At these occasions, a screening questionnaire was filled in to confirm the subjects' eligibility in the study. Subsequently, the eligible participants gave written informed consent, and appointments for later visits were scheduled. Exclusion criteria were currently taking diuretics, following prescribed dietary therapy, being enrolled in another study in the same period, not being able to read or speak the national language, being pregnant, lactating, having diabetes mellitus or kidney disease, and donating blood or plasma during or less than four weeks before the study. *para*-Aminobenzoic acid (PABA) was used to check the completeness of urine collections; therefore, subjects hypersensitive to PABA or taking antibiotics containing sulphonamides, which are PABA-antagonistic, were not eligible for the study.

Taking into account an anticipated dropout percentage of 20% and aiming at a net sample of 50 per stratum, a total of 60 men and 60 women were recruited per centre (*n* 600). The age range of subjects was 45 to 65 years, which was chosen to limit the heterogeneity of the sample. Furthermore, we aimed to include at least ten men and ten women in each of the three predetermined categories of education level (low, intermediate and high) per centre. We used country-specific classifications to define each category level.

We excluded one subject because no data for recall and biomarker collected on the same day were available. Therefore, the study population comprised 599 subjects (296 men and 303 women).

Study design

Wageningen University (the Netherlands) was, as the coordinating centre, responsible for the overall logistics of the validation study in the EFCOVAL consortium. For standardisation, all study procedures, i.e., on recruitment and fieldwork conditions, data processing formats, quality-control aspects and specimen collection, storage and transport details, were described in protocols. The recruitment of subjects and data collection in the Netherlands were performed from April to July 2007, six months before the other four centres, in order to test all the procedures of the fieldwork beforehand and to be able to refine the protocols. The other centres started the fieldwork in October or November 2007 with the last centre finalising the collection by April 2008.

At the beginning of the study, subjects had their body weight and height measured in the study centres. Then, a 24-h recall and a 24-h urine collection were obtained covering the same reference day. Subjects were aware of the days of data collection but not of the purpose of the interviews. The second recall and urine collection were obtained at least one month after the first one.

Dietary data

The two 24-h recalls were collected using two modes of administration: one by phone and one face-to-face at the centre since it is likely that future food consumption monitoring surveys will be conducted in both ways across European countries. The order of the two modes of administration was randomly allocated among the subjects.

Furthermore, the appointments for the dietary recalls followed a randomised schedule, which included all days of the week. This randomisation allowed the same person to have the same recalled weekday for both interviews by chance. Interviewers in each centre were nutritionists or dietitians who were trained in interviewing skills and working with EPIC-Soft in the context of the validation study. They were guided by qualified local trainers who were previously trained by staff from the Wageningen coordination centre and the National Institute for Public Health and the Environment in the Netherlands (RIVM). Interviewers were aware of the objectives of the study. The centres were allowed to organise their data collection in the same way they would do in a future performance of their nutritional surveillance system. An example is that interviewees were permitted to check food packages and household measures in their home for more detailed information during the phone interview while this was not possible during the face-to-face interview at the study centre. Another example is that dietary recalls in Belgium, the Czech Republic and the Netherlands were not conducted on Sundays. Therefore, Saturdays' intake was recalled two days later, on Mondays.

The two 24-h recalls were collected using EPIC-Soft (version 9.16). The structure and standardisation procedure of EPIC-Soft have been described elsewhere[18; 19]. Briefly, EPIC-Soft is a computer-assisted 24-h dietary recall that follows standardised steps when describing, quantifying, probing and calculating the food intakes[18]. All the participating countries had an existing version of EPIC-Soft available, except the Czech Republic for which a new country-specific version was developed. In addition, EPIC-Soft databases were adapted for each centre in terms of some common specifications for the EFCOVAL study (e.g., soups were treated as recipes rather than food items). Furthermore, the centres generated or updated a list of the single food items and recipes expected to be consumed by their participants. Modifications of such lists were needed afterwards based on notes made during the interview. The methods of estimation of portion size included household measures, weight/volume, standard units and portions, bread shapes and photographs. The set of photographs was developed in the context of the EPIC study[20]. Each centre chose from the EPIC portfolio of photographs the pictures that best represented their national food habits.

In the absence of harmonised recent food composition tables (FCT) including all countries of our assessment, protein and potassium contents in foods were calculated using country-specific FCT[21-24]. Carbohydrates, total fat, saturated fat, alcohol and dietary fibre intake as well as energy content were also computed. We calculated energy values by summing the contributions from protein, carbohydrates, fat and

alcohol and using related Atwater factors (17, 17, 37 and 29 kJ per gram, respectively). In the Czech Republic, the national FCT was published about 20 years ago. Therefore, a FCT was compiled for EFCOVAL purposes in the Czech Republic with composition of most foods based on the Slovakian tables[25]. In all the centres, missing nutrient data for a food was imputed from a similar food or another FCT, based on country-specific decisions, but in a few cases, this was not possible for potassium, saturated fat, dietary fibre and alcohol. The percentage of missing values was less than 6% of all reported foods for all nutrients.

Twenty-four hour urine collections and recovery biomarkers

The subjects were instructed not to make use of acetaminophen painkillers, such as paracetamol, and sulphonamide drugs, during the days of urine collection. To check the completeness of urinary collections, one tablet of 80 mg PABA (PABAcheck, Laboratories for Applied Biology, London, UK) had to be taken three times on the day of the urine collection: with the morning, midday and evening meals. Hence, we expected that 240 mg of PABA would be almost completely excreted within 24-h[26; 27]. The collection of the 24-h urine started with voiding and discarding the first urine in the morning after waking up. Subsequently, the urine excreted during the next 24-h, up to and including the first voiding of the following day, was collected. For this purpose, each subject received labelled containers (at least two), one funnel to help the collection, one safety pin to be fixed in the underwear as a reminder for collection and a diary scheme booklet to register the timing, observations (e.g., use of medication and supplements) and possible deviations (e.g., missing urine) of the urine collection protocol. Boric acid (3g/2 litre bottle) was used as preservative. The subjects provided their urine samples to the dietitians at the study centre when a face-to-face dietary recall was scheduled. If the 24-h recall interview was by phone, urine samples were collected at the subject's home or delivered to the study centre. When a long period was anticipated between the end of the collection and the receiving of samples, subjects were instructed to keep the urine samples at approximately 4°C, which in most cases was not more than 12-h. To verify the stability of PABA in urine, a pooled urine sample of three participants from the Netherlands were kept at four different temperatures (-20, 6, 20 and 30°C) for 8 days. At five moments (days 0, 1, 2, 4 and 7), PABA concentrations were measured. No significant changes in PABA concentrations were observed during the storage period at each temperature. The regression equation for PABA content as a function of time during storage at 20°C

(assumed to be the most common storage temperature) was as follows: PABA (mg/L) = 140.2 - 0.8 (time in days) with the 95% confidence interval for the time coefficient being [-2.5,0.8].

At the laboratory of the local centres, urine was mixed, weighed and aliquoted. Then, the specimens were stored at -20°C until shipment on dry ice to the central laboratory at Wageningen University, where they were kept at the same temperature.

Chemical analysis

On the day of chemical analysis, aliquots were rapidly thawed at room temperature. Urinary nitrogen was determined colorimetrically by the Kjeldahl technique on a Kjeltec 2300 analyser (Foss, Hilleroed, Denmark) after destruction of the sample with concentrated sulphuric acid. Urinary potassium was measured by an ion-selective electrode on a Beckman Synchron LX20 analyser (Beckman Coulter, Mijdrecht, The Netherlands). PABA was measured by colorimetry[28]. The intra-assay precision, expressed as coefficient of variation (CV), of these three analyses was less than 2%. Taking into account the extra-renal losses (approximately 19%) and the fact that protein on average contains 16% of nitrogen, urinary protein was calculated as [6.25 x (urinary nitrogen/0.81)][15; 29]. Urinary potassium was estimated by dividing the measured value by 0.77, assuming that 77% of potassium intake is excreted through the urine when considering faecal excretion[16; 30].

Urine samples with PABA recoveries below 50% were treated as incomplete and excluded from the data analysis (*n* 14). Additionally, the subjects who took drugs containing sulphonamides or acetaminophen, or one who took less than three PABA tablets had their urine diaries checked for other deviations in the urine collection. In cases where other deviations were observed, namely urine loss during the collection or absent registration of collection time, samples were excluded from the analysis (*n* 4). Otherwise, samples were included (*n* 13) as we did not want to exclude potentially complete urines. Results of the present paper did not change by excluding these subjects. As described before[31], specimens containing between 50 and 85% of PABA recovery (*n* 105) had their urinary concentrations proportionally adjusted to 93% of PABA recovery. Recoveries above 85% were included in data analyses without adjustments (*n* 1062).

Data analysis

The analyses were performed using SAS statistical package, version 9.1 (SAS Institute Inc, Cary, NC, USA). The statistical analyses were stratified by sex and using the average of two days of intake and excretion, except for 18 subjects who only had one day of 24-h recall and biomarker. For these subjects, the 24-h recall matched with the day of the urine collection. To assess the presence of bias (systematic errors), the mean difference between nutrient intake and excretion was calculated. Analysis of covariance (ANCOVA) followed by the Tukey post hoc test was used for testing whether biases differed between the centres. The ANCOVA model included age (continuous), education level (three categories) and body mass index (BMI continuous), given that stratified analysis of these variables showed us differential performance of the method within and between the centres. To estimate and compare the distribution of usual intake and excretion of protein and potassium between the centres, the Multiple Source Method (MSM) was used as the measurement error model[32]. This model removes the effect of day-to-day variability and random error in the two 24-h recalls and biomarker estimates. The MSM was developed in the framework of the EFCOVAL study and enabled us to estimate individual usual intake. We decided not to use covariates in the calculation of usual intakes with the MSM. Plots of usual intake distributions based on the 24-h recall and biomarker were created using R software, version 2.8.1 (http://CRAN.R-project.org). The percentages of subjects consuming above certain cut-off points for each distribution curve were calculated. For both sexes, we specified eleven cut-off points to cover the whole range of protein and potassium intake among the five centres. For the evaluation of ranking of individuals, we computed Pearson's correlation coefficients. For adjusted correlations, we used usual intake and excretion data corrected for within-person variability, as estimated by the MSM method, and further corrected for age, BMI and education level by using partial Pearson correlations. Confidence intervals of the correlations were obtained using the Fisher Z-transformation[33]. Energy-adjusted correlations were calculated using the residual method[34]. To test the equality of correlations, pairwise comparisons were made using Fisher Z-transformation[33]. Pooled correlations of the five centres were calculated by first converting the correlations into a standard normal metric (Fisher's r-to-Z transformation). Next, the pooled average was calculated, in which each transformed correlation coefficient was weighted by its inverse variance, followed by the back transformation[33]. The cochrane Q test was used for testing the heterogeneity of the pooled correlation[35].

Results

The mean age of the subjects was similar in the five centres (**Table 1**). In both sexes, mean BMI was comparable across the centres (ranges: 23.2-25.5 kg/m² in women and 25.5-27.9 kg/m² in men). Subjects with moderate and high education levels were over-represented in the study compared with individuals with a low education level, especially men in Norway. The variations in energy intake across the centres were less pronounced than in macronutrients, especially for carbohydrates.

A degree of underestimation was seen in the assessment of protein intake in all the centres. Underestimation varied from 2.7% (Norway) to 12.4% (the Netherlands) in men and from 2.3% (Norway) to 12.8% (France) in women, based on the crude differences between intake and excretion (**Table 2**). After adjusting for age, BMI and education level, the bias did not differ between the centres for women. However, men in the Czech Republic had a significantly smaller bias compared with those in France and the Netherlands. For potassium, the underestimation varied from 1.7% in Norway to 17.1% in France for men and from 6.6% in the Netherlands to 13% in France for women. An overestimation of 5.9% for men and 1.6% for women was found in the Czech Republic. A statistically significant difference in the adjusted bias was seen in men between France and three other centres: the Czech Republic, the Netherlands and Norway. In women, differences were statistically significant only between France and the Czech Republic. BMI was the only factor influencing the differences between the countries at a significant level (p<0.01 for all analyses, except for potassium in women; p=0.16).

Table 1 – Characteristics of five European centres in the EFCOVAL validation study[*]

| | Men | | | | | | | | | | Women | | | | | | | | | |
| | BE[†] (*n* 63) | | CZ (*n* 58) | | FR (*n* 54) | | NL (*n* 59) | | NO (*n* 62) | | BE (*n* 60) | | CZ (*n* 60) | | FR (*n* 59) | | NL (*n* 62) | | NO (*n* 62) | |
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Age (years)	54	5.5	55	6.9	56	5.4	57	4.3	55	6.0	55	5.0	55	6.1	55	6.0	55	5.6	54	6.0
Weight (kg)	81.1	13.3	85.7	13.2	78.1	9.7	83.8	14.4	85.7	9.9	67.6	12.5	66.8	9.8	60.6	8.6	71.4	13.8	68.4	11.4
Height (cm)	175.6	7.1	175.4	6.4	174.8	7.0	177.5	8.8	179.9	7.2	163.6	6.8	163.8	6.1	161.6	6.7	167.6	8.8	166.0	6.8
BMI (kg/m²)	27.2	3.6	27.9	4.2	25.5	2.7	26.5	3.8	26.4	2.5	25.2	4.2	25.0	3.9	23.2	3.0	25.5	5.0	24.8	3.7
Energy (MJ/d)	11.0	0.3	12.1	0.5	10.4	0.3	11.2	0.4	11.8	0.4	8.4	0.3	8.4	0.2	8.1	0.2	8.6	0.3	8.4	0.3
Energy %																				
protein	16.0	0.4	14.5	0.3	15.9	0.4	15.8	0.4	17.2	0.5	16.1	0.4	14.8	0.4	16.0	0.3	15.4	0.4	17.9	0.5
total fat	35.2	0.8	34.7	0.8	35.8	0.8	34.1	0.8	36.0	1.1	33.8	0.8	34.0	1.0	39.3	0.9	34.6	0.9	38.6	1.0
carbohydrates	41.6	0.9	47.0	1.1	44.0	1.0	43.1	1.0	42.8	1.1	44.8	1.0	49.1	1.1	42.4	1.0	46.0	0.9	40.0	1.1
saturated fat	13.7	0.4	12.7	0.3	13.7	0.4	13.0	0.4	13.9	0.6	13.7	0.4	12.8	0.4	14.0	0.5	12.5	0.4	14.8	0.5
Alcohol (g/d)	30.2	4.2	17.8	3.4	15.1	2.5	27.6	3.4	16.5	2.8	17.3	2.7	6.3	1.3	6.9	1.3	12.3	2.0	10.7	2.1
Fibre (g/MJ/d)	2.3	0.1	2.5	0.1	2.2	0.1	2.4	0.1	2.5	0.1	2.7	0.1	3.1	0.1	2.7	0.1	3.0	0.1	2.7	0.1
Education % of total																				
Low	15.9		20.7		25.9		20.3		3.2		16.7		16.6		35.6		24.2		16.1	
Intermediate	23.8		24.1		24.1		20.3		30.7		25.0		46.7		27.1		40.3		19.4	
High	60.3		55.2		50.0		59.4		66.1		58.3		36.7		37.3		35.5		64.5	

[*] Dietary intake based on 2x24-h recalls

[†]BE=Belgium, CZ=the Czech Republic, FR=France, NL=the Netherlands, NO=Norway.

Table 2 – Protein and potassium intake and excretion (Mean ± SE) based on 2x24-h recalls and 2x24-h urinary biomarkers for five European centres in the EFCOVAL validation study

| | Men | | | | | | | | | | |
	BE* (*n* 63)		CZ (*n* 58)		FR (*n* 54)		NL (*n* 59)		NO (*n* 62)		p-value†
Protein (g)	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	
Intake	101.7	3.3	100.4	4.2	95.9	3.4	101.5	3.5	115.2	3.8	
Excretion‡	110.8	3.2	104.1	3.0	109.1	2.8	115.9	3.6	118.4	3.1	
% crude difference	-8.2		-3.5		-12.1		-12.4		-2.7		
Adjusted difference	-7.5[ab]	3.4	-1.4[a]	3.6	-14.7[b]	3.6	-14.1[b]	3.6	-2.3[ab]	3.6	0.02
Potassium (mg)											
Intake	4024	131	3726	164	3464	138	4326	139	4847	182	
Excretion§	4301	148	3517	143	4180	141	4491	157	4935	138	
% crude difference	-6.4		+5.9		-17.1		-3.7		-1.7		
Adjusted difference	-230[ab]	144	282[a]	150	-759[b]	153	-123[a]	150	-66[a]	151	<0.01

| | Women | | | | | | | | | | |
	BE (*n* 60)		CZ (*n* 60)		FR (*n* 59)		NL (*n* 62)		NO (*n* 62)		p-value
Protein (g)	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	
Intake	79.0	2.5	70.8	2.1	74.7	1.9	78.2	3.3	85.5	2.6	
Excretion	87.5	2.6	78.8	2.2	85.7	2.0	85.1	2.9	87.5	2.1	
% crude difference	-9.7		-2.7		-12.8		-8.2		-2.3		
Adjusted difference	-7.9	2.5	-7.9	2.5	-12.2	2.5	-6.3	2.4	-1.8	2.5	0.07
Potassium (mg)											
Intake	3513	148	3155	143	3146	141	3618	157	3630	138	
Excretion	3928	138	3150	111	3617	124	3871	142	3899	102	
% crude difference	-10.5		+1.6		-13.0		-6.6		-6.9		
Adjusted difference	-414[ab]	115	9[a]	113	-503[b]	114	-224[ab]	110	-274[ab]	114	0.02

*BE=Belgium, CZ=the Czech Republic, FR=France, NL=the Netherlands, NO=Norway.
†One-way ANCOVA (General Linear Model) based on mean difference between intake and excretion. Tukey's post-hoc test was used for pair-wise comparison between the countries. ANCOVA model included age, BMI, and educational level. Different letter superscripts correspond to differences between countries at p<0.05.
‡Urinary protein=(urinary nitrogen/0.81)x6.25.[15]
§Urinary potassium=(urinary potassium/0.77).[16]

Upon inclusion of energy intake into the ANCOVA model, the conclusion about the differences between the centres changed only for protein results in men, which lost statistical significance (p=0.08). Additionally, when we pooled the data from all the countries, no consistent trend in mean protein and potassium biases was observed across the different education levels and modes of administration (data not shown).

The bias in mean intake can also be observed when comparing the distributions of usual intake based on food consumption data with those obtained from excretion data (**Figures 1 to 4**). The intake data curve shifted somewhat to the left (underestimation of intake) for almost all the centres compared with the excretion data. Since the prevalence of subjects consuming below or above a certain cut-off point is an important indicator for a population's nutritional status, we assessed and compared the prevalence of subjects consuming above specific cut-off points for both usual intake and usual excretion distributions (**Figures 5 to 8**).

Overall, we found a fair agreement between prevalences estimated based on the intake and excretion data at the lower end of the usual protein and potassium intake distribution, but larger differences at middle cut-off levels. For protein in men, the smallest differences in prevalence between intake and excretion were seen in Norway (up to 15%) and the largest ones in France (up to 46%) and the Netherlands (up to 41%). For women, the smallest differences were seen in Norway (up to 11%) and the largest ones in the Czech Republic (up to 38%) and France (up to 55%). The smallest difference between potassium intake and excretion distribution in males was observed in the Netherlands (up to 7%) while the larger differences were seen in the Czech Republic and France (up to 21 and 40%, respectively). In women, France was the centre with the largest difference (up to 29%) between potassium usual intake and excretion, and the Netherlands the smallest (up to 17%).

Unadjusted Pearson correlation coefficients between average protein intake and its biomarker within centres ranged between 0.42 and 0.65 in men and between 0.46 and 0.57 in women (**Table 3**). After adjusting for within person variability, age, BMI and education level, correlations ranged between 0.43 and 0.67 in men and between 0.39 and 0.63 in women. For potassium, unadjusted correlations ranged between 0.45 and 0.65 in men and between 0.31 and 0.69 in women. Adjusted correlations ranged between 0.40 and 0.69 in men and between 0.37 and 0.68 in women. For both protein and potassium, adjusting only for the within-person variability slightly increased the correlations between intake and excretion (data not shown). Statistically significant differences between correlation coefficients were only found between Belgium and the Czech Republic (p=0.04) for unadjusted correlations of potassium in women.
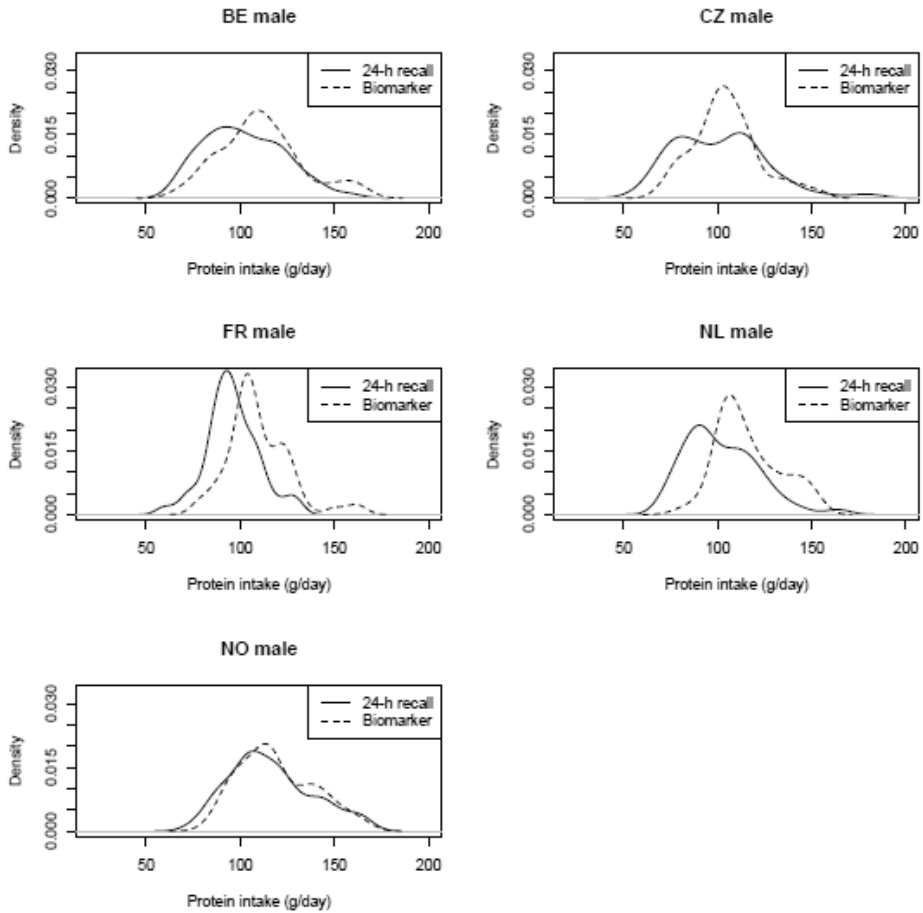
Figure 1 – Estimated distribution of usual protein intake in men, based on 2x24-h dietary recall and biomarker for five European centres in the EFCOVAL validation study

BE=Belgium, CZ=the Czech Republic, FR=France, NL=the Netherlands, NO=Norway
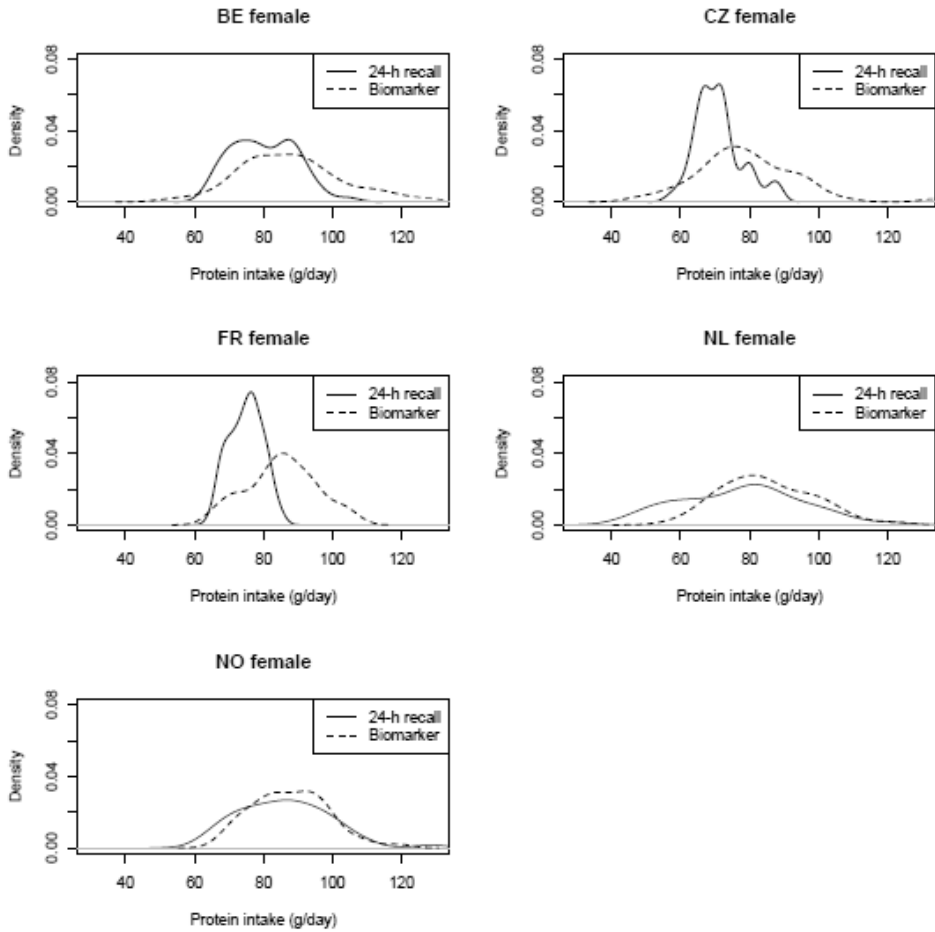
Figure 2 – Estimated distribution of usual protein intake in women, based on the 2x24-h dietary recall and biomarker for five European centres in the EFCOVAL validation study

BE=Belgium, CZ=the Czech Republic, FR=France, NL=the Netherlands, NO=Norway

Figure 3 – Estimated distribution of usual potassium intake in men, based on the 2x24-h dietary recall and biomarker for five European centres in the EFCOVAL validation study
BE=Belgium, CZ=the Czech Republic, FR=France, NL=the Netherlands, NO=Norway
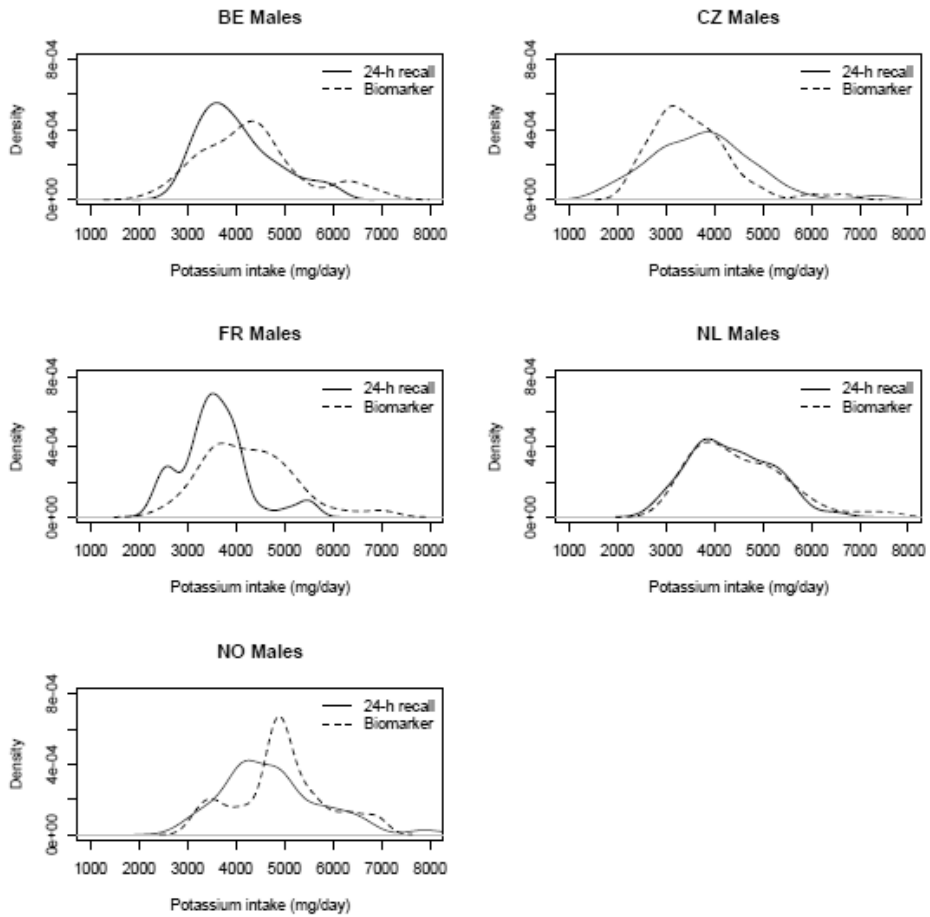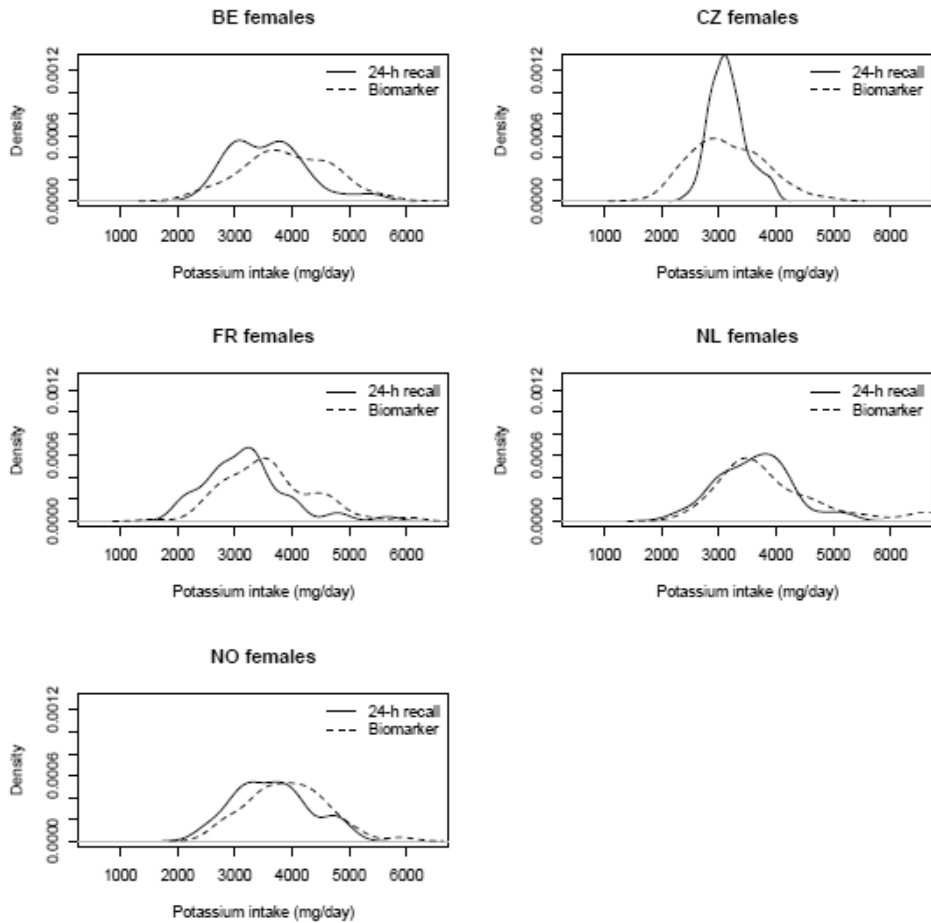
Figure 4 – Estimated distribution of usual potassium intake in women, based on the 2x24-h dietary recall and biomarker for five European centres in the EFCOVAL validation study

BE=Belgium, CZ=the Czech Republic, FR=France, NL=the Netherlands, NO=Norway

Belgium

the Czech Republic

the Netherlands

Norway

France
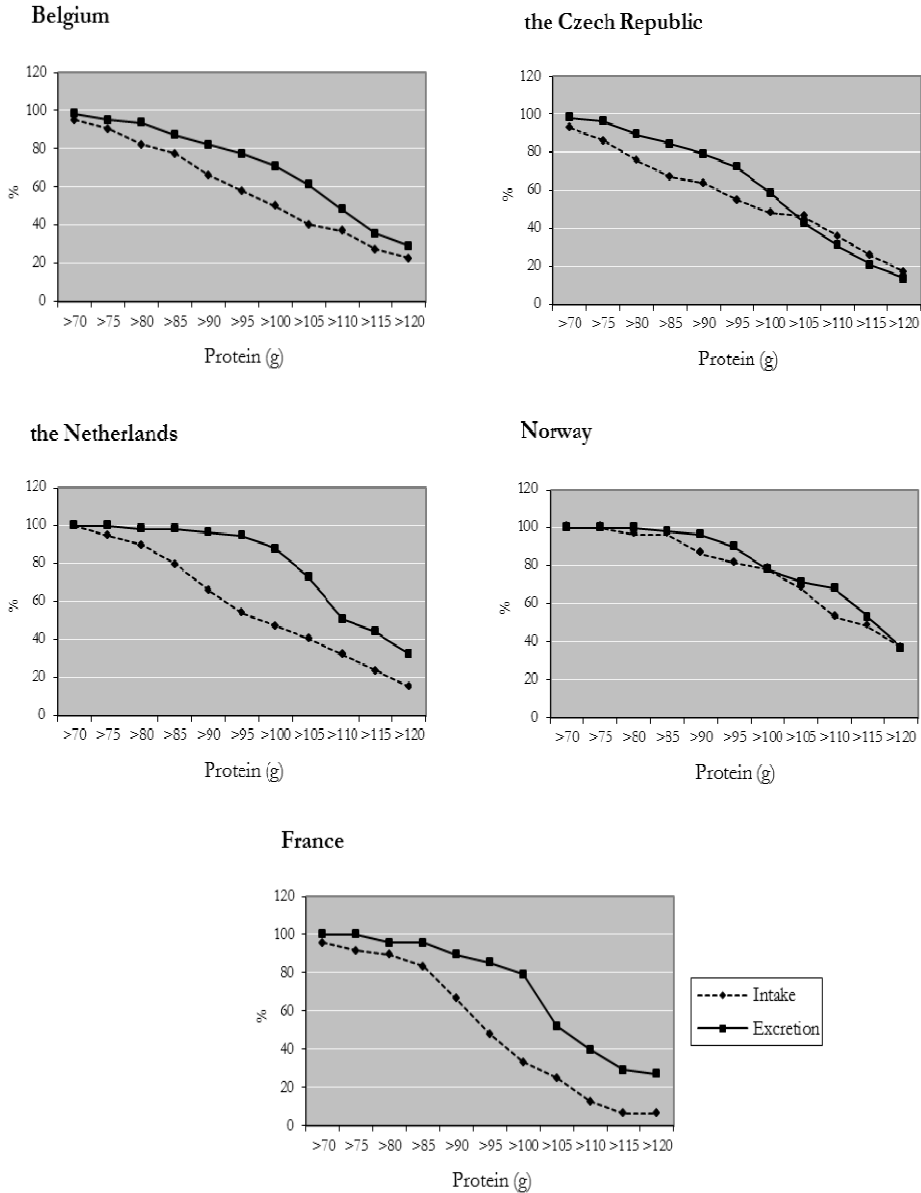
Figure 5 – Prevalence of men consuming above specific amounts of protein as estimated by usual intake distributions* from 24-h recalls (intake) and biomarkers (excretion) for five European centres in the EFCOVAL validation study

---

* Usual intake/excretion distribution estimated by MSM method.

Belgium

the Czech Republic

the Netherlands

Norway

France
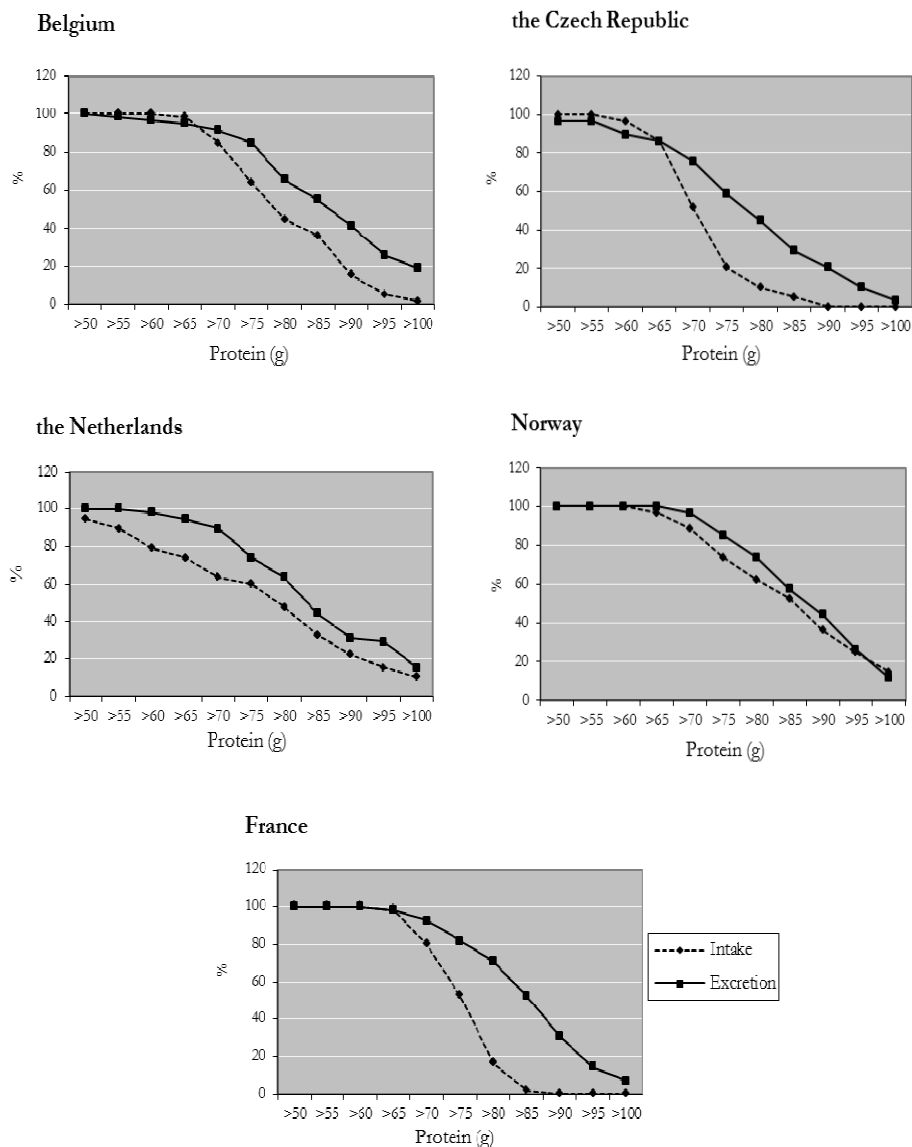
Figure 6 – Prevalence of women consuming above specific amounts of protein as estimated by usual intake distributions* from 24-h recalls (intake) and biomarkers (excretion) for five European centres in the EFCOVAL validation study

---

* Usual intake/excretion distribution estimated by MSM method.

Belgium

Czech Republic

the Netherlands

Norway

France

Figure 7 – Prevalence of men consuming above specific amounts of potassium as estimated by usual intake distributions* from 24-h recalls (intake) and biomarkers (excretion) for five European centres in the EFCOVAL validation study

---

* Usual intake/excretion distribution estimated by MSM method.

Figure 8 – Prevalence of women consuming above specific amounts of potassium as estimated by usual intake distributions* from 24-h recalls (intake) and biomarkers (excretion) for five European centres in the EFCOVAL validation study
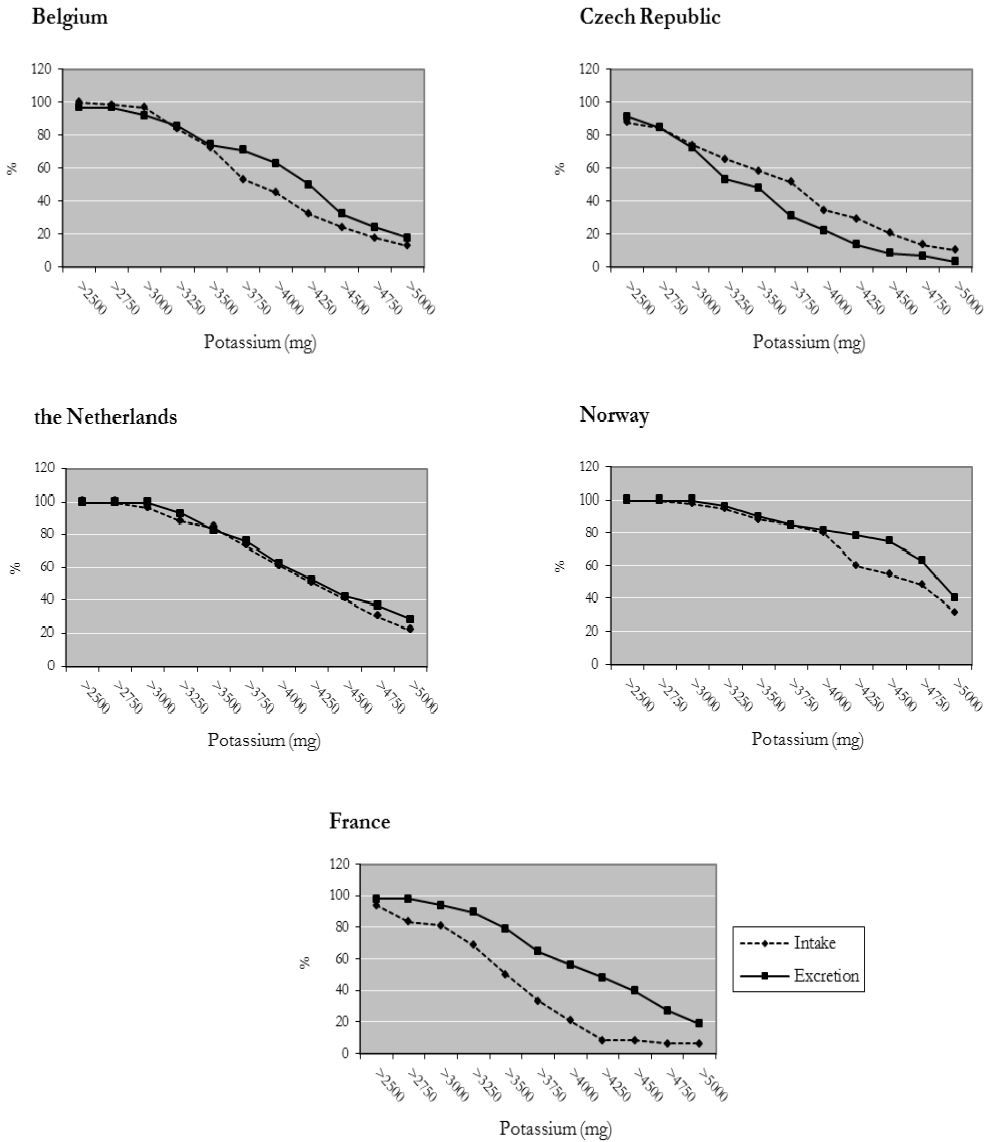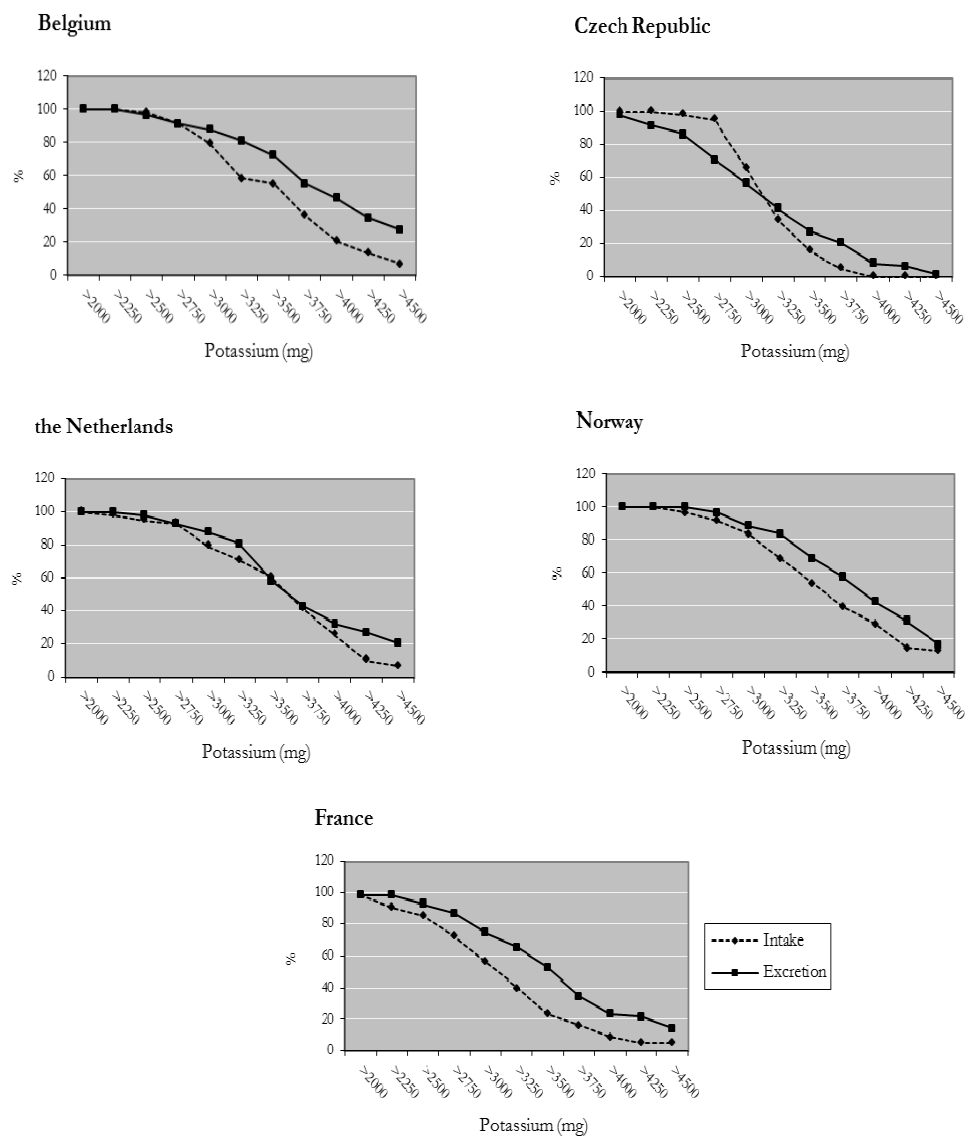
---

* Usual intake/excretion distribution estimated by MSM method.

Table 3 – Pearson coefficients of correlation (confidence interval) between protein intake and urinary excretion[*] for five European centres in the EFCOVAL validation study[†]

		PROTEIN INTAKE						
		Men				Women		
Centres	*n*	*Unadjusted*	*Adjusted[‡]*	*Energy-Adjusted[§]*	*n*	*Unadjusted*	*Adjusted*	*Energy-Adjusted*
Belgium	58	0.48 (0.27,0.65)	0.49 (0.27,0.67)	0.48 (0.26, 0.66)	62	0.57 (0.37,0.72)	0.57 (0.35,0.72)	0.35 (0.13,0.59)
the Czech Republic	58	0.50 (0.28,0.67)	0.43 (0.18,0.62)	0.25 (-0.01, 0.49)	58	0.56 (0.35,0.71)	0.57 (0.35,0.72)	0.49 (0.29,0.69)
France	55	0.65 (0.46,0.78)	0.67 (0.47,0.81)	0.65 (0.44, 0.79)	48	0.46 (0.23,0.64)	0.39 (0.13,0.60)	0.51 (0.27,0.69)
the Netherlands	58	0.42 (0.18,0.61)	0.51 (0.29,0.68)	0.47 (0.24, 0.65)	59	0.51 (0.29,0.67)	0.63 (0.44,0.77)	0.34 (0.15,0.60)
Norway	61	0.52 (0.32,0.69)	0.47 (0.24,0.65)	0.50 (0.27, 0.67)	60	0.53 (0.33,0.69)	0.52 (0.30,0.68)	0.41 (0.20,0.62)
Pooled[**]	290	0.52 (0.40, 0.63)	0.51 (0.39,0.63)	0.50 (0.38, 0.62)	287	0.53 (0.41,0.64)	0.60 (0.42,0.66)	0.45 (0.33,0.57)

		POTASSIUM INTAKE						
		Men				Women		
Centres	*n*	*Unadjusted*	*Adjusted*	*Energy-Adjusted*	*n*	*Unadjusted*	*Adjusted*	*Energy-Adjusted*
Belgium	58	0.54 (0.33,0.69)	0.53 (0.32, 0.69)	0.42 (0.18, 0.61)	62	0.69 (0.53,0.81)	0.68 (0.51,0..80)	0.60 (0.40, 0.75)
the Czech Republic	58	0.45 (0.21,0.63)	0.40 (0.15, 0.60)	0.37 (0.12, 0.58)	58	0.31 (0.01,0.52)	0.37 (0.12-0.58)	0.36 (0.11, 0.57)
France	55	0.62 (0.42,0.76)	0.64 (0.42, 0.78)	0.63 (0.42, 0.78)	48	0.61 (0.42,0.75)	0.63 (0.43, 0.77)	0.62 (0.41, 0.76)
the Netherlands	58	0.65 (0.47,0.76)	0.69 (0.52, 0.80)	0.66 (0.48, 0.79)	59	0.61 (0.42,0.74)	0.60 (0.40, 0.75)	0.36 (0.10, 0.57)
Norway	61	0.50 (0.28,0.67)	0.50 (0.28, 0.68)	0.62 (0.43, 0.76)	60	0.49 (0.28,0.66)	0.51 (0.29, 0.68)	0.49 (0.26, 0.66)
Pooled	290	0.55 (0.44,0.62)	0.56 (0.44, 0.68)	0.56 (0.44, 0.68)	287	0.55 (0.44,0.67)	0.57 (0.45, 0.68)	0.51 (0.40, 0.63)

[*] Average intake and excretion based on two days of collection

[†] Pairwise comparisons between countries (by Fisher Z transformation) suggested differences for: unadjusted correlations between Belgium and the Czech Republic in females; and between France and the Czech Republic for energy-adjusted correlations in males.

[‡] Adjusted for the within person variability using the usual intake/excretion data as estimated by MSM method (see methods section); and adjusted for age, BMI, and educational level using Partial Pearson correlations.

[§] Same adjustments as previous correlation plus energy-adjustment by residual method.

[**] P-value for heterogeneity were not significant for all the analyses (p>0.05).

However, after adjusting the correlations for energy, we found a significant difference between the Czech Republic (r=0.25) and France (r=0.65) for protein intake in men (p= 0.01). The pooled adjusted correlations in males and females were 0.51 and 0.60 for protein and 0.56 and 0.57 for potassium intake, respectively.

Discussion

In the present study, we compared the validity of usual protein and potassium intake estimated from two non-consecutive standardised 24-h recalls between five selected centres in Europe. On average, men and women underreported protein intake from the two 24-h recalls by 8%. For potassium intake, average underestimation was 7% for men and 4% for women.

Protein intake was markedly underestimated (~12%) in French and Dutch men, especially when compared with Czech Republic men. The same is true for potassium intake in French men. In women, underestimation of mean protein intake was present in all the centres and appeared to be comparable across the centres. For potassium intake, however, the underestimation observed in the French centre was not comparable to that of the other centres, particularly to the overestimation observed in the Czech Republic. Furthermore, we assessed the agreement between the percentage of subjects above a certain cut-off point based on 24-h recall and biomarker data. We found a fair agreement for cut-off points at the lower end of the distribution (less than 10% difference), but larger differences at other points of the intake distribution (up to 55% difference for protein in French females). Finally, we observed moderate correlations for the ranking of individuals, which were likely to be comparable across the centres.

The results from the EPIC study, using EPIC-Soft in different centres, revealed a similar or even higher underestimation of protein intake collected from a single day (average of 13% in men and 19% in women)[17]. The OPEN study in the United States, which assessed the structure of dietary measurement error in 24-h recalls collected twice, has also shown a similar underestimation of protein intake (11-15%)[36]. A few other studies indicated overestimation of protein (about 7% for the whole population)[37]. For potassium, studies indicated overestimation of intake up to 20%[38-40], similar to what we observed in the Czech Republic. Nevertheless, because of methodological differences, the comparison of bias estimates between the present study and other studies is not straightforward. For example, adjustment of nitrogen and potassium excretions to extra-renal losses was not consistently performed among

the studies. In addition, the completeness of 24-h urine collections was not always assessed. Although we acknowledge the differences in methodology between the studies, the performance of these two standardised 24-h recalls on assessing the mean protein and potassium intake appeared to provide alike or even more accurate results than what have been presented in the literature so far.

In terms of assessing the whole distribution of intake, two 24-h recalls used in the study by Freedman et al.[40] underestimated the usual protein intake in all points of the distribution, especially at the lower end. Moreover, they found a good agreement between potassium intake and excretion in the whole range of percentiles. In contrast, moderate to large discrepancies were found between 24-h recall and biomarker data distributions in the present study, but not at the lower end of the distribution. The present results suggest that the assessment of protein and potassium inadequacy at the population level by two non-consecutive 24-recalls in healthy European populations is, therefore, appropriate.

Independent of the size of the bias, the correct classification of individuals according to their intake is also informative on the quality of the dietary assessment. The correlations in the present paper are considerably higher compared with many other studies[36; 41-43]. Based on this, we conclude that the method performed sufficiently for the ranking of individuals, adding evidence to the use of this standardised 24-h recall. When we adjusted the nutrient values for energy intake, this changed the correlations in both directions and resulted in more noticeable differences across the centres. We doubt, however, whether energy-adjusted values will be our main exposure of interest in future monitoring surveys and whether individual energy intake was correctly estimated using only two days of 24-recall. Therefore, we do not base the conclusions of the present paper on the energy-adjusted results.

We suppose that the differences found in the size and direction of the bias (i.e., overestimation of potassium intake in the Czech Republic and underestimation of both potassium and protein in the other centres) between the centres may be explained by reasons related to characteristics of the population and of the method itself. We have controlled our statistical analyses for the influence of age, education level and BMI. As a result, BMI was the only factor significantly influencing the differences between the countries. This is in accordance with our expectations since other studies have revealed a differential under-reporting of dietary intake by subgroups of BMI[40; 44]. Nevertheless, other aspects of the population could have affected the validity of the method between the centres in a different manner, i.e., factors related to the food pattern of the centres. Due to cultural differences in food

pattern, it is expected that predominant food items contributing to protein and potassium intake across European countries will be different[45; 46]. For example, the food group 'dairy products' was one of the major contributors (>22%) to the protein intake in the Netherlands and Norway (in males only) whereas in the other three centres 'meat products' was distinctly the major contributor (>30%). Knowing that the errors in the assessment of different food groups differ, as for instance in the portion size estimation[47], differences in validity between the centres could be expected. Likewise, differences in the consumption of composite foods could have had an effect since it is more difficult to recall all ingredients of composite foods than a single food item[48; 49].

Another important factor that could explain the differences between countries is the use of not harmonize FCT across the centres. Use of different conversion factors as well as of distinct laboratory analyses to produce food nutrient contents across the tables are just some examples which could have caused biases not to be comparable. For instance, for three of the FCT used in EFCOVAL, protein figures were calculated from nitrogen contents using the so-called 'Jones conversion factors'[50] or slight modifications of them. However, in the Dutch tables only two of these factors were used (6.38 for milk products and 6.25 for all other foods) and in the compiled Czech table only one factor (6.25) was applied (Slovakian tables). Since errors attributed to these differences can be proportional to the level of intake, it is impossible to conclude on the influence of using different conversion factors in the comparison between the countries. Nevertheless, further investigation about the use of these conversion factors in FCT for comparisons of nutrient intake between countries is warranted.

The present study adds value to the current knowledge of collecting dietary information using standardised 24-h recalls for possible use in national monitoring surveys. An important strength of the present study was the collection of two days of both dietary intake and biomarkers allowing the quantification of the within-person variability and to estimate the usual intake distributions. A potential limitation of the present study is that a health-conscious sample may have been included, hampering the extrapolation of the results to the general population. However, the present results suggested that extrapolation to other populations could be done irrespective of their education level. In addition, the generalisability of protein and potassium results to other nutrients of interest should be done with care. Although we might want to assume that the validation results of a single nutrient can be used as a proxy to other nutrients, there is evidence nowadays that some foods and consequently related

nutrients might be selectively misreported[47; 51]. Besides, only two days of 24-h recall were used in our assessment while the inclusion of more than two days may be necessary to improve the use of this 24-h recall in the assessment of other nutrient intake distributions, particularly the infrequently consumed ones[52]. The statistical adjustments performed with the MSM intended to remove the day-to-day variation in intakes and assess the usual distributions of intake. But, if the variance of the nutrient intake is not reliably estimated from two days of intake, then the observed intake may shrink too much or too little toward the group mean intake, resulting in an inaccurate usual intake distribution[53]. The use of food frequency questionnaires combined with 24-h recalls may be an option in future monitoring surveys for the calculation of usual intakes of infrequently consumed nutrients, as more days of 24-h recalls are demanding and expensive. Furthermore, the reliability of the conversion factors used to adjust urinary protein and potassium in our analyses can be questioned. With the assumption that subjects were in nitrogen balance, these factors have been based on rigorously controlled feeding studies[15; 16] and in the case of protein confirmed by Kipnis et al.[54]. Lastly, we have collected data in the Netherlands six months before the other centres and this may have influenced the results. Nevertheless, while the data for the Netherlands were collected in spring/summer, the data for the other countries were collected in the winter/spring. However, since minor adjustments were done in the study protocols, and differences in the seasonality were small for protein and potassium intakes, it is unlikely that a different period influences the present results.

To conclude, first, the ability of the two non-consecutive standardised 24-h recalls using EPIC-Soft software appears to be sufficiently valid for assessing and comparing the mean protein and potassium intake across the centres. When comparing populations in a future nutrition monitoring system, the variability in the nutrient biases of 4-7% across the centres needs to be considered. Second, the method seems to be sufficiently valid for assessing and comparing the protein and potassium inadequacy of healthy populations across the centres, and less appropriate to assess other points of the intake distribution. Third, the ability to rank the individuals according to protein and potassium intakes within the centres is comparable between them, which substantiates the validity of the method. Therefore, this standardised two non-consecutive 24-h recalls, further adapted and validated in the EFCOVAL project, appears appropriate to be used in the context of a future pan-European dietary monitoring system. Built on EFCOVAL and EPIC experiences, improvements may be possible for the employment of this methodology by an even higher standardisation setting (e.g., conversion factors), which could result in an enhanced validity of the method, and thus comparability between the countries.

Acknowledgments

References

1. EFSA (2009) General principles for the collection of national food consumption data in the view of a pan-European dietary survey. *The EFSA Journal* 7, 1-51.
2. Charzewska J (1994) Gaps in dietary-survey methodology in eastern Europe. *Am J Clin Nutr* 59, 157S-160S.
3. Pietinen P & Ovaskainen ML (1994) Gaps in dietary-survey methodology in Western Europe. *Am J Clin Nutr* 59, 161S-163S.
4. Verger P, Ireland J, Moller A *et al.* (2002) Improvement of comparability of dietary intake assessment using currently available individual food consumption surveys. *Eur J Clin Nutr* 56 Suppl 2, S18-24.
5. Brussaard JH, Johansson L & Kearney J (2002) Rationale and methods of the EFCOSUM project. *Eur J Clin Nutr* 56 Suppl 2, S4-7.
6. Biro G, Hulshof KF, Ovesen L *et al.* (2002) Selection of methodology to assess food intake. *Eur J Clin Nutr* 56 Suppl 2, S25-32.
7. Brussaard JH, Lowik MR, Steingrimsdottir L *et al.* (2002) A European food consumption survey method--conclusions and recommendations. *Eur J Clin Nutr* 56 Suppl 2, S89-94.
8. Slimani N & Valsta L (2002) Perspectives of using the EPIC-SOFT programme in the context of pan-European nutritional monitoring surveys: methodological and practical implications. *Eur J Clin Nutr* 56 Suppl 2, S63-74.
9. Conway JM, Ingwersen LA & Moshfegh AJ (2004) Accuracy of dietary recall using the USDA five-step multiple-pass method in men: An observational validation study. *J Am Diet Assoc* 104, 595-603.
10. Statistics Canada (2004) *Canadian Community Health Survey, Cycle 2.2, Nutrition. A guide to accessing and interpreting the data*. Ottawa: Minister of Health.
11. McLennan W & Podger A (1995) *National Nutrition Survey User's Guide*. Canberra: Australian Goverment Printing Service.
12. Russell D, Parnell W & Wilson N (1999) *NZ Food: NZ People. Key results of the 1997 National Nutrition Survey*. Wellington: Minister of Health.
13. Willett W (1998) *Nutritional epidemiology*. 2nd ed. New York etc.: Oxford University Press.
14. Ocke MC & Kaaks RJ (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, 1240S-1245S.
15. Bingham SA & Cummings JH (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.
16. Tasevska N, Runswick SA & Bingham SA (2006) Urinary potassium is as reliable as urinary nitrogen for use as a recovery biomarker in dietary studies of free living individuals. *J Nutr* 136, 1334-1340.
17. Slimani N, Bingham S, Runswick S *et al.* (2003) Group level validation of protein intakes estimated by 24-hour diet recall and dietary questionnaires against 24-hour urinary nitrogen in the European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study. *Cancer Epidemiol Biomarkers Prev* 12, 784-795.
18. Slimani N, Deharveng G, Charrondiere RU *et al.* (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Methods Programs Biomed* 58, 251-266.
19. Slimani N, Ferrari P, Ocke M *et al.* (2000) 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* 54, 900-917.
20. van Kappel AL, Amoyel J, Slimani N *et al.* (1995) *EPIC-Soft picture book for estimation of food portion sizes*. Lyon, France: International Agency for Research on Cancer.
21. NUBEL (2004) *Belgische Voedingsmiddelentabel (Nutrients Belgium)*. Brussel VZW Nubel.
22. NEVO-TABEL (2006) *Nederlandse Voedings Middelen Tabel (Dutch Food Composition Tables)*. Den Haag: Voedingscentrum.
23. The Norwegian Food Safety Authority, the Norwegian Directorate of Health & and the University of Oslo (2006) *The Norwegian Food Composition Table*. Oslo; available at www.matvaretabellen.no/matvaretabellen

24. AFSSA/CIQUAL (2008) *French Food Composition Table* (*Belgium Food Composition Table*). Paris: French Food Safety Agency.
25. Food Research Institute (1997-2002) *Pozivatinove tabulky (Food Tables)*. Bratislava: Food Research Institute.
26. Bingham S & Cummings JH (1983) The use of 4-aminobenzoic acid as a marker to validate the completeness of 24 h urine collections in man. *Clin Sci (Lond)* 64, 629-635.
27. Runswick S, Slothouber B, Boeing H *et al.* (2002) Compliance with the urine marker PABAcheck in cancer epidemiology studies. *IARC Sci Publ* 156, 35-37.
28. Bingham SA, Williams R, Cole TJ *et al.* (1988) Reference values for analytes of 24-h urine collections known to be complete. *Ann Clin Biochem* 25 (Pt 6), 610-619.
29. Bingham SA (2003) Urine nitrogen as a biomarker for the validation of dietary protein intake. *J Nutr* 133 Suppl 3, 921S-924S.
30. Holbrook J, Patterson K, Bodner J *et al.* (1984) Sodium and potassium intake and balance in adults consuming self-selected diets. *Am J Clin Nutr* 40, 786-793.
31. Johansson G, Bingham S & Vahter M (1999) A method to compensate for incomplete 24-hour urine collections in nutritional epidemiology studies. *Public Health Nutr* 2, 587-591.
32. German Institute of Human Nutrition (2009) The Multiple Source Method (MSM). https://nugo.dife.de/msm/
33. Kleibaum DG, Kupper LL, Nizam A *et al.* (2008) *Applied regression analysis and other multivariable methods*. 4th ed. Belmont, California: Thomson Brooks/Cole.
34. Willett W, Howe G & Kushi L (1997) Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* 65, 1220S-1228.
35. Field AP (2005) Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychol Methods* 10, 444-467.
36. Subar AF, Kipnis V, Troiano RP *et al.* (2003) Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol* 158, 1-13.
37. Kahn HA, Whelton PK, Appel LJ *et al.* (1995) Validity of 24-hour dietary recall interviews conducted among volunteers in an adult working community. *Ann Epidemiol* 5, 484-489.
38. Bingham SA & Day NE (1997) Using biochemical markers to assess the validity of prospective dietary assessment methods and the effect of energy adjustment. *Am J Clin Nutr* 65, 1130S-1137S.
39. Heerstrass DW, Ocke MC, Bueno-de-Mesquita HB *et al.* (1998) Underreporting of energy, protein and potassium intake in relation to body mass index. *Int J Epidemiol* 27, 186-193.
40. Freedman LS, Midthune D, Carroll RJ *et al.* (2004) Adjustments to improve the estimation of usual dietary intake distributions in the population. *J Nutr* 134, 1836-1843.
41. Bingham SA, Gill C, Welch A *et al.* (1997) Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers. *Int J Epidemiol* 26 Suppl 1, S137-151.
42. Shai I, Rosner BA, Shahar DR *et al.* (2005) Dietary evaluation and attenuation of relative risk: multiple comparisons between blood and urinary biomarkers, food frequency, and 24-hour recall questionnaires: the DEARR study. *J Nutr* 135, 573-579.
43. Olafsdottir AS, Thorsdottir I, Gunnarsdottir I *et al.* (2006) Comparison of women's diet assessed by FFQs and 24-hour recalls with and without underreporters: associations with biomarkers. *Ann Nutr Metab* 50, 450-460.
44. Lissner L, Troiano RP, Midthune D *et al.* (2007) OPEN about obesity: recovery biomarkers, dietary reporting errors and BMI. *Int J Obes (Lond)* 31, 956-961.
45. Slimani N, Fahey M, Welch AA *et al.* (2002) Diversity of dietary patterns observed in the European Prospective Investigation into Cancer and Nutrition (EPIC) project. *Public Health Nutr* 5, 1311-1328.
46. Halkjaer J, Olsen A, Bjerregaard LJ *et al.* (2009) Intake of total, animal and plant proteins, and their food sources in 10 countries in the EPIC. *Eur J Clin Nutr* 63 Suppl 4, S16-36.
47. Rumpler WV, Kramer M, Rhodes DG *et al.* (2008) Identifying sources of reporting error using measured food intake. *Eur J Clin Nutr* 62, 544-552.
48. O'Brien MM, Kiely M, Galvin M *et al.* (2003) The importance of composite foods for estimates of vegetable and fruit intakes. *Public Health Nutr* 6, 711-726.
49. Cosgrove M, Flynn A & Kiely M (2005) Impact of disaggregation of composite foods on estimates of intakes of meat and meat products in Irish adults. *Public Health Nutr* 8, 327-337.
50. Jones DB (1941) *Factors for converting percentages of nitrogen in foods and feeds into percentages of protein*. Washington, DC: US Department of Agriculture-circ.
51. Pryer JA, Vrijheid M, Nichols R *et al.* (1997) Who are the 'low energy reporters' in the dietary and nutritional survey of British adults? *Int J Epidemiol* 26, 146-154.
52. Palaniappan U, Cue RI, Payette H *et al.* (2003) Implications of day-to-day variability on measurements of usual food and nutrient intakes. *J Nutr* 133, 232-235.
53. Carriquiry AL (2003) Estimation of usual intake distributions of nutrients and foods. *J Nutr* 133, 601S-608S.
54. Kipnis V, Midthune D, Freedman LS *et al.* (2001) Empirical Evidence of Correlated Biases in Dietary Assessment Instruments and Its Implications. *Am. J. Epidemiol.* 153, 394-403.

biasproteinpotassiummultilev
elcentres24hrecallefcovalepic
european analysisv
ariati nitro
genfit ibrati
onlab mrec
overy effec
tmod dexe

Chapter 3

Bias in protein and potassium intake collected with 24-h recalls (EPIC-soft) is rather comparable across European populations

Sandra P. Crispim, Anouk Geelen, Jeanne H. M. de Vries,
Heinz Freisling, Olga W Souverein, Paul J. M. Hulshof, Marga C. Ocke,
Hendriek Boshuizen, Lene F. Andersen, Jiri Ruprich, Willem De Keizer,
Inge Huybrechts, Lionel Lafay, Maria Santucci de Magistris, Fulvio Ricceri,
Rosario Tumino, Vittorio Krogh, H. Bas Bueno-de-Mesquita,
Joline WJ Beulens, Marie-Christine Boutron-Ruault, Androniki Naska,
Francesca L Crowe, Heiner Boeing, Alison McTaggart, Rudolf Kaaks,
Pieter van't Veer and Nadia Slimani

ducationindividualandcentrel
evelvariablesrandomeffectsva
riationdayoftheweekmodeofa
dministrationbmifoodcompos
itiontableslikelihoodratiocent
rerandomeffectvpcvarianceex

Abstract

We investigated whether group level bias of a 24-h recall estimate of protein and potassium intake, as compared to biomarkers, varied across European centres and whether this was influenced by characteristics of individuals or centres. The combined data from EFCOVAL and EPIC studies included 14 centres from 9 countries (*n* 1841). Dietary data was collected using a computerised 24-h recall (EPIC-Soft). Nitrogen and potassium in 24-h urine collections were used as reference method. Multilevel linear regression analysis was performed, including individual-level (e.g., BMI) and centre-level (e.g., food pattern index) variables. For protein intake, no between-centre variation in bias was observed in men while it was 5.7% (i.e., coefficient of variation) in women. For potassium intake, the between-centre variation in bias was 8.9% in men, and null in women. BMI was an important factor influencing the biases across centres (p<0.01 in all analyses). In addition, mode of administration (p=0.06 in women) and day of the week (p=0.03 in men and p=0.06 in women) may have influenced the bias in protein intake across centres. After inclusion of these covariables, between-centre variation in bias in protein intake disappeared for women, whereas for potassium it increased slightly in men (to 9.5%). Centre-level variables did not influence the results. In conclusion, the results suggest that group-level bias in protein and potassium (for women) collected with 24-h recalls does not vary across centres and to a certain extent varies for potassium in men. BMI and study design aspects, rather than centre-level characteristics, affected the biases across centres.

Introduction

There is an increasing interest in identifying and understanding geographical variations in dietary intake. For instance, a number of international studies have been performed in Europe with the purpose of investigating dietary exposure and testing hypotheses on diet-disease associations assessing dietary intake collected in different geographical areas[1-3]. Another example is that dietary intake data collected through national food consumption surveys across different European countries can be used to develop and evaluate nutritional policies under the EU commission framework[4]. However, to correctly estimate the variation in dietary intake across populations in those investigations, it is necessary to obtain data that are as accurate and comparable as possible.

The collection of dietary data for comparisons between populations should preferably be performed using the same and standardised dietary assessment method. To that end, a repeated non-consecutive 24-h dietary recall interview using EPIC-Soft has been recommended for assessing dietary intake in future national food consumption surveys[4; 5]. Subsequently, the evaluation of this method was performed within the European Food Consumption Validation (EFCOVAL) study[6].

An established approach to evaluate the validity of dietary assessment instruments is to compare self-reported dietary intake with its related biomarker estimates. In particular, recovery-based biomarkers have a precisely known quantitative relation to absolute daily intake and are a valid reference to estimate the bias in dietary intake reports[7]. Moreover, recovery biomarkers provide reference estimates of dietary intake with errors that are likely to be uncorrelated with the errors of self-reported dietary methods[8]. Two of the few available recovery biomarkers to assess the bias in nutrient intake are urinary nitrogen and potassium[9; 10].

Previously, the accuracy of protein as estimated by one 24-h dietary recall using EPIC-Soft has been evaluated using urinary nitrogen in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. In this study, protein intake was underestimated at the group level and varied across European centres, i.e., ratios between nitrogen intake and excretion ranged from 0.69 (Greece) to 0.99 (Ragusa-Italy) in men and from 0.54 (Greece) to 0.92 (Paris-France) in women[11]. More recently, in the EFCOVAL study, the average of two non-consecutive days of protein and potassium intake assessed with this computerised 24-h recall and compared to their respective biomarkers revealed underestimation that ranged across five European centres between 2 and 13% for protein intake and between 4 and 17% for potassium intake[12]. These results suggested that differences in the performance of the 24-h recall may exist across European countries.

A number of reasons have been hypothesized to explain the observed variation in biases in protein and potassium intake between-centres in the EPIC and EFCOVAL studies. For instance, differences in characteristics at the centre (e.g., food pattern) or individual level (e.g., socioeconomic status, BMI) could explain differential misreporting of dietary intake. However, an evaluation of the potential effect of characteristics at the individual and centre (country) level on the validity of the method was lacking. The analyses initially conducted in the EPIC and EFCOVAL data on protein and potassium bias used a single-level model with 'fixed effects', which did not allow for simultaneous separation of within- and between-centre

variance. These previous analyses also did not consider all possible explanatory variables at the individual and centre levels to be included in the model. Therefore, to gain a more in-depth understanding of the accuracy of nutrient intake assessed by the 24-h recall across European centres, the individual and centre level, ought to be considered simultaneously. For that purpose, multilevel modelling can be used by means of 'random effect models'. The random effect model approach allows to estimate the effects of individual- and centre-level characteristics, and assess their impact on the estimates of between-centre variation in the nutrient bias[13].

Furthermore, pooling the data from the EFCOVAL and EPIC studies increased the number of geographical regions considered, the heterogeneity of the dietary patterns and the statistical power to evaluate the bias in protein and potassium intake collected with 24-h recalls across European populations using multilevel analysis. Therefore, the objective of this paper was to further investigate whether the group-level bias in intake of protein and potassium collected with 24-h recalls using EPIC-Soft varied across European centres and whether this was affected by characteristics at the individual and centre level.

Subjects and Methods

Study Population

This study combines study populations from two European studies, the EPIC calibration sub-study and the EFCOVAL validation study, together representing 9 European countries. Previous publications described in detail the rationale and methods of the studies[1; 14-16]. Within the EPIC cohort, ~37,000 individuals comprised the subsample of the calibration sub-study. Between 1995 and 2000, these individuals were randomly chosen from the EPIC cohorts for completing a single standardised 24-h dietary recall (EPIC-Soft) to calibrate baseline food frequency questionnaires (FFQ)[1]. More details about the study population from the calibration study are reported elsewhere[11; 17]. In a convenience subsample of the calibration study, 24-h urines were collected from 1386 participants from 12 EPIC centres in 6 countries (Paris in France; Florence, Naples, Ragusa, Varese and Turin in Italy; some combined regions in Greece; Cambridge and Oxford in the United Kingdom; Bilthoven in the Netherlands; and Heidelberg and Potsdam in Germany). Urine was collected over the

same day as the 24-h recall (44%) or within a maximum of 6 days afterwards (56%). Furthermore, lifestyle information was collected at baseline from all EPIC study participants. To optimize the sample sizes in some centres, the initial 12 centres from the EPIC administrative areas were redefined into 9 centres, labelled hereafter as Heidelberg, Potsdam, Paris, Greece, Central/Southern Italy (including Florence, Naples and Ragusa), Northern Italy (including Varese and Turin), Bilthoven, Cambridge, and Oxford. In the EFCOVAL validation study, dietary information was collected in five European centres, i.e., Ghent (Belgium), Brno (the Czech Republic), Nice (France), Wageningen (the Netherlands), and Oslo (Norway), in the years 2007 and 2008. In total, 600 participants underwent two standardised 24-h recall interviews using EPIC-Soft software. In addition, two 24-h urines, covering the same days as the 24-h recalls, were collected. Both studies were conducted according to the guidelines laid down in the Declaration of Helsinki, and procedures involving human subjects were approved by ethical committees of the centres involved in the data collection. In the combined assessment, 1986 participants from 14 European centres (9 from the EPIC study) were initially used. From these, 145 participants were excluded from the protein analyses and 176 from the potassium analyses. Reasons for exclusion were that data of the 24-h recall (*n* 18), urinary protein (*n* 13) or potassium (*n* 44) was not available, participants were on a restricted diet (*n* 51) or <50% of *para*-Aminobenzoic acid (PABA) was recovered (*n* 63) – see details in the 24-h urine collection section. Thus, the final sample in the dataset included 1841 participants for the data analyses of protein and 1810 for potassium. An overview of the two studies and the pooled data is given in **Table 1**.

Dietary data

In both the EPIC and the EFCOVAL study, the 24-h recalls were collected using EPIC-Soft version 9.16. The structure and standardisation procedure of EPIC-Soft have been described elsewhere[12; 18; 19]. Briefly, EPIC-Soft is a computer-assisted 24-h dietary recall that follows standardised steps when describing, quantifying, probing and calculating dietary intake[18].

Table 1 – Overview of EPIC and EFCOVAL studies and the pooled database

PARAMETER	EPIC (1995-2000)	EFCOVAL (2006-2009)	Pooled
n	600	1386	1841 after exclusions
24-h Recall			
Number of administrations	1	2	1st
Mode of administration	FF	FF/T (at random)	Both FF/T
Days of the week	All included, uneven	All included, with small differences	All included, uneven
EPIC-Soft Version	9.16	9.16	9.16
Photo booklets	Full version developed at IARC	Country-specific selection with new pictures on bread shapes and household measurements	
Nutrient values	Standardised European database (ENDB)	Country-specific FCT	Different levels of standardisations (see below)
Protein	Assumed the laboratory analyses used to assess protein in foods are comparable (mostly by Kjeldahl)	Assumed the laboratory analyses used to assess protein in foods are comparable (mostly by Kjeldahl)	Prot/N data between countries are comparable in terms of lab analysis
Conversion factor Nitrogen -> Protein	Harmonised PROT values by standardizing CF as follows: - If N available then PROT = N x 6.25; otherwise: - If N_CF available then N=Prot / N_CF and new Prot=N x 6.25	Different CF used: FR, BE and NO (Jones Factors). CZ: 6.25 and NL 6.38 for dairy and all other foods 6.25	Non-standardised and standardised CF; In EFCOVAL: NL and FR standardised, others not. (see methods section)
Potassium	Assumed the laboratory analyses used to assess K in foods are comparable	Assumed the laboratory analyses used to assess K in foods are comparable	Assumed the laboratory analyses used to assess K in foods are comparable
Retention factor: Losses in K when foods are cooked	RF applied: Cooked single foods linked to raw foods were adjusted by retention factors (food group specific)	K losses were not considered when some cooked foods were linked to raw foods	K contents of single foods in EFCOVAL were adjusted as done in EPIC

PARAMETER	EPIC (1995-2000)	EFCOVAL (2006-2009)	Pooled
Biomarker	1x24h urine collection	2x24h urine collections	1st urine corresponding 1st recall
Urinary Nitrogen (UN)	Kjeldahl method (lab in UK)	Kjeldahl method (lab in the NL)	Laboratorial comparison in a subsample
Urinary potassium (UK)	Flame photometry (lab in UK)	Ion Electrode (lab in the NL)	Laboratorial comparison in a subsample
PABA correction	excluded < 70% and >110%	excluded < 50%	excluded <50%
	PABA adjustment between 70-85%	PABA adjustment between 50-85%	PABA adjustment between 50-85%
Other non-dietary data			
Educational level	5 categories: none, primary, technical/professional school, secondary and longer education (inc. university)	3 categories: low, intermediate and high	4 categories: none, low, intermediate and high
Weight and Height	Measured and self-reported that have been corrected, except for Paris sample (See Haftenberger et al. 2002)[27]	Measured	Measured and self-reported that have been corrected, except to Paris sample

CF = Conversion factor
EFCOVAL = European Food Consumption Validation
ENDB = EPIC Nutrient Database
EPIC = European Prospective Investigation into Cancer and Nutrition
FCT = Food composition tables
FF = Face-to-Face interview
FR = France; BE = Belgium; NO = Norway; CZ = the Czech Republic; NL = the Netherlands, UK = United Kingdom
IARC = International Agency for Research on Cancer
K = Potassium
N = Nitrogen
PABA = *para*-Aminobenzoic acid
PROT = Protein
RF = Retention factor
T = Telephone interview
UK = Urinary Potassium
UN = Urinary Nitrogen

The 24-h recalls were collected by trained dietitians through face-to-face interviews in the EPIC centres. In EFCOVAL, one telephone and one face-to-face interview were applied in random order in each subject. In both studies, dietary data of all days of the week were collected. In EPIC, protein and potassium food composition values from each national food composition database were standardised across countries within the European Nutrient Database (ENDB) project, a collaboration with national compilers and other international experts[20]. For EFCOVAL, protein and potassium intake were calculated using country-specific food composition databases.

In the pooled dataset, we only used the first 24-h recall data from the EFCOVAL participants. Consequently, the EFCOVAL measurements consisted of 24-h recalls collected by telephone and face-to-face interviews. Furthermore, an attempt has been made to standardize food composition values between EPIC and EFCOVAL studies. Similar to what has been done within the ENDB framework, losses in the potassium values of cooked single foods, that have been linked to raw foods in the food composition data, were adjusted by applying the same retention factors than those initially used for the EPIC data. For protein, standardisation of the EPIC data was performed by applying the 6.25 conversion factor (CF) instead of food-specific CFs to convert nitrogen into protein intake. Within EFCOVAL data, such standardisation was only possible for the data from Wageningen (NL) and Nice (FR) because it was not possible to retrieve the original CF information applied in the protein composition of the foods in the other centres. Energy values were computed by adding the contributions from protein, carbohydrates, fat and alcohol intake and using related Atwater factors (17, 17, 37 and 29 kJ per gram, respectively)[21].

There were some differences between the databases used in the EPIC-Soft software in the EPIC and EFCOVAL studies. These differences were mainly related to the upgrade of food lists, standard units, descriptors for food identification[22], and selection of food pictures for food quantification. Nevertheless, the purpose of updating these databases in EPIC-Soft was to take into account actual differences in consumption between the centres while the procedures to collect them were still standardised.

24-h urine collection and recovery biomarkers for protein and potassium intake

24-h urine collections were verified for completeness by using *para*-Aminobenzoic acid (PABA) tablets (PABAcheck, Laboratories for Applied Biology, London). Complete logistics of 24-h urine collections and laboratory analyses are described

elsewhere[11]. In brief, after collection, the 24-h urines were transported to the study centres where they were weighed and aliquoted. Then, specimens were stored at -20°C until shipment on dry ice to the central laboratories in Cambridge (EPIC) and Wageningen (EFCOVAL). Urinary nitrogen was determined by the Kjeldahl technique in both studies. Urinary potassium was determined using an IL 943 flame photometer (Instrumentation Laboratory) in EPIC and using an ion-selective electrode on a Beckman Synchron LX20 analyser in EFCOVAL. PABA was measured by colorimetry in both studies[11; 12; 23]. Urine samples with PABA recoveries below 50% were treated as incomplete and excluded from the data analyses. Specimens containing between 50-85% of PABA recovery had their urinary protein and potassium concentrations proportionally adjusted to 93 per cent[24]. Furthermore, we did not exclude participants with PABA recovery above 110%, as we assumed that those collections were complete. This procedure for dealing with PABA recovery is different from previously published data in the EPIC study[11], resulting in a larger sample sizes for some EPIC centres. Taking into account extra-renal losses (~19%) and the fact that protein on average contains 16% of nitrogen, urinary protein was calculated as $[6.25 \times (\text{urinary nitrogen}/0.81)]$[9; 25]. Urinary potassium was estimated by dividing the measured value by 0.77, assuming that 77% of potassium intake is excreted through the urine when considering faecal excretion[10; 26].

Laboratory calibration study

With the purpose of harmonizing biomarker laboratory data, a calibration study was conducted among laboratories that performed analyses in the EPIC and EFCOVAL studies. Therefore, during the Summer of 2008, 45 urine samples of the EPIC study that were previously analysed for protein and potassium content by the MRC Dunn Clinical Nutrition Centre in Cambridge (UK) were reanalysed by the laboratory at Wageningen University (NL). The results obtained from the two laboratories were compared. In addition, comparability of laboratory methods used in EPIC and EFCOVAL labs was further substantiated by evaluating standard reference materials and quality control procedures (e.g., inter-laboratory proficiency tests) of each lab measurement. A report of the laboratory comparison between studies is presented in **Appendix 1**. Shortly, we did not observe statistically significant differences between the measurements by the two labs for nitrogen or potassium. Therefore, calibration of data between both studies was not necessary and original biomarker data of the two studies was used in our analyses.

Anthropometrics and educational level

In both studies, measurements of body weight and height were collected for the calculation of body mass index (BMI). In EPIC, some measurements were self-reported and were corrected by prediction equations, as described in Haftenberger et al.[27]

Furthermore, a general lifestyle questionnaire, including educational level information, was applied at the start of each study. Educational level was categorized using different categories in the EPIC and EFCOVAL studies (**see table 1**). The proposed classification for the pooled data analyses included the following categories: none, low, intermediate and high, in which technical and secondary groups of education from the EPIC data were treated as intermediate levels.

Explanatory variables

Based on pre-existing knowledge, we selected full sets of explanatory variables to be included in the models, which we expected to vary across individuals or centres and be correlated with the nutrient bias or intake or biomarker levels. Variables at the individual level were age (in years), educational level (categorical), BMI (in kg/m²), mode of administration of the 24-h recall (face-to-face vs. telephone), day of the week of the 24-h recall (week- vs. weekend days) and year of recruitment. Explanatory variables at centre level were study (EFCOVAL vs. EPIC), human development index (HDI[28]) and a food pattern index. We used the HDI as a proxy for identifying socioeconomic differences across the centres. The HDI statistic is composed from national data on life expectancy, education and per-capita gross domestic product, as an indicator of standard of living, at the country level. Thus, centres in the same country had the same HDI. To capture the variability existing in food pattern across the European centres, a food pattern index was calculated for each individual and averaged out for each centre. From the 'Diet quality index-international' (DQI-I)[29], we used the variety index component to indicate the diversity in food group intake between the centres. This index assesses whether intake comes from diverse sources both across and within food groups, and varies from 0 to 20 points. It is divided in two parts. First, the overall food group variety is assessed by inclusion of at least one serving food per day from each of five food groups (meat/poultry/fish/egg, dairy, grains, fruits, and vegetables). Second, variety within protein sources is evaluated i.e., number of protein sources. The lowest food index score in our assessment was

attributed to Oxford (vegetarians)-UK (10.5 points) and the highest to the 3 Spanish centres (> 18.5 points).

Statistical analyses

Multilevel linear regression models were used to assess the variation in group-level bias of protein and potassium intake across the centres and to estimate the effects of individual- and centre-level explanatory variables on this variation. Individuals were set at the first level and centres at the second. Statistical analyses were conducted separately for men and women since our previous single-level analyses showed different group-level bias for each gender[12]. The number of centres in the analysis of each gender is different since the research centre in Paris only included women.

Bias was expressed as the ratio between nutrient intake and its excretion. We chose the ratios instead of absolute values to take into account differences that were related to high or low protein and potassium intake across centres. These ratios were treated as the dependent variable in the regression models and were log-transformed to improve normality (ln(individual ratio)).

We fitted three regression models in an increasing order of complexity (**Appendix 2**). Model (i) included a random effect to model between-centre variation of protein and potassium biases across centres (i.e., random intercepts) without explanatory variables. Therefore, we were able to estimate the between-centre variances in group-level bias in a crude model. In model (ii), individual-level explanatory variables were added to the fixed part of the model, whereas in model (iii) centre-level variables were also included. Full sets of individual- and centre-level explanatory variables were included in their respective regression models and the optimal subsets of variables were chosen by using a backward selection. The fit of the models was tested by the likelihood ratio test, which compared minus twice the difference of the maximum likelihood (ML) of that model with the preceding nested model[13]. The likelihood ratio test statistic was compared to a chi-square distribution with degrees of freedom equal to the number of extra parameters in the more complex model[13]. Results are only presented for models that showed a statistically significant improvement. Furthermore, we also attempted to include random slopes to allow the effects of age and BMI to vary between centres, but their results suggested homogeneity of the effects and they were, therefore, not included in the paper.

The total variance of each model was partitioned in two components, the between-centre variance (or centre random effect - σ_{u0}^2), and the within-centre between-individual variance (or individual random effect - σ_{e0}^2). To quantify the variation in nutrient biases across centres, we looked at the between-centre random effect obtained across the fitted models. Even though zero between-centre variation in bias may have been observed in a simpler model, we proceeded with the more complex ones to check whether the variance estimates would change by including different terms into the model (e.g., inclusion of explanatory variables). To interpret the contribution of between-centre variance, we used two approaches, the variance partition coefficient (VPC) and the coefficient of variation (CV) between centres. The VPC was calculated as the proportion of total variance that is due to differences between centres[13].

$$\text{VPC} = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_{e0}^2}$$

The CV expresses the variation in the bias between centres as a percentage, relative to the intake according to the reference method. Because the analysis of the bias was done on the logarithmic scale and the ratios on the centre level were close to one,

$$\text{CV} = \sqrt{\sigma_{u0}^2}$$

Statistical analyses were carried out using SAS statistical package, version 9.1 (SAS Institute Inc., Cary, NC, USA).

Results

All centres combined, both men and women underreported protein intake from one-day 24-h recall by 3 and 5% (ratio intake/excretion=0.97 and 0.95), respectively (**Table 2**). In men, the ratio between protein intake and excretion varied from 0.89 in Wageningen (NL) to 1.03 in Central/Southern Italy (IT). In women, the ratio varied from 0.84 in Greece (GR) to 1.05 in Oslo (NO). Average underestimation of potassium intake was 1% in men and 3% in women. In men, the lowest ratio between potassium intake and biomarker excretion was observed in Nice (FR) and Heidelberg (DE) with 0.86 whereas the highest ratio was seen in Northern Italy (IT) with 1.17. In women, the lowest ratio was 0.90 in Potsdam (DE) and the highest ratio was 1.08 in Greece (GR).

Protein intake

Based on the centre random effect, between-centre variance in protein bias was null ($\sigma_{u0}^2 \sim 0$) in men (**Table 3**). In women, the between-centre CV in protein biases was initially 5.7%, which was 3% of the total variance, and Greece (GR), Paris (FR) and Oslo (NO) were the centres with a group-level bias deviating from the overall mean bias (**Table 4**). After inclusion of individual explanatory variables, especially BMI, the between-centre variance in bias was reduced by 78% (from 0.0032 to 0.0007) in women (p<0.001). In addition, the remaining between-centre variance in protein biases (CV=2.6%) was not significant anymore and no centre appeared to deviate from the mean bias. Other variables that may have contributed to the reduction of between-centre variance in protein biases in women were 'day of the week' (p=0.06) and 'mode of administration' (p=0.06). When we added centre-level variables (e.g., HDI), we did not observe a significant improvement of the model's fit neither for men nor women (data not shown). Therefore, model ii (random intercepts to model the centre effect with inclusion of variables at the individual level) was retained as the most adequate model to the data (**Tables 3 and 4**).

Table 2 – Protein and potassium intake* based on 24-h recalls and urinary biomarkers and their ratios for European centres participating in the EPIC and EFCOVAL studies.

		Protein (g/day)						Potassium (mg/day)					
		Intake		Biomarker		Ratio†			Intake		Biomarker		Ratio
Centre	n	Mean	10th-90th	Mean	10th-90th	Mean	10th-90th	n	Mean	10th-90th	Mean	10th-90th	Mean	10th-90th
Men													
BE: Ghent	62	102.1	70.4-141.0	111.0	77.0-148.6	0.95	0.65-1.33	62	4098	2711-5762	4119	2774-6105	1.03	0.71-1.41
CZ: Brno	57	103.7	58.5-151.1	104.1	76.5-132.6	1.02	0.64-1.43	57	3821	2092-5226	3566	2394-4885	1.04	0.62-1.53
DE: Heidelberg	41	91.5	48.0-141.1	102.1	70.2-133.6	0.97	0.52-1.42	41	3943	2323-5631	4648	3509-5768	0.86	0.53-1.21
DE: Potsdam	78	90.9	53.5-130.5	100.4	78.2-124.3	0.93	0.55-1.30	60	3732	2493-5338	3935	2323-5606	1.02	0.66-1.59
FR: Nice	53	100.8	60.6-157.8	107.7	81.8-132.9	0.94	0.65-1.38	53	3510	2059-4936	4183	2595-5728	0.86	0.55-1.20
GR: Greece	49	81.9	37.5-134.3	84.4	58.6-117.7	0.99	0.52-1.44	49	3180	1186-4927	2587	1412-4057	0.99	0.75-1.31
IT: Central	24	114.8	63.5-173.7	113.9	78.2-136.7	1.03	0.51-1.48	24	4135	2785-5611	4008	2595-6408	1.12	0.52-1.59
IT: Northern	56	113.5	71.5-165.1	113.1	84.9-147.3	1.02	0.68-1.45	56	4217	2666-6137	3687	2722-5098	1.17	0.77-1.62
NL: Bilthoven	81	105.9	61.3-170.0	109.9	82.1-137.6	1.00	0.50-1.61	81	4293	2392-6475	4293	2767-5321	1.05	0.58-1.61
NL: Wageningen	58	101.5	61.5-149.1	117.2	82.6-157.2	0.89	0.56-1.28	58	4422	2971-6714	4572	3001-6228	0.99	0.75-1.31
NO: Oslo	62	112.0	79.9-148.9	116.9	82.0-158.2	0.99	0.65-1.47	62	4719	3152-6464	4969	3338-6387	0.98	0.60-1.48
UK: Cambridge	154	90.4	56.8-121.9	95.7	70.4-123.2	0.98	0.60-1.43	154	3949	2622-5384	4174	2656-5636	1.00	0.62-1.50
UK: Oxford	42	91.7	61.3-133.2	98.2	74.2-120.9	0.95	0.63-1.36	42	3969	2734-5274	4596	2960-6560	0.94	0.52-1.33
All centres	817	98.8	57.5-145.2	104.7	74.0-140.3	0.97	0.60-1.43	799	4013	2451-5821	4122	2575-5858	0.99	0.62-1.53

Women	n	Mean	10th-90th	Mean	10th-90th	Mean	10th-90th	n	Mean	10th-90th	Mean	10th-90th	Mean	10th-90th
BE: Ghent	59	79.6	46.4-109.5	84.6	59.4-114.6	0.98	0.60-1.39	59	3651	2325-5699	3948	2486-5169	0.97	0.62-1.53
CZ: Brno	60	70.7	45.5-98.9	77.4	57.1-94.0	0.94	0.52-1.38	60	3140	2184-4352	3226	2002-4352	0.97	0.61-1.38
DE: Heidelberg	48	72.8	35.2-102.4	80.8	45.8-119.0	0.96	0.47-1.40	48	3378	1880-4887	3665	2163-5291	0.98	0.58-1.30
DE: Potsdam	56	67.7	40.6-101.4	76.0	55.6-103.8	0.92	0.52-1.33	43	3269	1994-4572	3799	2188-5113	0.90	0.57-1.32
FR: Nice	57	76.8	54.9-100.6	82.4	61.8-102.1	0.96	0.67-1.31	57	3251	2214-4056	3685	2571-4933	0.93	0.59-1.28
FR: Paris	116	86.6	55.1-122.0	86.8	64.4-112.7	1.03	0.58-1.57	116	3459	2315-4877	3737	2412-5230	0.96	0.66-1.33
GR: Greece	52	56.9	29.7-93.6	69.6	47.1-91.7	0.84	0.38-1.32	52	2428	1257-3793	2378	1691-3377	1.08	0.54-1.78
IT: Central	71	78.1	50.9-113.8	88.5	62.1-115.9	0.90	0.63-1.28	71	3148	2003-4580	3428	2326-4722	0.98	0.60-1.54
IT: Northern	46	75.8	48.2-108.8	91.4	65.5-120.8	0.85	0.52-1.32	46	3066	1892-4127	3073	2168-4118	1.03	0.64-1.49
NL: Bilthoven	116	78.7	49.2-110.7	84.7	57.9-112.4	0.95	0.63-1.38	116	3539	2210-5094	3840	2300-5768	0.98	0.61-1.42
NL: Wageningen	60	78.8	47.6-108.7	83.1	62.5-108.4	0.96	0.62-1.26	60	3622	2314-4820	3933	2704-5387	0.95	0.71-1.36
NO: Oslo	62	86.3	52.9-114.9	84.7	61.1-109.2	1.05	0.74-1.37	62	3695	2505-5082	3839	2769-4946	0.99	0.69-1.41
UK: Cambridge	174	74.8	47.4-102.2	79.0	56.8-101.2	0.98	0.61-1.38	174	3418	2266-4721	3859	2620-5464	0.93	0.57-1.29
UK: Oxford	47	59.6	43.8-82.6	71.6	55.4-90.5	0.85	0.59-1.17	47	3169	2221-3993	3621	2463-4636	0.91	0.64-1.24
All centres	1024	75.8	45.4-108.0	81.9	58.3-108.0	0.95	0.58-1.36	1011	3348	2095-4721	3639	2300-5078	0.97	0.61-1.38

*Mean values and inter-quintiles (10th-90th percentiles)

†Intake/excretion

Table 3 – Multilevel regression analysis of the log-transformed ratio between protein intake and excretion in men across 13 European centres participating in the EPIC and EFCOVAL studies.

	Model i	Model ii	
Model*	Random intercept for centre – no explanatory variables	Random intercept for centre – explanatory variables at the individual level[a]	
n	817	817	
Likelihood ratio	673	644	
Likelihood ratio test[†]	-	P<0.001	
σ_{u0}^2 – Centre random effect ± SE (p-value)	0.000	0	
CV (%, relative to reference method)	0%	0%	
σ_{e0}^2 – Within centre random effect ± SE (p-value)	0.133 ± 0.007 (<0.001)	0.129 ± 0.006 (<0.001)	
VPC – Variance partition coefficient	0	0	
[a]Individual variables – effect (p-values)	-	BMI Week- vs. weekend days	-0.02 (<0.001) -0.06 (0.03)
Proportion of between centre variance explained[†]	-	0%	
Centres with bias deviating from the mean log-transformed ratio	None	None	

[*] Fit of model iii was not significantly better than the previous one. Therefore, results are not presented.

[†] Compared to the previous fitted model.

Table 4 – Multilevel regression analysis of the log-transformed ratio between protein intake and excretion in women from 14 European centres from the EPIC and EFCOVAL studies.

	Model i	Model ii	
Model*	Random intercept for centre – no explanatory variables	Random intercept for centre – explanatory variables at the individual level[a]	
n	1024	1024	
Likelihood ratio	751	713	
Likelihood ratio test[†]	-	p<0001	
σ^2_{u0} – Centre random effect ± SE (p-value)	0.0032 ± 0.002 (0.05)	0.0007 ± 0.001 (0.24)	
CV (%, relative to reference method)	5.7%	2.6%	
σ^2_{e0} – Within centre random effect ± SE (p-value)	0.120 ± 0.005 (<0.001)	0.117 ± 0.005 (<0.001)	
VPC – Variance partition coefficient	0.03	0.006	
[a]Individual variables – effect (p-values)	-	BMI Week- vs. weekend days Mode of administration	-0.01 (0.001) -0.05 (0.06) 0.06 (0.06)
Proportion of between centre variance explained[†]	-	78%	
Centres with bias deviating from the mean log-transformed ratio	Greece (GR), Paris (FR), Oslo (NO)	None	

[*] Fit of model iii was not significantly better than the previous one. Therefore, results are not presented.

[†] Compared to the previous fitted model.

Potassium intake

In men, the between-centre CV in potassium biases was initially 8.9% (Model i), which was about 5% of the total variance (**Table 5**). When applying model ii, the between-centre CV slightly increased to 9.5%. Furthermore, the biases from 4 centres, i.e., Greece (GR), Heidelberg (DE), Nice (FR) and Northern Italy (IT) seemed to differ from the overall mean potassium bias. Individual BMI was a factor influencing the between-centre variance in men (p=0.002). No between-centre variance ($\sigma_{u0}^2 = 0$) was initially observed in the potassium biases in women (**Table 6**). After including individual variables in the model, BMI predicted the bias and there was still no significant variation across centres in women (CV=1.7%). As for the protein analyses, inclusion of centre-level variables (model iii) did not improve the fit of the model, for men and women.

Discussion

In this paper, we investigated the variation in group-level bias in self-reported protein and potassium intake collected with the computerised 24-h recall (EPIC-Soft) across European adult populations. By using a multilevel modelling approach, we observed that the bias in protein intake did not vary across centres in men, but varied among women (5.7% of variation) in the crude model with random intercepts. Bias in potassium intake differed between centres in men (8.9% of variation), but not in women. Explanatory variables at the individual level (i.e., BMI, day of the week and mode of administration) predicted and explained the between-centre variation of bias in protein and potassium intake. When those were included in the model, the bias in protein intake in women did not significantly vary anymore, and the bias in potassium intake remained with variations across centres (9.5% of variation). Selected centre-level variables (i.e., HDI) did not influence the between-centre variations in bias in our assessment.

The major advantage of using multilevel analysis was that we were able to separate the two variance components (i.e., within- and between-centre) in protein and potassium bias in one sole model, which is important for a reliable comparison of populations[30; 31]. In addition, in this unique setting of combining datasets from two European studies, we were able to use dietary and biomarker measurements that were collected using standardised methodologies. A comparison of laboratory measurements was performed to overcome possible inter-laboratory errors and an important level of standardisation was achieved by estimating protein and potassium intake from food

composition tables across the different European centres, although not completely for Ghent (BE), Brno (CZ) and Oslo (NO). Furthermore, the large number of centres originating from different regions of Europe, allowed us to compare populations with different dietary intake profiles.

Yet, our study has limitations that should be considered in the interpretation of our findings and in the development of future research. First, we cannot assume that these results can be extrapolated for other points of the distribution of protein and potassium intake, which are important to assess prevalence above or below a certain cut-off point[32]. As previously shown, we may expect that the accuracy of other points of the distribution, between the mean and the ends of the tails, is inferior compared to the mean bias at the population level[12]. Nevertheless, this has been the first attempt of using a multilevel approach to validate dietary intake in an international context, and an important understanding of between-centre variation in nutrient intake bias as well as factors that can influence the performance of the method has been achieved. Second, we were not able to completely harmonize the food composition data for protein in EFCOVAL. However, when we excluded centres with non-standardised protein composition data from our main analysis, the results for protein did not change. Third, it can be questioned whether we have properly dealt with the results of the laboratory comparison, considering the small sample size in the calibration study. Based on the non-statistically significant differences obtained with the t-test, we opted not to calibrate the laboratory estimates. However, multilevel analysis with and without calibration of protein and potassium biomarker values resulted in similar results. At last, the generalisation of these results to other nutrients is not warranted given that foods and related nutrients might be differently misreported[33-35].

In other analysis with EFCOVAL and EPIC data[11; 12], the group-level bias of protein and potassium intake assessed with 24-h recalls varied across centres. A number of reasons were suggested to explain this variation in bias, as for instance a difference in BMI. Differential underreporting of dietary intake by overweight and obese individuals is expected based on the literature[36-38]. Indeed, BMI was the explanatory variable predicting most of the bias in protein and potassium intake in this analysis as well as explaining the variation of bias across the centres; thus, confirming the importance of considering BMI when performing the 24-h recalls in Europe.

Table 5 – Multilevel regression analysis of the log-transformed ratio between potassium intake excretion in men from 13 European centres from the EPIC and EFCOVAL studies.

Model*	Model i	Model ii
	Random intercept for centre – no explanatory variables	Random intercept for centre – explanatory variables at the individual level[a]
n	799	799
Likelihood ratio	715	706
Likelihood ratio test[†]		p=0.002
σ^2_{u0} – Centre random effect ± SE (p-value)	0.008 ± 0.004 (0.03)	0.009 ± 0.005 (0.02)
CV (%, relative to reference method)	8.9%	9.5%
σ^2_{e0} – Within centre random effect ± SE (p-value)	0.139 ± 0.007 (<0.001)	0.138 ± 0.006 (<0.001)
VPC – Variance partition coefficient	0.05	0.06
[a]Individual variables – effect (p-values)	-	BMI -0.01 (0.002)
Proportion of between centre variance explained[†]	-	0%
Centres with bias deviating from the mean log-transformed ratio	Greece (GR), Heidelberg (GE), Nice (FR), Northern Italy (IT)	Greece (GR), Heidelberg (GE), Nice (FR), Northern Italy (IT)

[*] Fit of model iii was not significantly better than the previous one. Therefore, results are not presented.

[†] Compared to the previous fitted model.

Table 6 – Multilevel regression analysis of the log-transformed ratio between potassium intake and excretion in women from 14 European centres from the EPIC and EFCOVAL studies.

	Model i	Model ii
Model*	Random intercept for centre – no explanatory variables	Random intercept for centre – explanatory variables at the individual level[a]
n	1011	1011
Likelihood ratio	642	629
Likelihood ratio test[†]	p<0.001	p<0.001
σ^2_{u0} – Centre random effect $\pm$ SE (p-value)	0.0000	0.0003 $\pm$ 0.0006 (0.34)
CV (%, relative to reference method)	0%	1.7%
σ^2_{e0} – Within centre random effect $\pm$ SE (p-value)	0.110 (0.005) <0.001	0.109 (0.005) <0.001
VPC – Variance partition coefficient	0	0.003
[a]Individual variables – effect (p-values)	-	BMI -0.01 (0.003)
Proportion of between centre variance explained[†]	-	0%
Centres with bias deviating from the mean log-transformed ratio	None	None

* Fit of model iii was not significantly better than the previous one. Therefore, results are not presented.
† Compared to the previous fitted model.

Besides BMI, the day of the week (week- vs. weekend days) and the mode of administration (face-to-face vs. telephone) appeared to influence the bias in protein intake across centres, but not in potassium. An explanation for this difference may be that potassium is a nutrient present in a greater variety of foods/food groups and more equally distributed among different food groups than protein[10]. Moreover, higher protein intake has been observed during weekend days across European populations when compared to weekdays[39]. What regards the comparability of different modes of administration, comparable results between telephone and face-to-face interviews could be expected[40-42], but perhaps populations with different dietary intake patterns respond differently to these two modes of administration. Actually, within the EFCOVAL study, we observed that 24-h recalls collected by telephone interviews seemed to provide a more accurate assessment than by face-to-face interviews in some research centres[43].

Furthermore, we observed a between-centre variation in group-level bias in potassium intake in men, but not in women. As differential reporting bias is suggested among genders, we speculate that improvements of the reported 24-h recalls might be expected if the person who does the shopping and/or the cooking of the foods is involved in the dietary interview.

We hypothesized that certain centre characteristics (e.g., food pattern index, HDI) could influence the variation of group-level biases in protein and potassium intake across the European centres. However, we observed almost no variation in biases across the centres, except for bias in potassium intake in men. Therefore, there was not much variation in bias to be explained by characteristics at the centre level. Nevertheless, we suppose that these characteristics may be relevant in the assessment of less regularly consumed nutrients and, especially, for foods and food groups, as we may expect a larger variation in the dietary intake assessment between populations in Europe than was found for the nutrients we assessed[44]. For that, more insight into food pattern indexes to represent country differences would be valuable, as the index we have used in this assessment may have not been sufficiently accurate.

In conclusion, the present results appear to bring us a step further to understand and quantify the variation in bias in the assessment of protein and potassium intake collected with 24-h recalls across European centres. Remarkably, almost no variation in protein and potassium biases of the 24-h recalls using EPIC-Soft was observed across the centres. In addition, the results of this study suggest that the group-level bias in protein intake for both genders and potassium intake for women does not vary across centres and to a certain extent varied for potassium intake in men.

Furthermore, the large number of centres originating from different regions of Europe, allowed us to compare populations with different dietary intake profiles. In view of that, the data to be collected in future pan-European food consumption surveys should be properly analysed and interpreted considering the characteristics that may influence the report of protein and potassium intake across countries. Above all, we suggest that it may be of special importance to additionally explore the between-centre effect in the ranking of self-reported food groups and infrequently consumed nutrients across countries.

Acknowledgements

References

1. Riboli E, Hunt KJ, Slimani N *et al.* (2002) European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* 5, 1113-1124.
2. Moreno LA, Gonzalez-Gross M, Kersting M *et al.* (2008) Assessing, understanding and modifying nutritional status, eating habits and physical activity in European adolescents: the HELENA (Healthy Lifestyle in Europe by Nutrition in Adolescence) Study. *Public Health Nutr* 11, 288-299.
3. Pigeot I, Barba G, Chadjigeorgiou C *et al.* (2009) Prevalence and determinants of childhood overweight and obesity in European countries: pooled analysis of the existing surveys within the IDEFICS Consortium. *Int J Obes (Lond)* 33, 1103-1110.
4. EFSA (2009) General principles for the collection of national food consumption data in the view of a pan-European dietary survey. *The EFSA Journal* 7, 1-51.
5. Brussaard JH, Lowik MR, Steingrimsdottir L *et al.* (2002) A European food consumption survey method--conclusions and recommendations. *Eur J Clin Nutr* 56 Suppl 2, S89-94.
6. de Boer E (2011) Rationale and methods of the EFCOVAL project. *Eur J Clin Nutr (in press)*.
7. Kaaks RJ (1997) Biochemical markers as additional measurements in studies of the accuracy of dietary questionnaire measurements: conceptual issues. *Am J Clin Nutr* 65, 1232S-1239S.
8. Ocke MC & Kaaks RJ (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, 1240S-1245S.
9. Bingham SA & Cummings JH (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.
10. Tasevska N, Runswick SA & Bingham SA (2006) Urinary potassium is as reliable as urinary nitrogen for use as a recovery biomarker in dietary studies of free living individuals. *J Nutr* 136, 1334-1340.

11. Slimani N, Bingham S, Runswick S *et al.* (2003) Group level validation of protein intakes estimated by 24-hour diet recall and dietary questionnaires against 24-hour urinary nitrogen in the European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study. *Cancer Epidemiol Biomarkers Prev* 12, 784-795.
12. Crispim SP, de Vries JH, Geelen A *et al.* (2011) Two non-consecutive 24 h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres - results from the European Food Consumption Validation (EFCOVAL) study. *Br J Nutr* 105, 447-458.
13. Hox J (2010) *Multilevel analysis. Techniques and applications.* 2nd ed. Utrecht: Routledge Academic.
14. Kaaks R, Plummer M, Riboli E *et al.* (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, 245S-250S.
15. Slimani N, Kaaks R, Ferrari P *et al.* (2002) European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study: rationale, design and population characteristics. *Public Health Nutr* 5, 1125-1145.
16. Ferrari P, Day NE, Boshuizen HC *et al.* (2008) The evaluation of the diet/disease relation in the EPIC study: considerations for the calibration and the disease models. *Int J Epidemiol* 37, 368-378.
17. Ferrari P, Roddam A, Fahey MT *et al.* (2009) A bivariate measurement error model for nitrogen and potassium intakes to evaluate the performance of regression calibration in the European Prospective Investigation into Cancer and Nutrition study. *Eur J Clin Nutr* 63 Suppl 4, S179-187.
18. Slimani N, Deharveng G, Charrondiere RU *et al.* (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Methods Programs Biomed* 58, 251-266.
19. Slimani N, Ferrari P, Ocke M *et al.* (2000) 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* 54, 900-917.
20. Slimani N, Deharveng G, Unwin I *et al.* (2007) The EPIC nutrient database project (ENDB): a first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study. *Eur J Clin Nutr* 61, 1037-1056.
21. FAO (2003) *Food energy - methods of analysis and conversion factors.* 77 Rome, Italy: Food and Agriculture Organization of the United Nations.
22. Crispim SP, Geelen A, Le Donne C *et al.* (2010) Dietary exposure to flavouring substances: from screening methods to detailed assessments using food consumption data collected with EPIC-Soft software. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess* 27, 433-446.
23. Bingham SA, Williams R, Cole TJ *et al.* (1988) Reference values for analytes of 24-h urine collections known to be complete. *Ann Clin Biochem* 25 (Pt 6), 610-619.
24. Johansson G, Bingham S & Vahter M (1999) A method to compensate for incomplete 24-hour urine collections in nutritional epidemiology studies. *Public Health Nutr* 2, 587-591.
25. Bingham SA (2003) Urine nitrogen as a biomarker for the validation of dietary protein intake. *J Nutr* 133 Suppl 3, 921S-924S.
26. Holbrook JT, Patterson KY, Bodner JE *et al.* (1984) Sodium and potassium intake and balance in adults consuming self-selected diets. *Am J Clin Nutr* 40, 786-793.
27. Haftenberger M, Lahmann PH, Panico S *et al.* (2002) Overweight, obesity and fat distribution in 50- to 64-year-old participants in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Public Health Nutr* 5, 1147-1162.
28. UNDP (2009) *Human Development Report. Overcoming barriers: Human mobility and development* New York: United Nations Development Programme.
29. Kim S, Haines PS, Siega-Riz AM *et al.* (2003) The Diet Quality Index-International (DQI-I) provides an effective tool for cross-national comparison of diet quality as illustrated by China and the United States. *J Nutr* 133, 3476-3484.
30. Duncan C, Jones K & Moon G (1998) Context, composition and heterogeneity: using multilevel models in health research. *Soc Sci Med* 46, 97-117.
31. Diez Roux AV (2002) A glossary for multilevel analysis. *J Epidemiol Community Health* 56, 588-594.
32. EFSA (2010) Statistical modelling of usual intake. Scientific report. EFSA-Q-2009-00841. 1-23.
33. Pryer JA, Vrijheid M, Nichols R *et al.* (1997) Who are the 'low energy reporters' in the dietary and nutritional survey of British adults? *Int J Epidemiol* 26, 146-154.
34. Subar AF, Kipnis V, Troiano RP *et al.* (2003) Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol* 158, 1-13.
35. Rumpler WV, Kramer M, Rhodes DG *et al.* (2008) Identifying sources of reporting error using measured food intake. *Eur J Clin Nutr* 62, 544-552.
36. Heerstrass DW, Ocke MC, Bueno-de-Mesquita HB *et al.* (1998) Underreporting of energy, protein and potassium intake in relation to body mass index. *Int J Epidemiol* 27, 186-193.
37. Lissner L, Troiano RP, Midthune D *et al.* (2007) OPEN about obesity: recovery biomarkers, dietary reporting errors and BMI. *Int J Obes (Lond)* 31, 956-961.

38. Heinz F, van Bakel M, Biessy C *et al.* (2011) Recovery biomarkers for protein and potassium intake show no differences in the effect of BMI on the validity of self-reported diet across 6 European countries. *Br J Nutr (in press).*
39. Halkjaer J, Olsen A, Bjerregaard LJ *et al.* (2009) Intake of total, animal and plant proteins, and their food sources in 10 countries in the European Prospective Investigation into Cancer and Nutrition. *Eur J Clin Nutr* 63 Suppl 4, S16-36.
40. Fox TA, Heimendinger J & Block G (1992) Telephone surveys as a method for obtaining dietary information: a review. *J Am Diet Assoc* 92, 729-732.
41. Casey PH, Goolsby SL, Lensing SY *et al.* (1999) The use of telephone interview methodology to obtain 24-hour dietary recalls. *J Am Diet Assoc* 99, 1406-1411.
42. Brustad M, Skeie G, Braaten T *et al.* (2003) Comparison of telephone vs face-to-face interviews in the assessment of dietary intake by the 24 h recall EPIC SOFT program--the Norwegian calibration study. *Eur J Clin Nutr* 57, 107-113.
43. Crispim SP, Geelen A, Siebelink E *et al.* (2011) Design aspects of 24-h recall assessments may affect the estimates of protein and potassium intake in dietary surveys. *submitted.*
44. Crispim SP, Geelen A, de Vries JH *et al.* (2011) Biomarker-based evaluation of two 24-h recalls (EPIC-Soft) for estimating and comparing fish, fruit and vegetable intakes across five European centers: results from the European Food Consumption Validation (EFCOVAL) study. *Eur J Clin Nutr (in press).*

APPENDIX 1 - Report on the comparison of laboratory measurements

With the purpose of harmonizing biomarker data used in the paper 'Bias in protein and potassium intake collected with 24-h recalls (EPIC-Soft) is rather comparable across European populations', a calibration study was conducted among laboratories that performed chemical analyses of urine samples from the EPIC and EFCOVAL studies.

Nutrients and specimens in the calibration study: The amount of urinary nitrogen and urinary potassium of 45 samples from the EPIC cohort were determined at both the MRC Dunn Clinical Nutrition Centre in Cambridge (EPIC) in the early 2000's and at the Division of Human Nutrition at Wageningen University (EFCOVAL) in 2008.

Methodology: Agreement between the two biomarker measurements determined by the Cambridge and Wageningen laboratories was assessed using Bland-Altman plots and Pearson's coefficients of correlation. The paired t-test was used to test for significant differences ($p<0.05$) between the mean biomarker measurements from the two laboratories. In addition, comparability of laboratory methods used in EPIC and EFCOVAL labs was further substantiated by evaluating standard reference materials and quality control procedures (e.g., inter-laboratory proficiency tests) of the lab measurements.

Linear regression analyses were carried out to generate calibration equations between the biomarker measurements from the two laboratories. Protein and potassium values which were +/- 3SD from the mean were considered outliers (n = 1 for nitrogen and for potassium) and excluded from the regression analyses.

Results: A good agreement and correlation was seen between the measurements from the two laboratories based on the Bland-Altman plots (**Figure A1 and A2**) and Pearson's correlation coefficients (r>0.97), respectively (Table A1). No significant differences were seen between the mean nitrogen (p=0.10) and potassium (p=0.68) measurements from the two laboratories.

Based on those results, we judged that calibration of data between EFCOVAL and EPIC studies was not necessary and original biomarker data of the two studies was used in the main analyses of the paper.

Because the calibration study was limited by the rather small sample size, we still estimated the calibration equations and performed a sensitivity analysis using the calibrated data in the multilevel analysis. The EFCOVAL data was calibrated using the following regression formulas:

$$\text{EFCOVAL\_N}_{\text{calibrated}} = 0.92 \text{ x EFCOVAL\_N} + 0.03$$

$$\text{EFCOVAL\_K}_{\text{calibrated}} = 0.92 \text{ x EFCOVAL\_K} + 1.87$$

Standard reference materials and inter-laboratory proficiency tests of the laboratory procedures in the two laboratories indicated no discrepancies between the analyses performed. Thus, there was no evidence to opt for having one or the other laboratory as reference in the calibration procedure, and the Cambridge lab (from the EPIC study) was chosen.

Table A1 – Comparison of means (SD) and Pearson's correlation coefficients between biomarker measurements determined by the Wageningen and Cambridge laboratories in the calibration study

Biomarker	_n_	Wageningen	Cambridge	p*	r†
Urinary nitrogen (g/100ml)	45	0.70 (0.3)	0.68 (0.3)	0.10	0.97
Urinary potassium (mmol/L)	45	42.1 (17.3)	40.6 (16.2)	0.68	0.99

* Paired t-test
† Pearson correlation
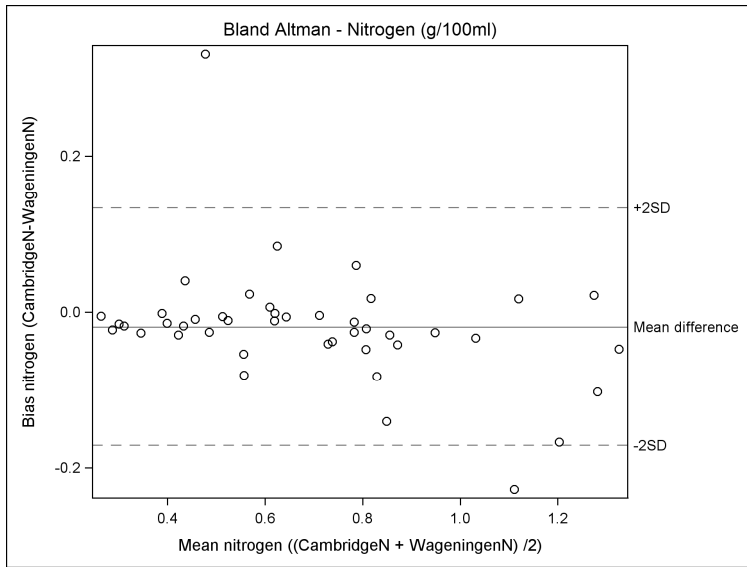
Figure A1 - Bland-Altman plot for nitrogen measurements from the Cambridge and Wageningen laboratories.
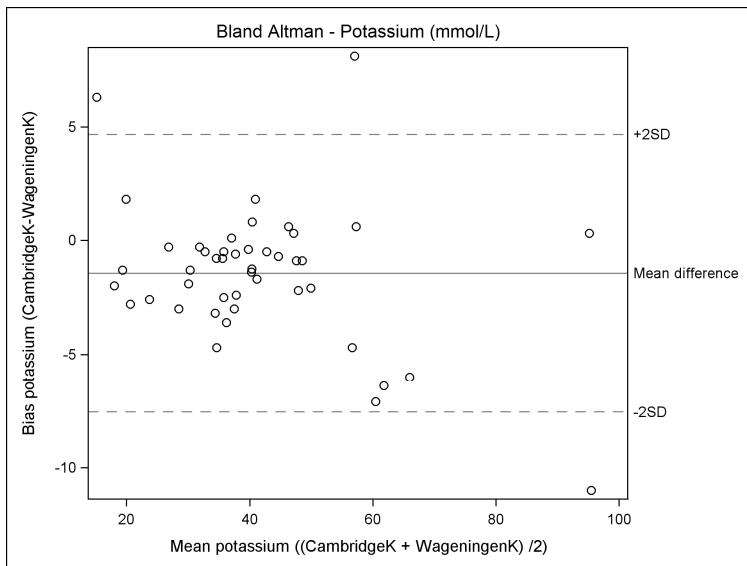
Figure A2 - Bland-Altman plot for potassium measurements from the Cambridge and Wageningen laboratories.

APPENDIX 2 – Specification of models used in the multilevel approach

The following regression model represents model iii (random intercepts with individual- and centre-level explanatory variables) in the assessment:

$$Y_{ij} = \alpha_j + \beta_1 X_{1ij} \ldots \beta_n X_{nij} + \gamma_1 Z_{1j} \ldots \gamma_n Z_{nj} + e_{ij}$$
$$\alpha_j = \alpha + u_{0j}$$

$$u_{0j} \sim N(0, \Omega_u)$$
$$e_{ij} \sim N(0, \sigma_e^2)$$

Where,

j = the index for the centres (j = 1,…,N)

i = the index for the individuals within the centres (i = 1,…,nj)

Y_{ij} = log ratio between dietary intake and biomarker for ith individual in the jth centre

α = the overall mean of log ratio between intake and biomarker across all centres

$\beta_{1...}\beta_n$ = effects of individual explanatory variables $X_{1ij...} X_{nij}$

$\gamma_{1...}\gamma_n$ = fixed effects of the centre-level explanatory variables $Z_{1j...} Z_{nj}$

u_{0j} = centre-level random effects on the mean of the intercept of Y

e_{ij} = residual error term, assumed to have a mean of zero and a variance (σ_e^2 = individual random effect)

Thus, this model has fixed-effect parameters $(\alpha, \beta_n, \gamma_n)$ as well as zero-mean random coefficients (u_{0j}, e_{ij}).

In model ii (random intercepts with only individual explanatory variables), the coefficients $\gamma_{1...}\gamma_n$ of the centre-level variables $Z_{1j...} Z_{nj}$ are zero. Model I additionally constrained to zero the coefficients $\beta_{1...}\beta_n$ from individual variables $X_{1ij...} X_{nij}$.

Chapter 4

Biomarker-based evaluation of two 24-h recalls for comparing usual fish, fruit and vegetable intakes across European centres in the EFCOVAL study

Sandra P. Crispim, Anouk Geelen, Olga W. Souverein, Paul J. M. Hulshof, Jiri Ruprich, Marcela Dofkova, Inge Huybrechts, Willem De Keyzer, Inger T. L. Lillegaard, Lene F. Andersen, Lionel Lafay, Anne-Sophie Rousseau, Marga C. Ocke, Nadia Slimani, Pieter van 't Veer and Jeanne H. M. de Vries

Abstract

Background: A standardised methodology is important to enable consistent monitoring of dietary intake across European countries. For this reason, we evaluated the comparability of the assessment of usual food intake collected with two non-consecutive computerised 24-h recalls and a food propensity questionnaire (FPQ) between five European centres.

Subjects/Methods: Two 24-h recalls using EPIC-Soft software were performed to determine fish, fruit and vegetable consumption of 600 adults in Belgium (BE), the Czech Republic (CZ), France (FR), the Netherlands (NL), and Norway (NO) in a validation study. An FPQ was used to identify non-consumers. Information from the 24-h recalls and FPQ were used to estimate individual usual food intake by the Multiple-Source-Method (MSM). Blood samples were drawn to determine fatty acids in phospholipids and serum carotenoids as biomarkers of fish, and fruit plus vegetable (FV) intake, respectively.

Results: The pooled correlation between usual fish intake and EPA plus DHA in phospholipids was 0.19 in men and 0.31 in women (p for heterogeneity > 0.50) and centre-specific correlations ranged between 0.08 (CZ) and 0.28 (BE and NO) in men and between 0.19 (BE) and 0.55 (FR) in women. For usual FV intake, the pooled correlation with serum carotenoids was 0.31 in men and 0.40 in women (p for heterogeneity > 0.10); the centre-specific correlations varied between 0.07 (NO) and 0.52 (FR) in men and between 0.25 (NL) and 0.45 (NO) in women.

Conclusion: Two standardised 24-h recalls using EPIC-Soft and an FPQ appeared to be appropriate to rank individuals according to their fish and fruit and vegetable intake in a comparable way between five European centres.

Introduction

Dietary data from national food consumption surveys are useful to develop and evaluate policies on nutrition and food safety. In Europe, national food consumption data are important to assess the variability of food patterns among different countries. However, European countries performing national surveys use different methodologies such as 24-h dietary recalls and food diaries to collect dietary data[1]. In addition, differences exist in a number of aspects such as the food classification

system used across countries. For instance, olives can be considered as a fruit in one food classification and as a vegetable in another[2].

European countries are expected to provide similar dietary indicators if harmonised food consumption data are collected in future national surveys[3]. For this reason, the European Food Consumption Survey Method (EFCOSUM) consortium recommended the collection of food consumption data using two non-consecutive standardised 24-h recalls as the most appropriate method in future pan-European surveys[4]. Furthermore, the consortium recommended the use of EPIC-Soft software for standardisation and defined a set of dietary components including, besides specific nutrients, vegetables, fruits, bread, and fish and shellfish[4-7] to serve as nutritional indicators.

Because the use of a standardised and valid methodology is crucial to enable consistent monitoring of diet across European countries, the European Food Consumption Validation (EFCOVAL) consortium aimed to further develop and validate the methodology proposed for pan-European dietary monitoring. To that end, our previous work showed that two non-consecutive days of dietary intake collected with 24-h recalls (EPIC-Soft) were considered sufficiently valid for comparing usual protein and potassium intake between five European centres[8]. In the present study, we intended to further evaluate the dietary intake collected with respect to the comparability of food group assessment across different European populations. A food propensity questionnaire (FPQ) was included in the assessment to offer covariate information in complementing the 24-h recalls during the estimation of usual intake of food groups[9].

Assessment of intake of fruits & vegetables and fish & shellfish can be evaluated using, respectively, serum carotenoids[10-12] and n-3 fatty acids in e.g., phospholipids[13-15] as concentration biomarkers. Concentration biomarkers are related to dietary intake but not as directly as recovery biomarkers because their concentrations are the result of complex metabolic processes[16]. Therefore, their use in validation studies is restricted to their associations, commonly as correlations, with self-reported dietary intakes. The strength of these correlations is often lower (<0.6) than that of recovery biomarkers[17].

This paper aims to evaluate and compare the assessment of ranking of individuals according to their usual fish and fruit & vegetable consumption estimated with two non-consecutive standardised 24-h recalls and an FPQ between five selected centres

in Europe, using fatty acids in phospholipids and serum carotenoids as biomarkers of intake, respectively.

Subjects and Methods

Study population

The study population consisted of 297 men and 303 women, between 45 to 65 y old, from five selected centres from Belgium, the Czech Republic, France (Southern part), the Netherlands, and Norway. These centres were chosen to represent the large diversity of food patterns across Europe. Participants were recruited by convenience sampling and were healthy individuals representing all educational levels. Eligible participants were able to read and speak the national language, not following prescribed dietary therapy, not pregnant or lactating, and not enrolled in another study at the same period. In addition, we did not allow subjects in the study who were donating blood or plasma during or less than four weeks before the study, institutionalised persons or more than one member of the same household. More details about the study populations, including recruitment and sampling procedures are described elsewhere[8].

Study design

The period of data collection was from April to July 2007 in the Netherlands and from October or November 2007 to April 2008 in the other four centres. Ethical committees in each centre approved the research protocol and participants signed an informed consent. At the beginning of the study, each participant filled out a screening and a general questionnaire with questions about lifestyle and food habits, including type and frequency of used supplements during the previous three months. Participants were then weighed and had their height measured in the study centres following standardised procedures. They also underwent a non-fasting venipuncture. Then, we collected two non-consecutive 24-h recalls with approximately one month in between. The time interval between blood sampling and the first 24-h recall was on average less than a week for all centres, except in the Czech Republic where the average was two weeks.

Dietary data

We collected the two 24-h recalls using EPIC-Soft software, version 9.16[18; 19]. In brief, EPIC-Soft is a computer assisted dietary tool that follows a standardised procedure to minimize measurement errors when describing, quantifying, probing and calculating food intakes across countries[18]. The two 24-h recalls were collected using two modes of administration: one by phone and one face-to-face. A randomisation schedule was created to consider a random order of the two modes of administration as well as the inclusion of all days of the week equally among the subjects. This randomisation allowed the same person to have the same day of the week recalled for both interviews by chance.

Interviewers in each centre were nutritionists or dietitians who were trained by qualified local trainers in interviewing skills and working with EPIC-Soft. Centres were allowed to organize their data collection in the same way as they would in a future performance of their national monitoring survey. For example, dietary recalls in Belgium, the Czech Republic and the Netherlands were not conducted on Sundays and, therefore, Saturday's intake was recalled two days later on Mondays. Furthermore, interviewees were permitted to check food packages and household measures in their home for detailed information during the phone interview while this was not possible during the face-to-face interview at the study centre. All centres used an existing version of EPIC-Soft software, which had already been used in a national survey or within the EPIC study, except the Czech Republic for which a new version was developed. Methods of estimation of portion size included household measures, weight/volume, standard units and portions, and photographs in a picture book. Furthermore, dietary supplement-use information of the recalled day was collected at the end of the 24-h recall interview. If a supplement was taken, subjects reported on the physical state (e.g., capsule), the number of units per consumption occasion, and the frequency. If known, the brand name was also reported. In addition, an FPQ including one question per food group was used to identify frequency of usual consumption of fish, fruits and vegetables over the past year.

Food groups were defined as suggested by EFCOSUM[2]. To this end, the foods as reported by the recalls were regrouped by including or excluding specific subgroups of the EPIC-Soft food classification[18]. Fruit intake was defined not to include nuts, seeds, olives and fruit juices other than freshly squeezed juices. Vegetable intake was defined to include herbs but not pulses and potatoes. Fish intake was defined to include shellfish. Fish was classified in lean fish (<4g of fat/100g of edible part such

as cod, tuna, tilapia, and carp) or fatty fish (>= 4g of fat/100g of edible part such as salmon, herring, and mackerel) using country-specific food composition tables.

Venipuncture and biomarkers

We provided participants with guidelines to have a low-fat breakfast before blood sampling. We requested subjects to rest before a trained lab technician drew blood (2x9ml) from the antecubital vein. The blood was then allowed to clot for 30 minutes at room temperature (20-22°C) and centrifuged for 15 minutes at 1200 xg. Serum samples from each subject were aliquoted into cryo-tubes for storage at –80°C until shipment on dry ice to the central laboratory at Wageningen University, where analyses took place.

After thawing and mixing the samples, fatty acids in the phospholipid fraction were measured by extracting and separating the lipid classes. Briefly, the phospholipid fraction was separated from the other lipid classes on an aminopropyl column according to the procedure described by Kaluzny et al.[20]. Fatty acid methyl ester profile was prepared according to Metcalfe et al.[21]. Serum carotenoids were analysed as described by Khan et al.[22]. This method does not adequately separate lutein and zeaxanthin; consequently these two carotenoids are presented together. Total cholesterol was measured spectrofotometrically on a Synchron LX20 clinical analyzer (Beckman Coulter, Mijdrecht, the Netherlands).

The fatty acid composition of phospholipids was used as concentration biomarker of fish intake, namely the percentage of eicosapaentaenoic acid (EPA; C20:5n-3) and docosahexaenoic acid (DHA; C22:6n-3) in relation to the total area of measured fatty acids (=36 fatty acids). The sum of serum carotenoids, including α-carotene, β-cryptoxanthin, β-carotene, lutein, and zeaxanthin, was used as marker of fruit and vegetable intake. To further explore the correlations across centres, both α-carotene and β-carotene were used as biomarkers of fruits & vegetables. Likewise, β-cryptoxanthin was used as biomarker of fruit intake[11; 23] and lutein plus zeaxanthin of vegetable intake[23; 24]. The intra-assay precision, expressed as coefficient of variation (CV), of EPA and DHA was less than 4% and of individual carotenoids between 5 and 8%.

Supplement use

Supplement users were identified as those who reported taking any supplements containing EPA, DHA, or carotenoids on one of the recalled days or during the past three months according to the general questionnaire. To identify the presence of fatty acids and carotenoids in the reported supplements, we: (1) searched companies' websites, (2) visited drugstores, (3) searched other sources such as national databases.

Statistical analyses

The statistical analyses were done for men and women separately. For evaluating the ranking of individuals, we calculated Pearson's correlation coefficients between the average intake of foods groups based on the two days and their respective biomarkers per centre. In addition, adjusted Pearson's correlation coefficients were estimated between usual intake of food groups and respective biomarker using partial correlations. For adjusted correlations, we used usual intake corrected for within-person variability together with the information from the FPQ (Adjusted$_1$), as estimated by the Multiple-Source-Method (MSM)[25]. We further corrected for the following covariables that were expected to be associated with the intake or excretion based on pre-existing knowledge: age, BMI, education level, alcoholic beverage intake, and smoking status (Adjusted$_2$). Fruit and vegetable intake analyses were also corrected for total serum cholesterol[26]. For the calculation of the correlations, intake of foods and concentrations of biomarkers were log-transformed to improve normality of the observed distributions. Considering that improvements in the normality of the distribution may have not been achieved, Spearman's correlations were also computed. However, only when conclusions based on Pearson's and Spearman's correlations differed, we presented the latter. Confidence intervals of the correlations were obtained using the Fisher Z-transformation[27].

The MSM is a statistical method for estimating usual dietary intake of nutrients and foods, including episodically consumed foods, for populations as well as individuals. In contrast to many other statistical methodologies, MSM first estimates individual usual intakes rather than constructing directly the population distributions of usual intake. The method can make use of covariate information such as consumption frequency information from an FPQ to improve the modelling of consumption probability and intake amount[25].

Pooled correlations of the five centres were calculated by first converting correlations into a standard normal metric (Fisher's r-to-Z transformation). Next, the pooled average was calculated, in which each transformed correlation coefficient was weighted by its inverse variance, followed by the back transformation[27]. Cochrane Q-test was used for testing heterogeneity of the pooled correlation[28].

The estimated intake did not include the amounts of EPA, DHA or carotenoids originating from supplement use. To help interpreting the main results, biomarker levels were presented separately for the total sample, users and non-users of supplements. Given the small number of subjects in each group, men and women were grouped together to optimize this part of the analysis.

Analyses were performed using SAS statistical package, version 9.1 (SAS Institute Inc., Cary, NC).

Results

Regarding the characteristics of our study population, the mean BMI of the French men and women was somewhat lower than those of the other four centres (**Table 1**). A larger prevalence of smokers was observed in Czech men (33%) and Norwegian women (23%) than in subjects of the other centres. Moreover, subjects with a low educational level were less represented than subjects with a moderate or high level, especially in Norwegian men. Belgian men reported the highest intake of alcoholic beverages (average of 30.2 g/day) and Czech women the lowest (average of 6.3 g/day). Furthermore, the total serum cholesterol concentration of the subjects did not vary substantially across the five centres in both genders.

In all centres, each day of the week was represented by between 12 to 17% of the 24-h recalls, except in France, where Saturday was less representative (8%) and Thursday more (19%), and in the Czech Republic, where almost 19% of the interviews were about the intake of a Sunday. The interval between the first and the second 24-h recalls was at least three weeks for all centres.

Mean fish intake was highest in the Norwegian centre; it was 3-4 times higher than the mean intakes in the Czech Republic or the Netherlands **(Table 2)**. Likewise, the highest mean percentage of EPA plus DHA in phospholipids was seen in Norway and the lowest in the Czech Republic and the Netherlands.

Table 1– Characteristics of subjects from the five European centres in the EFCOVAL validation study

	Men					Women				
	BE*	**CZ**	**FR**	**NL**	**NO**	**BE**	**CZ**	**FR**	**NL**	**NO**
n	63	58	54	60	62	60	60	59	62	62
Age (years)	54 ± 0.7	55 ± 0.9	56 ± 0.7	57 ± 0.6	55 ± 0.8	55 ± 0.7	55 ± 0.8	55 ± 0.8	55 ± 0.7	54 ± 0.8
BMI (kg/m²)	27.2 ± 0.5	27.9 ± 0.5	25.5 ± 0.4	26.5 ± 0.5	26.4± 0.3	25.2 ± 0.5	25.0 ± 0.5	23.2 ± 0.4	25.5 ± 0.6	24.8 ± 0.5
Smoking (% of total)										
Current	15.9	32.8	14.8	11.7	19.4	13.3	10.0	8.5	3.2	22.6
Former	47.6	17.2	25.9	61.6	40.3	28.3	21.7	23.7	37.1	38.7
Never	36.5	50.0	59.3	26.7	40.3	58.4	68.3	67.8	59.7	38.7
Education (% of total)										
Low	15.9	20.7	25.9	20.0	3.2	16.7	16.6	35.6	24.2	16.1
Intermediate	23.8	24.1	24.1	20.0	30.7	25.0	46.7	27.1	40.3	19.4
High	60.3	55.2	50.0	60.0	66.1	58.3	36.7	37.3	35.5	64.5
Alcoholic beverage intake (g/day)	30.2 ± 4.2	17.8 ± 3.4	15.1 ± 2.5	27.1± 3.4	16.5 ± 2.8	17.3 ± 2.7	6.3 ± 1.3	6.9 ± 1.3	12.3 ± 2.0	10.7 ± 2.1
Total serum cholesterol (mmol/L)	5.5 ± 0.1	5.5 ± 0.1	5.8 ± 0.1	5.5 ± 0.1	5.8 ± 0.1	5.5 ± 0.1	5.6 ± 0.1	5.8 ± 0.1	5.6 ± 0.1	5.9 ± 0.1

* BE=Belgium, CZ=the Czech Republic, FR=France, NL=the Netherlands, NO=Norway. Results in mean ± standard error, unless otherwise stated.

Evaluation of fish, fruit & vegetable intake collected with 2x24-h

A low proportion of fatty to lean fish intake (excluding fish products) was observed in Czech and French men (ratio ≤ 0.3) as compared to the other three centres, where the ratio ranged between 0.9 and 1.1. A high proportion of fatty to lean fish (ratio 4.3) was observed in Dutch women. Shellfish and roe products contributed between 13 to 26% of total fish intake in Belgium and France, whereas the Czech Republic did not report any consumption of it (data not shown). The percentage of fish consumers identified by 24-h recalls was lower (**Table 2**) than by FPQ, which showed nearly 95% of consumers in all centres (data not shown).

The lowest crude correlation between fish intake and the biomarker (**Table 3**) was observed in the Czech Republic in both genders (r=-0.04 in men and 0.24 in women). When we analysed usual fish intake by adjusting intakes for within person variability and including FPQ data (**See table 3 *Adjusted$_1$***), evident improvement of correlations was seen across the centres, with the exception of the correlation coefficients in Belgian and Czech women that decreased, respectively, from 0.34 to 0.19 and from 0.24 to 0.21. Further adjustment of the correlations for possible confounders did not explain the differences across centres (**See table 3 *Adjusted$_2$***). Nevertheless, although the adjusted correlation for fish intake was still considerably lower in Czech men (r=0.08) than in the other centres, no statistically significant heterogeneity of correlations was found between the centres (p>0.20 in both gender).

The largest average intake of both fruit and vegetables was seen in France for both men and women (**Table 2**). While the lowest fruit intake was reported in Belgium for men and in Norway for women, the lowest vegetable intake was observed in the Czech Republic and Norway for both genders. Additionally, there were large differences in the types of fruit and vegetable consumed across centres for both men and women. Cooked fruits and vegetables were less consumed by Czech subjects (~25%) and more by French (~45%). The average amounts of citrus fruits consumed were larger in the Czech Republic and France than in the other three centres. Leafy vegetables were clearly less consumed in the Czech Republic and Norway than in Belgium, France, and the Netherlands. The percentage of fruit & vegetable consumers identified by the 24-h recalls was the same as by the FPQ, and was nearly 100%. In relation to the biomarker, the highest mean concentration of carotenoids was observed in France and the lowest in the Czech Republic.

Crude correlation coefficients of fruit & vegetable intake with the sum of carotenoids were between 0.05 in the Norwegian men to 0.54 in the Czech men **(Table 3)**.

Table 2 – Intakes of fish, fruits and vegetables estimated from the 2x24-h recalls and related biomarkers (mean ± standard error) of 600 participants in the EFCOVAL validation study

		Men				
		Belgium	the Czech Republic	France	the Netherlands	Norway
Fish intake (g/day)		52 ± 7.6	20 ± 5.4	47 ± 8.0	25 ± 4.8	82 ± 10.5
% Fish consumers		63	26	59	42	77
% EPA+DHA of total fatty acids in phospholipids		5.3 ± 0.2	4.4 ± 0.1	5.2 ± 0.2	4.5 ± 0.2	7.3 ± 0.2
Ratio fatty/lean fish intake*		1.0	0.2	0.3	1.1	0.9
Fruit intake (g/day)		163 ± 18.4	207 ± 23.2	228 ± 27.5	198 ± 21.4	199 ± 23.7
Vegetable intake (g/day)		220 ± 13.7	162 ± 15.9	222 ± 19.4	194 ± 12.1	168 ± 13.5
% cooked fruits and vegetables		50	27	35	38	31
Subgroups† (g/day)	Citrus fruits	39 ± 7.7	61 ± 10.8	55 ± 11.6	22 ± 6.4	43 ± 8.9
	Non-citrus fruits	107 ± 15.7	132 ± 16	139 ± 17.9	159 ± 17.2	154 ± 20.1
	Leafy vegetables	34 ± 5.2	7± 3.8	39 ± 7.1	35 ± 5.6	10 ± 2.4
	Fruiting vegetables	63 ± 9.3	54 ± 10	77 ± 10.3	69 ± 7.9	56 ± 7.7
	Root vegetables	26 ± 4.4	29 ± 4.4	38 ± 6.6	10 ± 2.7	30 ± 6.4
	Cabbages	25 ± 6.6	44 ± 8.5	18 ± 5.4	27 ± 7.0	35 ± 6.4
	Onion and garlic	33 ± 5.3	14 ± 2.1	20 ± 3.4	24 ± 4.0	9 ± 1.6
Sum of serum carotenoids‡ (mcg/100ml)		77 ± 4.8	60 ± 3.1	121 ± 8.5	87 ± 5.4	77 ± 3.2
Lutein + zeaxanthin (mcg/100ml)		28.5 ± 1.3	26.3 ± 1.4	42.2 ± 2.7	31.5 ± 1.7	29.5 ± 2.2
β-cryptoxanthin (mcg/100ml)		15.3 ± 1.1	12.8 ± 1.1	22.4 ± 2.1	16.5 ± 1.3	13.3 ± 0.8
α-carotene (mcg/100ml)		7.2 ± 0.7	5.0 ± 0.5	12.9 ± 1.2	6.4 ± 1.6	7.3 ± 0.6
β-carotene (mcg/100ml)		26.0 ± 2.7	15.9 ± 1.4	42.9 ± 3.7	32.7 ± 2.9	26.7 ± 1.4

	Women				
	Belgium	**the Czech Republic**	**France**	**the Netherlands**	**Norway**
Fish intake (g/day)	32 ± 4.7	24 ± 5.0	43 ± 6.6	22 ± 4.6	65 ± 8.8
% Fish consumers	65	40	64	40	82
% EPA+DHA of total fatty acids in phospholipids	5.3 ± 0.2	4.1 ± 0.1	5.8 ± 0.2	4.8 ± 0.2	7.0 ± 0.3
Ratio fatty/lean fish intake*	1.2	0.8	0.8	4.3	0.9
Fruit intake (g/day)	206 ± 18.2	226 ± 20.1	265 ± 21.7	257 ± 19.0	194 ± 18.7
Vegetable intake (g/day)	215 ± 14.7	157 ± 12.8	254 ± 18.5	174 ± 10.0	166 ± 12.6
% cooked fruits and vegetables	43	24	34	29	25
Subgroups† (g/day) Citrus fruits	49 ± 10.6	67 ± 10.5	82 ± 11.8	41 ± 9.6	49 ± 8.1
Non-citrus fruits	150 ± 14.3	157 ± 17.2	162 ± 16.5	180 ± 14.6	131 ± 16
Leafy vegetables	41 ± 8.0	6 ± 2.3	59 ± 8.1	29 ± 4.4	17 ± 3.9
Fruiting vegetables	60 ± 7.3	51± 6.7	77 ± 10.1	63 ± 5.9	61 ± 8.2
Root vegetables	29 ± 4.7	29 ± 5.0	45 ± 6.9	21 ± 5.3	28 ± 5.2
Cabbages	29 ± 6.8	42 ± 7.1	16.4 ± 4.1	27 ± 5.5	25 ± 5.7
Onion and garlic	38 ± 5.7	12 ± 1.7	16.3 ± 2.4	13 ± 2.5	9 ± 1.7
Sum of serum carotenoids‡ (mcg/100ml)	102 ± 6.1	81 ± 5.4	151 ± 8.6	108 ± 5.9	100 ± 5.8
Lutein + zeaxanthin (mcg/100ml)	37.0 ± 2.2	26.7 ± 1.7	49.0 ± 2.6	36.2 ± 2.1	33.4 ± 2.0
β-cryptoxanthin (mcg/100ml)	21.3 ± 1.9	14.4 ± 1.3	32.6 ± 3.7	24.5 ± 1.9	14.7 ± 0.8
α-carotene (mcg/100ml)	9.4 ± 1.0	8.8 ± 1.2	15.0 ± 1.2	7.1 ± 0.6	11.0 ± 1.0
β-carotene (mcg/100ml)	34.1 ± 2.9	30.8 ± 3.2	54.5 ± 3.7	39.8 ± 2.9	40.6 ± 3.1

* Excluding fish products.

† Fruit and vegetable subgroups most contributing to the main food group intake. Other fruits not presented included mixed fruits (e.g., dry fruits). Other vegetables not presented included mushrooms, mixed, grain and stalk vegetables.

‡ α-carotene+β-cryptoxanthin+β-carotene+lutein+zeaxanthin.

Table 3 – Pearson's correlation coefficients (r) with confidence intervals (CI) between intakes of fish, fruit and vegetable from 24-h recalls and related biomarkers of men participants in the EFCOVAL validation study.

	Centre	Crude		Adjusted₁*		Adjusted₂†	
		r	CI	r	CI	r	CI
Fish vs. EPA+ DHA	Belgium	0.11	(-0.14,0.35)	0.32	(0.07,0.52)	0.28	(0.03,0.50)
	the Czech Republic	-0.04	(-0.29,0.23)	0.05	(-0.21,0.30)	0.08	(-0.19,0.34)
	France	0.22	(-0.05,0.46)	0.34	(0.08,0.56)	0.27	(-0.02,0.51)
	the Netherlands	0.13	(-0.13,0.37)	0.23	(-0.03,0.46)	0.21	(-0.06,0.46)
	Norway	0.22	(-0.04,0.44)	0.27	(0.02,0.49)	0.28	(0.01,0.50)
	Pooled‡	0.13	(0.01, 0.25)	0.24	(0.13,0.36)	0.19	(0.07,0.30)
Fruit & vegetables vs. sum of carotenoids§	Belgium	0.38	(0.15,0.57)	0.38	(0.14,0.57)	0.36	(0.11,0.57)
	the Czech Republic	0.54	(0.32,0.70)	0.47	(0.24,0.65)	0.52	(0.28,0.69)
	France	0.43	(0.18,0.62)	0.50	(0.26,0.67)	0.43	(0.16,0.64)
	the Netherlands	0.32	(0.06,0.53)	0.20	(-0.06,0.44)	0.16	(-0.12,0.41)
	Norway	0.05	(-0.20,0.30)	0.06	(-0.20,0.31)	0.07	(-0.20,0.33)
	Pooled	0.35	(0.23,0.46)	0.33	(0.21,0.45)	0.31	(0.20,0.41)
Vegetable intake vs. sum of carotenoids	Belgium	0.21	(-0.04,0.44)	0.24	(-0.02,0.45)	0.20	(-0.06,0.44)
	the Czech Republic	0.47	(0.23,0.65)	0.55	(0.34,0.71)	0.63	(0.42,0.77)
	France	0.37	(0.11,0.58)	0.31	(0.04,0.53)	0.33	(0.04,0.56)
	the Netherlands	0.09	(-0.17,0.33)	0.01	(-0.24,0.27)	-0.01	(-0.28,0.26)
	Norway	0.16	(-0.10,0.39)	0.12	(-0.14,0.36)	0.06	(-0.21,0.32)
	Pooled	0.26	(0.14,0.38)	0.26	(0.14,0.37)	0.24	(0.12,0.34)
Fruit intake vs. sum of carotenoids	Belgium	0.27	(0.02,0.48)	0.21	(-0.04,0.43)	0.22	(-0.04,0.46)
	the Czech Republic	0.14	(-0.12,0.39)	0.28	(0.02,0.50)	0.30	(0.02,0.53)
	France	0.27	(0.01,0.50)	0.49	(0.25,0.67)	0.54	(0.29,0.72)
	the Netherlands	0.29	(0.03,0.50)	0.23	(-0.03,0.46)	0.16	(-0.12,0.41)
	Norway	-0.16	(-0.39,0.10)	-0.07	(-0.32,0.19)	-0.07	(-0.33,0.20)
	Pooled	0.16	(0.05,0.28)	0.23	(0.11,0.34)	0.17	(0.06,0.28)

Table 3 continues…

		Crude		Adjusted₁*		Adjusted₂†	
		r	CI	r	CI	r	CI
Fruit & vegetables vs. α-carotene	Belgium	0.41	(0.18,0.60)	0.45	(0.23,0.63)	0.39	(0.14,0.59)
	the Czech Republic	0.46	(0.23,0.64)	0.43	(0.19,0.61)	0.48	(0.23,0.66)
	France	0.55	(0.33,0.71)	0.59	(0.38,0.74)	0.50	(0.24,0.69)
	the Netherlands	0.23	(-0.03,0.46)	0.16	(-0.11,0.40)	0.11	(-0.17,0.37)
	Norway	0.11	(-0.15,0.35)	0.12	(-0.14,0.36)	0.09	(-0.19,0.34)
	Pooled	0.36	(0.27,0.44)	0.36	(0.24,0.48)	0.31	(0.20,0.41)
Fruit & vegetables vs. β-carotene	Belgium	0.44	(0.22,0.62)	0.43	(0.21,0.61)	0.38	(0.13,0.58)
	the Czech Republic	0.46	(0.22,0.64)	0.41	(0.17,0.60)	0.42	(0.16,0.62)
	France	0.45	(0.20,0.64)	0.52	(0.29,0.69)	0.39	(0.11,0.61)
	the Netherlands	0.17	(-0.09,0.41)	0.03	(-0.23,0.29)	-0.05	(-0.32,0.23)
	Norway	0.18	(-0.07,0.41)	0.13	(-0.13,0.37)	0.07	(-0.20,0.33)
	Pooled	0.33	(0.21,0.44)	0.31	(0.20,0.43)	0.26	(0.15,0.37)
Vegetable intake vs lutein+zeaxanthin	Belgium	-0.11	(-0.35,0.14)	-0.07	(-0.31,0.18)	-0.06	(-0.32,0.20)
	the Czech Republic	0.37	(0.12,0.57)	0.48	(0.25,0.66)	0.52	(0.28,0.69)
	France	0.23	(-0.04,0.47)	0.24	(-0.03,0.48)	0.22	(-0.07,0.48)
	the Netherlands	0.15	(-0.11,0.39)	0.07	(-0.19,0.32)	0.04	(-0.23,0.31)
	Norway	0.08	(-0.18,0.32)	0.06	(-0.19,0.31)	0.01	(-0.25,0.28)
	Pooled	0.14	(0.02,0.26)	0.16	(0.04,0.28)	0.17	(0.06,0.29)
Fruit intake vs β-cryptoxanthin	Belgium	0.38	(0.14,0.57)	0.28	(0.03,0.49)	0.35	(0.09,0.55)
	the Czech Republic	0.17	(-0.09,0.41)	0.31	(0.05,0.52)	0.35	(0.09,0.57)
	France	0.16	(-0.12,0.41)	0.41	(0.15,0.61)	0.48	(0.22,0.67)
	the Netherlands	0.26	(0.01,0.48)	0.18	(-0.09,0.42)	0.17	(-0.10,0.43)
	Norway	0.18	(-0.08,0.41)	0.03	(-0.23,0.28)	0.08	(-0.19,0.34)
	Pooled	0.23	(0.12,0.35)	0.24	(0.12,0.36)	0.21	(0.09,0.32)

* Adjusted for within person variability by Multiple-Source-Method, taking into account the food propensity questionnaire.
† Adjusted₁ + adjusted for age, BMI, educational level, alcoholic beverage, smoking status using Partial Pearson correlations. Cholesterol levels was also included in the fruit and vegetable analysis.
‡ P-value for heterogeneity was > 0.10 for all analyses in the table.
§ α-carotene+β-cryptoxanthin+β-carotene+lutein+zeaxanthin.

Table 4 – Pearson's correlation coefficients (r) with confidence intervals (CI) between intakes of fish, fruit and vegetable from 24-h recalls and related biomarkers of women participants in the EFCOVAL validation study.

	Centre	Crude		Adjusted$_1$*		Adjusted$_2$†	
		r	CI	r	CI	r	CI
Fish vs. EPA+ DHA	Belgium	0.34	(0.09,0.54)	0.19	(-0.07,0.42)	0.19	(-0.09,0.43)
	the Czech Republic	0.24	(-0.02,0.46)	0.21	(-0.04,0.44)	0.28	(0.02,0.51)
	France	0.37	(0.12,0.57)	0.54	(0.32,0.69)	0.55	(0.33,0.71)
	the Netherlands	0.30	(0.06,0.51)	0.34	(0.10,0.54)	0.35	(0.09,0.56)
	Norway	0.31	(0.07,0.52)	0.48	(0.25,0.65)	0.46	(0.22,0.64)
	Pooled‡	0.31	(0.20, 0.43)	0.36	(0.24,0.47)	0.31	(0.20,0.41)
Fruit & vegetables vs. sum of carotenoids§	Belgium	0.49	(0.27,0.66)	0.52	(0.30,0.68)	0.36	(0.10,0.57)
	the Czech Republic	0.33	(0.08,0.54)	0.39	(0.14,0.58)	0.35	(0.09,0.56)
	France	0.37	(0.13,0.57)	0.42	(0.18,0.61)	0.44	(0.19,0.64)
	the Netherlands	0.42	(0.18,0.60)	0.35	(0.11,0.55)	0.25	(-0.02,0.48)
	Norway	0.44	(0.21,0.62)	0.46	(0.23,0.64)	0.45	(0.21,0.64)
	Pooled	0.41	(0.30,0.53)	0.43	(0.31,0.54)	0.40	(0.29,0.49)
Vegetable intake vs. sum of carotenoids	Belgium	0.34	(0.09,0.54)	0.36	(0.11,0.56)	0.39	(0.14,0.59)
	the Czech Republic	0.19	(-0.07,0.42)	0.23	(-0.03,0.46)	0.21	(-0.06,0.45)
	France	0.44	(0.21,0.63)	0.47	(0.23,0.64)	0.58	(0.36,0.73)
	the Netherlands	0.46	(0.23,0.63)	0.34	(0.09,0.54)	0.16	(-0.10,0.41)
	Norway	0.44	(0.22,0.62)	0.53	(0.31,0.68)	0.50	(0.26,0.67)
	Pooled	0.38	(0.26,0.49)	0.39	(0.27,0.50)	0.38	(0.28,0.47)
Fruit intake vs. sum of carotenoids	Belgium	0.22	(-0.03,0.45)	0.40	(0.16,0.59)	0.18	(-0.10,0.42)
	the Czech Republic	0.31	(0.06,0.52)	0.32	(0.07,0.53)	0.27	(0.00,0.50)
	France	0.21	(-0.05,0.44)	0.29	(0.03,0.51)	0.26	(-0.01,0.50)
	the Netherlands	0.24	(-0.01,0.46)	0.32	(0.07,0.52)	0.23	(-0.03,0.47)
	Norway	0.12	(-0.13,0.36)	0.25	(-0.01,0.47)	0.21	(-0.06,0.45)
	Pooled	0.22	(0.11,0.34)	0.32	(0.20,0.43)	0.25	(0.14,0.36)

Table 4 continues…

		Crude		Adjusted₁*		Adjusted₂†	
		r	CI	r	CI	r	CI
Fruit & vegetables vs. α-carotene	Belgium	0.44	(0.21,0.62)	0.47	(0.24,0.64)	0.34	(0.08,0.56)
	the Czech Republic	0.14	(-0.12,0.38)	0.16	(-0.10,0.39)	0.14	(-0.14,0.39)
	France	0.38	(0.13,0.58)	0.42	(0.17,0.61)	0.41	(0.16,0.62)
	the Netherlands	0.52	(0.30,0.68)	0.44	(0.22,0.62)	0.38	(0.12,0.58)
	Norway	0.43	(0.20,0.61)	0.47	(0.25,0.65)	0.43	(0.19,0.62)
	Pooled	0.37	(0.27,0.50)	0.40	(0.28,0.51)	0.34	(0.23,0.44)
Fruit & vegetables vs. β-carotene	Belgium	0.41	(0.17,0.60)	0.44	(0.21,0.62)	0.30	(0.03,0.52)
	the Czech Republic	0.22	(-0.04,0.45)	0.26	(0.01,0.48)	0.24	(-0.03,0.48)
	France	0.33	(0.08,0.54)	0.38	(0.13,0.58)	0.39	(0.12,0.59)
	the Netherlands	0.21	(-0.04,0.44)	0.14	(-0.11,0.38)	-0.03	(-0.29,0.23)
	Norway	0.44	(0.21,0.62)	0.46	(0.24,0.64)	0.51	(0.27,0.68)
	Pooled	0.34	(0.23,0.46)	0.34	(0.23,0.46)	0.30	(0.19,0.40)
Vegetable intake vs lutein+zeaxanthin	Belgium	0.39	(0.15,0.59)	0.45	(0.21,0.63)	0.48	(0.24,0.66)
	the Czech Republic	0.15	(-0.11,0.39)	0.15	(-0.11,0.39)	0.17	(-0.10,0.42)
	France	0.26	(0.01,0.48)	0.29	(0.03,0.51)	0.43	(0.17,0.63)
	the Netherlands	0.45	(0.22,0.63)	0.36	(0.12,0.56)	0.21	(-0.06,0.45)
	Norway	0.32	(0.07,0.52)	0.41	(0.17,0.60)	0.33	(0.07,0.55)
	Pooled	0.32	(0.20,0.43)	0.34	(0.22,0.45)	0.29	(0.18,0.39)
Fruit intake vs β-cryptoxanthin	Belgium	0.20	(-0.06,0.43)	0.41	(0.17,0.60)	0.19	(-0.08,0.43)
	the Czech Republic	0.57	(0.36,0.72)	0.57	(0.36,0.72)	0.59	(0.38,0.74)
	France	0.28	(0.03,0.50)	0.29	(0.03,0.51)	0.22	(-0.06,0.46)
	the Netherlands	0.27	(0.02,0.49)	0.40	(0.17,0.59)	0.32	(0.06,0.54)
	Norway	0.05	(-0.21,0.29)	0.15	(-0.11,0.39)	0.09	(-0.18,0.35)
	Pooled	0.28	(0.17,0.40)	0.37	(0.26,0.49)	0.35	(0.26,0.46)

* Adjusted for within person variability by Multiple-Source-Method, taking into account the food propensity questionnaire.
† Adjusted₁ + adjusted for age, BMI, educational level, alcoholic beverage, smoking status using Partial Pearson correlations. Cholesterol levels was also included in the fruit and vegetable analysis.
‡ P-value for heterogeneity was > 0.10 for all analyses in the table.
§ α-carotene+β-cryptoxanthin+β-carotene+lutein+zeaxanthin.

A smaller range of correlations (0.33 in the Czech Republic to 0.49 in Belgium) was observed in women (**Table 4)**. Pearson correlations based on usual intakes, as estimated by the MSM method, barely differed from the crude ones (*Adjusted₁*).

Adjusted correlations including possible confounders varied in different directions and did not explain the differences across the centres (*Adjusted₂*). The adjusted correlation of fruit & vegetable intake in the Norwegian men was 0.07 while all other centres presented correlations ranging from 0.16 to 0.52. However, we did not identify deviating correlations when assessing the heterogeneity of the pooled correlations (p>0.10 for all comparisons). Overall, the correlations of the combined intake of fruit & vegetables with the sum of carotenoids were higher than of fruit and vegetables separately, especially for fruit. Correlations between fruit & vegetable intake with α-carotene were higher than with the sum of carotenoids in some subpopulations but lower in others like in the Dutch men, who happened to have the lowest consumption of carrots (results not shown). When using β-carotene as biomarker of fruit and vegetable intake, correlations were often lower than the sum of carotenoids, particularly in the Netherlands. The correlations between vegetable intake and lutein plus zeaxanthin and the correlations between fruit intake and β-cryptoxanthin did not explain the low correlation observed in Norway. Nevertheless, when using Spearman correlations, there was an improvement of those correlations in Norway. Furthermore, after including all types of juices in the fruit & vegetable group, the correlations modestly increased between fruit & vegetable intake and the sum of carotenoids in some centres, but not for all (data not shown). The major changes were seen in Norway for men, where for instance the correlation between fruit intake and β-cryptoxanthin increased from 0.03 to 0.32.

The percentage of fish oil supplement users was high in Norway (63%) as compared to the other four centres (<14%) (not shown in tables). In line with this, the percentage of EPA plus DHA in phospholipids of subjects, who reported not taking any fish-oil supplement, was substantially lower than of the supplement users and the total group in Norway (**Table 5**). Supplements containing carotenoids were less often consumed than those with fish oil, with the highest number of users in the Czech Republic and the Netherlands (11 subjects each) and the lowest in Norway (1 subject). As a result, mean serum carotenoid concentrations of non-supplement users were similar to those of the total group.

Table 5 – Biomarker levels (mean ± SE) of the total sample, users and non-users of specific supplements in the EFCOVAL validation study

	Belgium		the Czech Republic		France		the Netherlands		Norway	
	% EPA+DHA*	n	% EPA+DHA	n	% EPA+DHA	n	% EPA+DHA	n	% EPA+DHA	n
All subjects	5.3 ± 0.1	123	4.2 ± 0.1	118	5.5 ± 0.2	111	4.7 ± 0.1	120	7.2 ± 0.2	121
Supplement users†	6.5 ± 0.5	6	4.7 ± 0.3	17	5.8 ± 0.8	10	5.6 ± 0.5	13	7.9 ± 0.2	76
Non-supplement users	5.2 ± 0.1	117	4.1± 0.1	101	5.4 ± 0.1	101	4.6 ± 0.1	107	5.9 ± 0.3	45
	Serum carotenoids‡	n	Serum carotenoids	n	Serum carotenoids	n	Serum carotenoids	n	Serum carotenoids	n
All subjects	89.1 ± 4.0	123	70.6 ± 3.3	118	136.8 ± 6.3	111	98.2 ± 4.1	120	88.4 ± 3.5	121
Supplement users§	79.1 ± 19.2	2	94.4 ± 11.4	11	140.2 ± 62	5	127.5 ± 17.4	11	197	1
Non- supplement users	89.3 ± 4.1	121	68.1 ± 3.3	107	136.6 ± 6.1	106	95.2 ± 4.1	109	87.5 ±3.4	120

* of total fatty acids in phopholipids.

† Fish oil supplements.

‡ α-carotene+β-cryptoxanthin+β-carotene+lutein+zeaxanthin.

§ Supplements containing carotenoids.

Discussion

We compared the assessment of usual fish and fruit & vegetable consumption of adults estimated with two non-consecutive standardised 24-h recalls in combination with a FPQ between five centres in Europe. Overall, we observed weak to moderate associations between fish and fruit & vegetable intake and biomarkers. In men, correlations for fish intake in the Czech Republic and for fruit & vegetable intake in Norway were distinctly lower than those in the other centres. In women, the correlations across centres were rather comparable.

One of the major strengths of this study is the replicate collection of 24-h recalls, allowing the application of statistical adjustment to obtain an estimate of the individual usual intake. Welch et al.[29] have shown that one 24-h recall was less consistent in providing an association between fish intake and serum fatty acids than food frequency questionnaires or a 7-day diary. This can be explained by the fact that 24-h recalls are not able to reflect usual intakes of infrequently consumed foods. Indeed, we have observed that correlations substantially improved when considering usual intakes including FPQ data for fish. Another important strength of this study regards the unique setting of data collection that has provided standardised dietary intake and biomarker information for different countries.

One of the limitations of our study is that given the limited number of participants, large confidence intervals were observed. The sample size may also have limited the interpretation of the Cochrane Q-test. We found no statistically significant heterogeneity between correlations, but this could be caused by the relatively small sample size. Nevertheless, this is not very likely because the observed p-values for that test were rather high, especially in the assessment of fish intake ($p > 0.50$). In addition, data collection was performed in a different season in the Netherlands than in the other four centres. Considering that carotenoid contents in fruits and vegetables may differ between seasons[30], this could have led to a different performance of the method in the Dutch population. Nevertheless, we expect that both intake and biomarker assessment have been affected, thus minimizing the possible influence of seasonality on the correlations. Another potential limitation may be that the five centre populations are not representative of their respective country populations, because they can be expected to consist of health-conscious subjects. This hampers the extrapolation of our results to the general population. Furthermore, the individual usual intakes of foods estimated with MSM can be questioned. A study by Souverein et al.[31] showed that when applying methods such as MSM to groups of small sample size, the estimates of usual intake distributions are highly uncertain. Yet, the accuracy

of individual usual intakes estimated by MSM remains unclear, as it has not been investigated.

Apart from sampling variation, the differences in correlations between centres can be attributed to differences in the range of food intake, different compositions of the foods consumed, and the presence of other determinants or modifiers of the concentration biomarkers. The range of food intakes differed across centres, especially for fish intake. The low correlation for fish in Czech men may be explained by their very low amounts of fish consumption (on average 20 g/day) by only few subjects (26%). In the Czech population, the fish consumption is traditionally very low and our finding agrees with the low intake of 13 g/day according to a household budget survey[32]. In terms of differences in types of foods consumed, studies have shown that the correlation between fatty fish with serum and plasma fatty acids was stronger than for total fish[29; 33]. We were unable to present these correlations for our data given the high number of non-consumers for fish subgroups and the lack of specific FPQ data. Even so, we observed a very low ratio between fatty and lean fish intake in Czech men and in France. However, in France a considerable amount of shellfish and roe products was reported, which also contributes to n-3 fatty acids. The low consumption of shellfish and roe products together with the low ratio between fatty and lean fish intake might explain why the Czech Republic presented a very low correlation between fish intake and the biomarker, and France did not. Furthermore, substantial differences were observed in the types of fruits and vegetables consumed across centres. Because the contents and bioavailability of carotenoids in foods can differ depending on harvest conditions, degree of maturity, storage, and physical state[30], populations with different fruit and vegetable intakes, being more represented by one specific carotenoid than another, may have different carotenoid profiles as well. As a consequence, the sum of carotenoids may not sensibly represent the carotenoid content of fruits and vegetables of a specific population. We did observe different correlations across centres when using specific carotenoids as biomarkers of specific fruit and vegetable intake, but these did not substantially explain any observed differences in the comparison of fruit & vegetable intake vs. sum of carotenoids across centres. Furthermore, other dietary sources may contribute to n-3 fatty acids or carotenoids in the blood. For example, some oils are rich sources of α-linolenic acid (ALA), which may to a low extent be converted to EPA and DHA[34], and coloured foods, such as cheeses, can contain carotenoids[35; 36]. Non-fresh fruit and vegetable juices, which were not included in the fruit & vegetable group, and fortified foods may also contribute to concentrations of carotenoids. These sources may partly explain the low associations between intake and biomarker, as observed in Norway.

In addition, the interpretation of our results demands understanding of aspects influencing not only the assessment of intake but also that of the biomarker. Concentration biomarkers do not reflect absolute intake and their quantitative relationship with diet may vary between populations, depending on the presence and relative impact of determinants, such as genetic variation and lifestyle factors[17; 37]. For instance, smoking and alcohol consumption have been inversely associated with levels of n-3 fatty acids[38-40] and serum carotenoids[41; 42]. Nevertheless, when adjusting our correlations for a number of potential confounders like alcohol intake, smoking status and total serum cholesterol concentrations, the outcomes were quite similar.

In addition, the association between food intake and biomarkers may have been influenced by the use of supplements[29; 33]. Although the Norwegian sample was not the centre with the most deviating association between fish intake and its biomarker, the percentage EPA plus DHA in phospholipids in supplement users was markedly higher than in non-supplement users and the total population in Norway. Nevertheless, it is questionable whether participants were able to recall the exact type and brand name of their supplements[43]. Therefore, a degree of uncertainty remains in the evaluation of supplement use in relation to food intake assessment and in the explanations of differences in correlations across centres.

The correlations in this paper are consistent with the results of other studies that have found weak to moderate correlations between fish intake and n-3 fatty acids in the blood[29; 44] and between fruit & vegetable intake and serum carotenoids[45; 46] when using 24-h dietary recalls (r for fish intake between 0.11 and 0.22 and for fruit & vegetable between 0.30 and 0.42).

Despite the limitations, we conclude that two standardised 24-h recalls using EPIC-Soft in combination with an FPQ appeared to be appropriate to rank subjects according to their usual fish and fruit and vegetable intake within the five European centres in a comparable manner.

Acknowledgements

References

1. EFSA (2010) Statistical modelling of usual intake. Scientific report. EFSA-Q-2009-00841. 1-23.
2. Ireland J, van Erp-Baart AM, Charrondiere UR *et al.* (2002) Selection of a food classification system and a food composition database for future food consumption surveys. *Eur J Clin Nutr* 56 Suppl 2, S33-45.
3. Brussaard JH, Johansson L & Kearney J (2002) Rationale and methods of the EFCOSUM project. *Eur J Clin Nutr* 56 Suppl 2, S4-7.
4. Brussaard JH, Lowik MR, Steingrimsdottir L *et al.* (2002) A European food consumption survey method-- conclusions and recommendations. *Eur J Clin Nutr* 56 Suppl 2, S89-94.
5. Biro G, Hulshof KF, Ovesen L *et al.* (2002) Selection of methodology to assess food intake. *Eur J Clin Nutr* 56 Suppl 2, S25-32.
6. Slimani N & Valsta L (2002) Perspectives of using the EPIC-SOFT programme in the context of pan-European nutritional monitoring surveys: methodological and practical implications. *Eur J Clin Nutr* 56 Suppl 2, S63-74.
7. Steingrimsdottir L, Ovesen L, Moreiras O *et al.* (2002) Selection of relevant dietary indicators for health. *Eur J Clin Nutr* 56 Suppl 2, S8-11.
8. Crispim SP, de Vries JH, Geelen A *et al.* (2011) Two non-consecutive 24 h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres - results from the European Food Consumption Validation (EFCOVAL) study. *Br J Nutr* 105, 447-458.
9. Subar AF, Dodd KW, Guenther PM *et al.* (2006) The food propensity questionnaire: concept, development, and validation for use as a covariate in a model to estimate usual food intake. *J Am Diet Assoc* 106, 1556-1563.
10. Macdonald HM, Hardcastle AC, Duthie GG *et al.* (2009) Changes in vitamin biomarkers during a 2-year intervention trial involving increased fruit and vegetable consumption by free-living volunteers. *Br J Nutr* 102, 1477-1486.
11. Polsinelli ML, Rock CL, Henderson SA *et al.* (1998) Plasma carotenoids as biomarkers of fruit and vegetable servings in women. *J Am Diet Assoc* 98, 194-196.
12. Yeum KJ, Booth SL, Sadowski JA *et al.* (1996) Human plasma carotenoid response to the ingestion of controlled diets high in fruits and vegetables. *Am J Clin Nutr* 64, 594-602.
13. Katan MB, Deslypere JP, van Birgelen AP *et al.* (1997) Kinetics of the incorporation of dietary fatty acids into serum cholesteryl esters, erythrocyte membranes, and adipose tissue: an 18-month controlled study. *J Lipid Res* 38, 2012-2022.
14. Leaf DA, Connor WE, Barstad L *et al.* (1995) Incorporation of dietary n-3 fatty acids into the fatty acids of human adipose tissue and plasma lipid classes. *Am J Clin Nutr* 62, 68-73.
15. Zock PL, Mensink RP, Harryvan J *et al.* (1997) Fatty acids in serum cholesteryl esters as quantitative biomarkers of dietary intake in humans. *Am J Epidemiol* 145, 1114-1122.
16. Freedman LS, Kipnis V, Schatzkin A *et al.* (2010) Can we use biomarkers in combination with self-reports to strengthen the analysis of nutritional epidemiologic studies? *Epidemiol Perspect Innov* 7, 2.
17. Jenab M, Slimani N, Bictash M *et al.* (2009) Biomarkers in nutritional epidemiology: applications, needs and new horizons. *Hum Genet* 125, 507-525.
18. Slimani N, Deharveng G, Charrondiere RU *et al.* (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Methods Programs Biomed* 58, 251-266.
19. Slimani N, Ferrari P, Ocke M *et al.* (2000) 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* 54, 900-917.
20. Kaluzny MA, Duncan LA, Merritt MV *et al.* (1985) Rapid separation of lipid classes in high yield and purity using bonded phase columns. *J Lipid Res* 26, 135-140.
21. Metcalfe LD, Schmitz AA & Pelka JR (1966) Rapid Preparation of Fatty Acid Esters from Lipids for Gas Chromatographic Analysis. *Analytical Chemistry* 38, 514-515.
22. Khan NC, West CE, de Pee S *et al.* (2007) The contribution of plant foods to the vitamin A supply of lactating women in Vietnam: a randomized controlled trial. *Am J Clin Nutr* 85, 1112-1120.
23. Jansen MC, Van Kappel AL, Ocke MC *et al.* (2004) Plasma carotenoid levels in Dutch men and women, and the relation with vegetable and fruit consumption. *Eur J Clin Nutr* 58, 1386-1395.
24. Al-Delaimy WK, Slimani N, Ferrari P *et al.* (2005) Plasma carotenoids as biomarkers of intake of fruits and vegetables: ecological-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Eur J Clin Nutr* 59, 1397-1408.
25. German Institute of Human Nutrition (2009) The Multiple Source Method (MSM). https://nugo.dife.de/msm/
26. Brady WE, Mares-Perlman JA, Bowen P *et al.* (1996) Human serum carotenoid concentrations are related to physiologic and lifestyle factors. *J Nutr* 126, 129-137.
27. Kleinbaum DG, Kupper LL, Muller KE *et al.* (2007) *Applied regression analysis and other multivariable methods.* Belmont, California: Duxbury Press.

28. Field AP (2005) Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychol Methods* 10, 444-467.

29. Welch AA, Bingham SA, Ive J *et al.* (2006) Dietary fish intake and plasma phospholipid n-3 polyunsaturated fatty acid concentrations in men and women in the European Prospective Investigation into Cancer-Norfolk United Kingdom cohort. *Am J Clin Nutr* 84, 1330-1339.

30. Maiani G, Caston MJ, Catasta G *et al.* (2009) Carotenoids: actual knowledge on food sources, intakes, stability and bioavailability and their protective role in humans. *Mol Nutr Food Res* 53 Suppl 2, S194-218.

31. Souverein OW, Dekkers A, Geelen A *et al.* (2011) Comparing four different methods to estimate usual intake distributions. *Eur J Clin Nutr (in press).*

32. Dofkova M, Kopriva V, Resova D *et al.* (2001) The development of food consumption in the Czech Republic after 1989. *Public Health Nutr* 4, 999-1003.

33. Hjartaker A, Lund E & Bjerve KS (1997) Serum phospholipid fatty acid composition and habitual intake of marine foods registered by a semi-quantitative food frequency questionnaire. *Eur J Clin Nutr* 51, 736-742.

34. Connor WE (2001) n-3 Fatty acids from fish and fish oil: panacea or nostrum? *Am J Clin Nutr* 74, 415-416.

35. Gordon HT & Bauernfeind JC (1982) Carotenoids as food colorants. *Crit Rev Food Sci Nutr* 18, 59-97.

36. Jones ST, Aryana KJ & Losso JN (2005) Storage stability of lutein during ripening of cheddar cheese. *J Dairy Sci* 88, 1661-1670.

37. Kaaks R, Riboli E & Sinha R (1997) Biochemical markers of dietary intake. *IARC Sci Publ*, 103-126.

38. Galan P, Viteri FE, Bertrais S *et al.* (2005) Serum concentrations of beta-carotene, vitamins C and E, zinc and selenium are influenced by sex, age, diet, smoking status, alcohol consumption and corpulence in a general French adult population. *Eur J Clin Nutr* 59, 1181-1190.

39. Pawlosky RJ, Hibbeln JR & Salem N, Jr. (2007) Compartmental analyses of plasma n-3 essential fatty acids among male and female smokers and nonsmokers. *J Lipid Res* 48, 935-943.

40. Simon JA, Fong J, Bernert JT, Jr. *et al.* (1996) Relation of smoking and alcohol consumption to serum fatty acids. *Am J Epidemiol* 144, 325-334.

41. Northrop-Clewes CA & Thurnham DI (2007) Monitoring micronutrients in cigarette smokers. *Clin Chim Acta* 377, 14-38.

42. Rock CL, Thornquist MD, Kristal AR *et al.* (1999) Demographic, dietary and lifestyle factors differentially explain variability in serum carotenoids and fat-soluble vitamins: baseline results from the sentinel site of the Olestra Post-Marketing Surveillance Study. *J Nutr* 129, 855-864.

43. Skeie G, Braaten T, Hjartaker A *et al.* (2009) Use of dietary supplements in the European Prospective Investigation into Cancer and Nutrition calibration study. *Eur J Clin Nutr* 63 Suppl 4, S226-238.

44. Saadatian-Elahi M, Slimani N, Chajes V *et al.* (2009) Plasma phospholipid fatty acid profiles and their association with food intakes: results from a cross-sectional study within the European Prospective Investigation into Cancer and Nutrition. *Am J Clin Nutr* 89, 331-346.

45. Al-Delaimy WK, Ferrari P, Slimani N *et al.* (2005) Plasma carotenoids as biomarkers of intake of fruits and vegetables: individual-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Eur J Clin Nutr* 59, 1387-1396.

46. Resnicow K, Odom E, Wang T *et al.* (2000) Validation of three food frequency questionnaires and 24-hour recalls with serum carotenoid levels in a sample of African-American adults. *Am J Epidemiol* 152, 1072-1080.

Chapter 5

Design aspects of 24-h recall assessments may affect the estimates of protein and potassium intake in dietary surveys

Sandra P. Crispim, Anouk Geelen, Els Siebelink,
Inge Huybrechts, Inger T.L. Lillegaard, Irene Margaritis,
Irena Rehurkova, Nadia Slimani, Marga C. Ocke,
Evelien de Boer, Pieter van 't Veer and Jeanne H. M. de Vries

Abstract

Objective: To evaluate the impact of different modes of administration (face-to-face vs. telephone), recall days (1st vs. 2nd), days of the week (week- vs. weekend days), and interview days (1 vs. 2 days later) on bias in protein and potassium intake collected with 24-h recalls.

Design: European Food Consumption Validation (EFCOVAL) study.

Setting: Five centres in Belgium, the Czech Republic, France, the Netherlands and Norway.

Subjects: 600 adults (45-65 y).

Methods: Two non-consecutive 24-h recalls (EPIC-Soft) were used to estimate protein and potassium intake by a face-to-face interview at the research centres and a telephone interview, and included all days of the week. Two 24-h urines were collected to determine biomarkers of protein and potassium intake.

Results: The bias in protein intake in the Czech Republic and Norway was smaller for telephone than face-to-face interviews (p=0.01). The second 24-h recall estimates of protein intake in France and potassium intake in Belgium were less accurate than the first 24-h recall estimates (p=0.01 and 0.04, respectively). In the Czech Republic, protein intake estimated during weekends and potassium intake estimated during weekdays were less accurate than during other days of the week (p=0.01). In addition, potassium intake collected two days later in the Czech Republic was likely to be overestimated.

Conclusions: 24-h recalls collected by telephone provide a more accurate assessment than by face-to-face interviews and second 24-h recalls seem to be less accurate than first recalls in some centres. In addition, it is suggested that the days of the week should be equally represented in dietary surveys.

Introduction

Standardisation of methods and field work is of crucial importance to compare dietary intake between European countries[1]. The European Food Consumption Validation (EFCOVAL) study (www.efcoval.eu) aimed to further develop and validate a European food consumption method using a standardised 24-h recall (EPIC-Soft software) for assessing dietary intake within and between European countries. The study was carried out in view of a future pan-European dietary monitoring system, which is foreseen to deliver detailed, harmonised and high quality food consumption

data for between-country comparisons[2]. In EFCOVAL, design aspects of 24-h recall assesments, such as mode of administration and day of the week, were shown to influence the variation in protein and potassium bias across European centres[3]. Thus, further investigating the different design aspects of fieldwork used for collecting 24-h recalls within different countries is relevant for future surveys.

In some countries, telephone interviews may be applied as an alternative to face-to-face interviews. A number of studies have shown that 24-h recalls administered by telephone and face-to-face yield similar data[4; 5]. However, to know whether they really provide similar results, the validity of interviews administered by telephone should be compared to that of face-to-face, especially in countries with limited experience in telephone interviews[6].

The collection of at least two non-consecutive days of intake to estimate habitual intake through statistical modeling has been advised by EFCOSUM[7]. A second dietary interview may be affected by a motivational or learning effect, however. Some studies have suggested that the subjects' motivation decreases with increasing number of days of data collection, leading to underreporting of intake[8; 9]. Besides, the results of the second recall may differ because subjects learned from their first recall. Therefore, it is important to investigate whether 1st and 2nd recall estimates provide comparable results.

Another important issue in future pan-European surveys concerns the dietary data collection on different days of the week. Food consumption on weekend days differs from weekdays in most European countries[6; 10]. It is therefore advisable that dietary assessments in surveys are randomly allocated over all days of the week among the population[6]. However, it is questionable whether the accuracy of the assessments of 24-h recalls is similar between week- and weekend days.

Furthermore, to carry out dietary interviews on a Sunday for recalling the diet of Saturday, is less feasible in some countries like the Netherlands and Spain. Reasons for this include problems with transportation of interviewers to remote areas when using face-to-face interviews and aspects of family privacy on Sundays[10]. An alternative is to collect data from Saturday on the following Monday (two days later), but whether those assessments provide comparable results to those on Sundays is to be investigated.

In this paper, we evaluated the bias in protein and potassium intake collected with 24-h recalls between different modes of administration (telephone vs. face-to-face),

recall days (1st vs. 2nd), days of the week (weekdays vs. weekend), and interview days (1 vs. 2 days later) in five European centres.

Subjects and Methods

Data was collected in the framework of the EFCOVAL validation study in five European centres: Belgium, the Czech Republic, France, the Netherlands and Norway. Ethical committees in each centre approved the research protocol and participants signed an informed consent. In brief, 600 subjects were interviewed twice to report their intake using the computerised 24-h recall method (EPIC-Soft software)[11; 12]. One recall was performed by telephone with participants at home and the other one face-to-face mostly in the study centre. The order of the two modes of administration of the 24-h recall was randomly assigned with at least four weeks between the recalls. Furthermore, dietary recalls followed a randomised schedule that equally included all days of the week. However, in Belgium, the Czech Republic and the Netherlands dietary recalls about Saturdays were not conducted on Sundays but on Mondays. The number of trained interviewers (i.e., dietitians or nutritionists) was four in Belgium, six in the Czech Republic, two in France, seven in the Netherlands and three in Norway. On the same days of which 24-h recall data were reported, the 24-h urines were collected to determine nitrogen and potassium excretion in urine. These were used as biomarkers of protein and potassium intake, respectively. *para*-Aminobenzoic acid (PABA) was used to check the completeness of urine collections. Complete logistics and details of the study were reported elsewhere[13; 14].

The data used in this paper include mostly repeated measurements of the same subject. The bias in protein and potassium intake was defined as the mean of individual ratios between nutrient intake from 24-h recalls and the excretion of its recovery biomarkers. The means were adjusted for interviewer using an ANCOVA model and then reported by centre and mode of administration (face-to-face vs. telephone interview), recall day (1st vs. 2nd), day of the week (week vs. weekend day), or interview day (1 vs. 2 days later – i.e., Saturday's intake collected on Mondays). Adjustment for subject characteristics and other design aspects were not necessary because of the well balanced dataset. Weekdays were defined to include Mondays, Tuesdays, Wednesdays and Thursdays, and weekends Fridays, Saturdays and Sundays. We also performed the analysis including Friday as a weekday and as the results were quite similar to the first definition, they are not presented. Hereafter, 'recall day' refers to 1st or 2nd day of application of the 24-h recall, 'day of the week' to the comparison

of week- and weekend days, and 'interview day' to after 1 vs. 2 days later of the dietary intake. ANCOVA was also used to test the differences between adjusted mean ratios of the subgroups in each centre. Analyses were performed using SAS statistical package, version 9.1 (SAS Institute Inc, Cary, NC).

Results

The bias in protein and potassium intakes, as represented by the ratios between intake and excretion, were comparable for face-to-face and telephone interviews in Belgium, France and the Netherlands (**Table 1**). In the Czech Republic and Norway, the bias in potassium intake was also comparable between these modes of administration, but not for protein. In these two centers, the bias for the assessment of protein was smaller by telephone than by face-to-face interviews (p=0.01 in both countries). However, while an overestimation of the mean protein intake collected with face-to-face interviews was observed in Norway, an underestimation was seen in the Czech Republic.

The protein and potassium intake collected on 1st and 2nd recall days yielded similar bias in the Czech Republic, Norway and the Netherlands (**Table 2**). However, protein intake in France and potassium intake in Belgium collected during the second 24-h recall were apparently less accurate than intakes from the first recall (p=0.01 and 0.04, respectively).

The bias in protein and potassium intakes collected on weekdays did not differ from weekend days, except in the Czech Republic (**Table 3**). Whilst protein intake was underestimated during weekdays in the Czech Republic, potassium intake was overestimated during weekends (p=0.01 for both).

The bias in protein and potassium intake from recalls collected on Mondays about Saturday's intake was similar to those of recalls about the other days of the week in the Netherlands and Belgium (not shown in tables). However, in the Czech Republic the bias in potassium intake from recalls performed two days later indicated overestimation of intake (ratio of 1.35 ± 0.09 for Saturdays' intake (*n* 28) vs. 1.14 ± 0.08 for the average of Fridays and Sundays (*n* 74) and 1.06 ± 0.08 for the average of Mondays to Thursdays (*n* 132)). Furthermore, removing Saturdays' potassium intake in the comparison of week- and weekend days reduced the difference observed in the Czech Republic (p=0.05, data not shown).

Table 1 – Comparison of mean* ratios of nutrient intake and excretion by mode of administration in the EFCOVAL validation study

	Ratios intake and excretion												
	Protein							Potassium					
	Face-to-face			Telephone				Face-to-face			Telephone		
Country	*n*	Mean	SE	*n*	Mean	SE	p-value	*n*	Mean	SE	*n*	Mean	p-value
Belgium	123	0.97	0.03	120	0.91	0.03	0.15	123	0.97	0.03	120	0.92 0.03	0.28
the Czech Republic	117	0.91	0.04	117	1.02	0.04	0.01	117	1.09	0.05	117	1.13 0.05	0.48
France	108	0.89	0.03	109	0.90	0.03	0.78	108	0.86	0.03	109	0.90 0.03	0.29
the Netherlands	118	0.92	0.04	120	0.93	0.04	0.80	118	1.00	0.03	120	0.98 0.04	0.60
Norway	123	1.07	0.03	122	0.97	0.03	0.01	123	1.00	0.03	122	1.01 0.03	0.99

Table 2– Comparison of mean* ratios of nutrient intake and excretion by recall day in the EFCOVAL validation study

	Ratios intake and excretion													
	Protein							Potassium						
	1st recall			2nd recall				1st recall			2nd recall			
Country	*n*	Mean	SE	*n*	Mean	SE	p-value	*n*	Mean	SE	*n*	Mean	SE	p-value
Belgium	122	0.97	0.04	121	0.93	0.03	0.34	122	1.00	0.04	121	0.91	0.03	0.04
the Czech Republic	118	0.98	0.04	116	0.94	0.04	0.38	118	1.11	0.05	116	1.10	0.05	0.75
France	110	0.94	0.02	107	0.85	0.03	0.01	110	0.90	0.03	107	0.87	0.03	0.48
the Netherlands	119	0.92	0.04	119	0.93	0.04	0.85	119	0.96	0.04	119	1.01	0.04	0.27
Norway	124	1.04	0.03	121	1.00	0.04	0.38	124	1.01	0.03	121	1.00	0.04	0.81

*Adjusted for interviewer

Table 3 – Comparison of mean[*] ratios of nutrient intake and excretion of recalls performed on week- or weekend days in the EFCOVAL validation study

	Ratios intake and excretion													
	Protein							Potassium						
	Weekday[†]			Weekend[‡]				Weekday			Weekend			
Country	*n*	Mean	SE	*n*	Mean	SE	p-value	*n*	Mean	SE	*n*	Mean	SE	p-value
Belgium	141	0.93	0.03	102	0.97	0.04	0.36	141	0.92	0.03	102	0.98	0.04	0.12
the Czech Republic	132	0.92	0.03	102	1.03	0.04	0.01	132	1.05	0.04	102	1.20	0.05	0.01
France	141	0.89	0.02	76	0.89	0.03	0.97	141	0.88	0.02	76	0.89	0.03	0.90
the Netherlands	143	0.92	0.03	95	0.95	0.04	0.50	143	1.00	0.03	95	0.97	0.04	0.50
Norway	141	1.00	0.03	104	1.05	0.04	0.28	141	0.98	0.03	104	1.04	0.04	0.13

[*] Adjusted for interviewer

[†] Monday-Thursday

[‡] Friday-Sunday

A significant interviewer effect was observed in some of the analyses (p<0.05), but it did not change the conclusions as compared to the crude analyses. An exception was seen for Belgium, where the bias in protein intake was only similar between the two modes of administration after adjustment for interviewer.

Discussion

In the present study, we compared the bias in protein and potassium intake estimated from a standardised 24-h recall between different modes of administration, recall days, days of the week, and interview days in five European centres. Overall, the biases in protein and potassium intake were comparable between face-to-face and telephone interviews, first and second recall days, week and weekend days, and interviews performed one or two days later in some, but not all, centres.

Other studies have indicated that dietary data collected by telephone are in good agreement with those by face-to-face interviews, especially when adjusted for interviewer[4; 5; 15]. However, these studies compared the intakes estimated by the two modes of administration rather than their validity. Contrarily, our validation results showed differences between the two modes of administration in the Czech Republic and Norway with significantly larger biases in protein intakes when face-to-face interviews were conducted. The fact that subjects were allowed to check foods consumed at home can hypothetically explain the better validity of recalls by telephone, as this was not possible during the face-to-face interviews performed at the study centre. Nevertheless, this study showed that bias in potassium intake was comparable between the two modes of administration in all centres.

In the OPEN study, first and second 24-h recall assessments of protein intake showed similar bias[16]. We, however, observed a less accurate performance of the method for second day assessments of protein or potassium intakes in France and Belgium, respectively. This difference is hypothetically explained by less motivation of the subjects for the second recall. However, also a learning effect may have affected the 2nd recalls. Thus, the absence of a difference in bias observed in some centres may be explained by the fact that the two proposed effects could have ruled each other out.

The Czech Republic was the only centre that did not present comparable biases in the assessments of protein and potassium intake week- and weekend days and between 24-h recalls collected one and two days after the intake. Reasons for these differences are not clear though.

Three possible explanations for the observed differences in bias between modes of administrations, recall days, days of the week and interview days within some of the centres are given. First, food composition data are known to be a source of errors in dietary assessments[13] and may have invariably influenced the bias between the different design aspects of the 24-h recall assessment. For example, different factors were used to convert nitrogen into protein contents in foods in the food composition tables applied in the centres. Furthermore, an official national food composition table in the Czech Republic was not available during the study and the nutrient composition of foods consumed needed to be borrowed from Slovak and other foreign tables. A second explanation may be that specific foods or food groups, of which the intake varied between centres because of a different dietary pattern, may have been differentially misreported. Third, the degree of experience in using EPIC-Soft may have caused differences in bias among the centres. Thus, it could be hypothesized that the centre's degree of experience possibly in combination with the quality of the figures in food composition tables and their dietary pattern caused the differences in bias of the different design aspects within the centres.

A limitation of our study is that we probably included a health conscious population, which may hinder the extrapolation of the results to the general population of the respective countries. Additionally, only two nutrients were evaluated. Nevertheless, as differences were observed in the performance of the method between different design aspects of the assessment, this may also be true for other nutrients and foods. Moreover, because of small sample sizes in the analysis of the interview day in this study, we may lack power to conclude on the comparability of data collected one or two days after the dietary intake.

To our knowledge, this is the first study describing the bias in protein and potassium intake between different modes of administration, recall days, days of the week, and interview days across different European populations. The results presented here can provide a greater understanding of the performance of the 24-h recall methodology, which may have implications for the planning of future dietary surveys and the analyses and interpretation of the collected data.

We conclude that 24-h recalls collected by telephone interviews seem to provide a more accurate assessment than by face-to-face interviews at a research centre in some European centres. In addition, second recall assessments may be less accurate than first recalls. Finally, it is suggested that the days of the week should be equally represented in dietary surveys or appropriately adjusted for during data analysis.

References

1. Brussaard JH, Lowik MR, Steingrimsdottir L *et al.* (2002) A European food consumption survey method-conclusions and recommendations. *Eur J Clin Nutr* 56 Suppl 2, S89-94.
2. EFSA (2010) Statistical modelling of usual intake. Scientific report. EFSA-Q-2009-00841. 1-23.
3. Crispim SP, Geelen A, de Vries J *et al.* (2011) Bias in protein and potassium intake collected with 24-h recalls (EPIC-Soft) is small and rather comparable across European populations. *submitted*.
4. Fox TA, Heimendinger J & Block G (1992) Telephone surveys as a method for obtaining dietary information: a review. *J Am Diet Assoc* 92, 729-732.
5. Casey PH, Goolsby SL, Lensing SY *et al.* (1999) The use of telephone interview methodology to obtain 24-hour dietary recalls. *J Am Diet Assoc* 99, 1406-1411.
6. De Henauw S, Brants HA, Becker W *et al.* (2002) Operationalization of food consumption surveys in Europe: recommendations from the European Food Consumption Survey Methods (EFCOSUM) Project. *Eur J Clin Nutr* 56 Suppl 2, S75-88.
7. Hoffmann K, Boeing H, Dufour A *et al.* (2002) Estimating the distribution of usual dietary intake by short-term measurements. *Eur J Clin Nutr* 56 Suppl 2, S53-62.
8. 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, 48-55.
9. Buzzard IM, Faucett CL, Jeffery RW *et al.* (1996) Monitoring dietary change in a low-fat diet intervention study: advantages of using 24-hour dietary recalls vs food records. *J Am Diet Assoc* 96, 574-579.
10. EFCOSUM group (2001) *European Food Consumption Survey Method. Final report.* Zeist: TNO Nutrition and Food Research.
11. Slimani N, Deharveng G, Charrondiere RU *et al.* (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Methods Programs Biomed* 58, 251-266.
12. Slimani N, Ferrari P, Ocke M *et al.* (2000) 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* 54, 900-917.
13. Crispim SP, de Vries JH, Geelen A *et al.* (2011) Two non-consecutive 24 h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres - results from the European Food Consumption Validation (EFCOVAL) study. *Br J Nutr* 105, 447-458.
14. Crispim SP, Geelen A, de Vries JH *et al.* (2011) Biomarker-based evaluation of two 24-h recalls (EPIC-Soft) for estimating and comparing fish, fruit and vegetable intakes across five European centers: results from the European Food Consumption Validation (EFCOVAL) study. *Eur J Clin Nutr (in press)*.
15. Brustad M, Skeie G, Braaten T *et al.* (2003) Comparison of telephone vs face-to-face interviews in the assessment of dietary intake by the 24 h recall EPIC SOFT program-the Norwegian calibration study. *Eur J Clin Nutr* 57, 107-113.
16. Subar AF, Kipnis V, Troiano RP *et al.* (2003) Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol* 158, 1-13.

Chapter 6

Dietary exposure to flavouring substances: from screening methods to detailed assessments using food consumption data collected with EPIC-Soft software

Sandra P. Crispim, Anouk Geelen, Cinzia Le Donne, Jeanne H. M. de Vries, Stefania Sette, Antonio Raffo, Els Siebelink, Marga Ocke, Pieter van 't Veer and Catherine Leclercq

Abstract

This study aimed to compare different methods of assessing dietary exposure to flavourings in the context of a stepwise approach. The dietary exposure to four flavourings - raspberry ketone, glycyrrhizinic acid, coumarin, and caffeine - was determined. When dietary exposure exceeded the safety limits, the need for more detailed assessment using less aggregated data was judged necessary. First, screening methods - maximized survey-derived daily intake (MSDI), single-portion exposure technique (SPET), and modified theoretical added maximum daily intake (mTAMDI) - were applied. Next, individual food consumption data were used for creating models with different levels of detail to identify the foods: a model based on food groups and models based on food items. These were collected from 121 Dutch adults using a standardised two 24-h dietary recall (EPIC-Soft) in the European Food Consumption Validation (EFCOVAL) study. Three food item models were developed: without improvements of the flavouring descriptor built in the software; with improvements; and with use of non-specified flavour descriptors. Based on the results of at least one of the three screening methods, refined assessment was necessary for raspberry ketone, glycyrrhizinic acid, and caffeine. When applying the food group model, the need for refinement was indicated for the four flavourings. When applying the food item models, only glycyrrhizinic acid and caffeine presented dietary exposure above the safety limits. In the raspberry ketone case, dietary exposure increased when improvements in food description were considered. The use of non-specified flavour descriptors hardly changed the results. The collection of detailed food consumption data at the individual level is useful in the dietary exposure assessment of these flavourings.

Introduction

More than 2700 flavouring substances (hereafter 'flavourings') are currently registered and can be added to foods and beverages in the European Union[1-4]. Accordingly, the Joint FAO/WHO* Expert Committee on Food Additives (JECFA) and the European Food Safety Authority (EFSA) have been working towards the safety evaluation of flavourings in order to provide a positive list of these substances[5; 6]. Within the safety evaluation procedure of any chemical substance, one crucial step is the dietary exposure assessment.

---

*Joint Food and Agriculture Organization/World Health Organization

A major pitfall of dietary exposure assessment to chemicals is the limited availability of the two types of information that are needed: food consumption data and chemical concentration in foods[7]. The ideal situation of performing a detailed dietary exposure assessment by collecting information at the individual level for every hazardous substance is neither practical nor cost-effective[8], especially when the objective is to verify that a safety limit is not exceeded. Consequently, dietary exposure should be evaluated through a stepwise approach[9].

The stepwise approach follows the premise of an assessment using the least refined method (screening) towards the most refined one, if necessary. The refinement of data is judged necessary when the dietary exposure assessed with a conservative method using highly aggregated data (i.e., the chemical is assumed to be present in specific food groups supposedly ingested by the whole population and there is no information about distribution of the consumption) exceeds the safety limits of the chemical. Once safety limits are surpassed, this indicates there is a possibility of safety concern and further investigation is needed by using less aggregated data (e.g., food consumption collected at the individual level). Then, the next step is performed using more detailed information on food consumption and/or concentration data in order to determine the right hand extreme of the distribution of dietary exposure. On the other hand, when the dietary exposure assessed using screening methods is under the safety limits, further refinement of the assessment is not needed[8-10]. In this way, wasting of resources by collecting a large amount of unneeded data is avoided. The most important characteristic of screening methods is that conservative assumptions regarding food consumption and concentration levels in food should be used in order to provide a good level of protection for the whole population by intentionally overestimating chronic dietary exposure[11].

The assessment is said to be refined when dietary exposure evaluations go beyond conservative assumptions of screening methods. In a refined assessment, the purpose of the evaluation often changes to provide an estimate of dietary exposure based on observed food consumption patterns and/or measured chemical concentration data rather than assumed values[11; 12]. The refinement of dietary exposure to chemicals should be designed in such a way that non-average individuals are considered in the assessment, and in particular those who consume relatively large quantities of foods containing higher concentrations of substances that may potentially lead to a health risk[11].

To consider the distribution of dietary exposure, it is important to collect food consumption information from individuals rather than base the assessment on average

population data. Methods available to collect individual dietary data include food records, food frequency questionnaires and 24-h dietary recalls[13]. Monitoring surveys aim to provide such type of information for nationally representative populations. However, dietary assessment methods are not standardised across countries[14] and the level of detail available in the data may differ considerably.

Furthermore, challenges may be encountered during refined dietary exposure assessment using information at the individual level. One of these challenges is the presence of uncertainties in the process of identifying and describing the consumption of foods. The non-identification of potential consumers of interest may occur due to the lack of ability of dietary methods, such as 24-h dietary recalls, on capturing sufficient information for the assessment of chemicals in the diet[12]. Additionally, the ability of interviewees on providing such information can be limited, resulting in misreporting or non-reporting of foods.

The 'European Food Consumption Validation' (EFCOVAL) project aims at validating a method for future monitoring surveys on the dietary intake in European countries. For this purpose, a duplicate 24-h recall using EPIC-Soft software has been chosen. A secondary objective is to adapt EPIC-Soft in such a way that food safety issues can be investigated. To explore this, the flavouring substances category has been chosen.

In this paper, we report the results of an explorative study aimed at comparing methods used to estimate the dietary exposure to flavourings in the context of a stepwise approach.

Material and Methods

Flavourings under investigation

Four flavourings were selected for the exercise of assessing dietary exposure to flavourings in the diet: raspberry ketone, glycyrrhizinic acid (excluding ammonium glycyrrhizinate), coumarin and caffeine. These flavourings represent different origins (naturally contained in food and/or added flavouring) and different production volumes when used as added flavouring.

Raspberry Ketone (4-(4-Hydroxyphenyl)butan-2-one; Chemical Abstracts Service (CAS) number 5471-51-2) is the primary aroma compound of raspberry and is also

found naturally in other berry fruits such as cranberry, blackberry, and loganberry[15; 16]. It is also used in flavour formulations of mixed berries and strawberries added to processed foods such as yoghurt and beverages[17; 18]. The safety limit for raspberry ketone is assumed to be 0.03 mg kg^{-1} body weight (bw) day^{-1}, considering the threshold of toxicological concern (TTC) of 1800 µg person^{-1} day^{-1} for flavourings classified in structural class I[19] and assuming a 60 kg adult. Structural class I suggests the lowest of three classes of toxicity of flavourings in their safety evaluation procedure by JECFA and was assigned to raspberry ketone in 2001[20].

Glycyrrhizinic acid (CAS number 1405-86-3) is found in foods and beverages as a natural constituent or as an added flavouring. Glycyrrhizinic acid is present in extracts of roots and rhizomes of the liquorice plant, *Glycyrrhiza glabra*. Liquorice confectionery and herbal teas are the main sources of dietary exposure to this substance[21; 22]. Although an acceptable daily intake (ADI) is not determined, safety evaluations of glycyrrhizinic acid performed by JECFA and the Scientific Committee on Food (SCF) have suggested that a dietary exposure to 100 mg day^{-1} would be unlikely to cause adverse effects in the majority of adults[23-25]. A safety factor of 10 has been used by Stormer et al.[21] to establish a safety limit with the 100 mg day^{-1} figure. This safety factor is used to account for inter-individual variability in susceptibility when toxicological information is available for humans. Based on this reference, a safety limit of 0.16 mg kg^{-1} bw day^{-1}, considering a 60 kg bw was used in the present paper for the sole scope of this study.

Coumarin (1,2-benzopyrone; CAS number 91-64-5) is a naturally occurring flavouring present in plants and spices. The main source of coumarin in the diet is cinnamon[26] although coumarin content can greatly differ between different types of cinnamon. Cassia cinnamon can contain up to 3000 mg kg^{-1} of coumarin whereas the most refined type of cinnamon, the Ceylon cinnamon, contains only about 8 mg kg^{-1} [27]. Other sources of coumarin include bilberry, celery, and green tea[28]. According to both the European Union and the USA legislation, coumarin cannot be added as such to foodstuffs, whereas it may be present in a foodstuff following the addition of cinnamon. For this reason, maximum permitted levels of coumarin in foodstuffs have been set[29]. Furthermore, EFSA suggests a tolerable daily intake (TDI) of 0.1 mg kg^{-1} bw[30; 31].

Caffeine (1,3,7-trimethylxanthine; CAS number 58-08-2) may be naturally present in foods or added to them. Beverages and foods containing caffeine include coffee, tea, guarana, cola nuts, cocoa, chocolate, energy drinks, and some plants (e.g., mate)[32]. In addition, caffeine may be added to a variety of both prescription and over-the-counter

drugs, which were not part of the present assessment. An officially established TDI or ADI for caffeine does not exist. A review published by Nawrot et al.[33], concluded that for the healthy adult population, moderate daily caffeine intake at a dose level up to 400 mg day^{-1} was not associated with adverse effects. Thus, for the sole scope of this study, the safety limit to caffeine was estimated to be 6.7 mg kg^{-1} bw when using an individual bw of 60 kg.

Food consumption data used for the refined assessment of dietary exposure

Food consumption data used in the refined dietary exposure assessment were collected in the Dutch sample of the EFCOVAL validation study. Between May and July 2007, trained dieticians carried out interviews using a standardised 24-h dietary recall method (EPIC-Soft software) on two non-consecutive days. The two 24-h dietary recalls were collected with at least one month in-between, taking into account weekday variations. The sample consisted of a total of 121 healthy Dutch adults (62 women and 59 men), aged between 45 and 65 years and with all educational levels being represented. However, the participants in the EFCOVAL validation study could not be considered a representative sample of the general population in these strata. The study protocol was approved by the Wageningen University Ethical Committee and informed consent was obtained from all study participants.

EPIC-Soft software

EPIC-Soft is a software program developed in the European Prospective Investigation into Cancer and Nutrition (EPIC) study to ensure the highest possible level of standardisation of 24-h dietary recalls. The structure and standardisation procedure of EPIC-Soft are described in detail elsewhere[34; 35]. An important feature of EPIC-Soft is the use of two complementary food description systems: *explicit* and *implicit*. In the explicit description, facets and descriptors are used during the process of food identification, which is based on the Langual coding system initially used to describe technological and toxicological food characteristics[36]. Facets are used to describe foods in more detail and this is done by means of standardised questions asked to the interviewee each time a food is reported. One of the facets available for a number of food categories is 'flavour or added component'. The descriptors, which are the country-specific terms associated with each facet, are used as pre-defined potential answers built in the database of the software (e.g., strawberry flavour or

strawberry added pieces). In addition, the descriptor 'unknown' may be used when the interviewees are not able to provide the expected level of detail (e.g., unknown flavour for a yoghurt that has been consumed). In the implicit description, the name of a food provides sufficient information to identify the food and no further detail is collected using the facet/descriptor system. For instance, the food name 'liquorice drops' implies the presence of liquorice so that there is no need to use the facet 'flavour' to indicate such presence.

A pre-existing list of facets and descriptors was available in the Dutch software's database since EPIC-Soft has been used in the Dutch National Food Consumption Survey[37]. However, this list was not aimed at the assessment of dietary exposure to flavourings. Therefore, within the EFCOVAL study adjustments were made in the list of descriptors and facets for the identification of foods containing raspberry flavouring. The facet 'flavour' was assigned to new food groups where raspberry may be present, and fourteen new descriptors were included: raspberry, blackberry, blueberry, cranberry, strawberry, cloudberry, loganberry, thimbleberry, bilberry, blackberry, mulberry, berries non-specified (n.s.), red fruits n.s., forest fruit. No further adaptations were made in the descriptors of glycyrrhizinic acid (i.e., liquorice), caffeine (i.e., coffee) and coumarin (i.e., cinnamon) flavourings.

Dietary exposure assessment

With the use of the stepwise approach, dietary exposure to the four flavourings was assessed in three different steps (**Table 1**).

Step 1 – Use of screening methods: Maximized Survey-Derived Daily Intake (MSDI) and Single-Portion Exposure Technique (SPET) as used by JECFA and modified Theoretical Added Maximum Daily Intake (mTAMDI) used by EFSA.

MSDI is also known as the 'per capita method' or 'per capita x 10' approach. Assumptions of the method are: that 60% of total production of flavourings is reported by the industry; that 10% of the total population are consumers of the flavouring; and that there is no variation in the intake of the particular flavouring among consumers. Accordingly, the following formula is used:

$$\textbf{Intake} = \frac{(\text{annual production of the flavouring, kg x } 10^9 \text{ µg kg}^{-1})}{(\text{population of consumers x 365 days})}$$

Final figures were converted into mg kg^{-1} bw day^{-1}

As for the safety evaluations performed by JECFA, the European Union population in this study was assumed to be 32×10^6. The annual production volumes of the flavourings considered were those used by JECFA[20] and SCF[23]: 19.500 kg y^{-1} for raspberry ketone and 1.956 kg y^{-1} for glycyrrhizinic acid, respectively. Poundage data for coumarin is not available since it cannot be used as an added flavouring substance. In the absence of European Union production volumes for caffeine, per capita dietary exposure in the USA was used as a proxy for per capita dietary exposure in the European Union.

The *SPET* method provides a dietary exposure assessment based on normal use levels and identifies the single food category containing the flavouring agent of interest that is likely to contribute to the highest dietary exposure from one 'standard portion'. The standard portion is taken to represent the mean food consumption amount within one eating event for consumers of that food category, assuming daily consumption of one portion over a long period. These standard portions can be found in the 67th and 69th report of JECFA [6; 38]. Thus, the general formula to derive the SPET figure is:

Intake (mg kg^{-1}) = Maximum (standard portion size (mg) x normal use level of the flavouring (mg/kg))

The industry normal use levels used in this study were the ones reported in the Fenaroli's handbook[17], except for glycyrrhizinic acid, for which only upper use levels are reported by industry[39]. For coumarin, the maximum permitted level in foods containing cinnamon[29] was used as replacement of the absent upper use level.

mTAMDI is calculated on the basis of standard portions and normal use levels for flavourable beverages and foods in general, i.e., foods and beverages that may contain the flavouring substance, and for five particular foods groups (exceptions a-e). For instance, exception$_a$ used in this calculation refers to candies and confectioneries[40]. The use levels considered for SPET calculations were also applied to calculate mTAMDI. The general formula used to estimate the mTAMDI[40] is:

Intake (mg kg^{-1}) = (normal use levels in beverages x 324) + (normal use levels in foods x 133) + (normal use levels in exception$_a$ x 27) + (normal use levels in exception$_b$ x 20) + (normal use levels in exception$_c$ x 20) + (normal use levels in exception$_d$ x 20) + (normal use levels in exception$_e$ x 2)

Where normal use levels are in mg kg^{-1}

Table 1 - Stepwise approach used for the assessment of dietary exposure to flavourings in the EFCOVAL study

	Methods	Data Assumptions		Step
		Food Consumption	**Concentration in food**	
Screening	MSDI* or per capita method	Assumption of 10% eaters in the population	Poundage data (industry)	
	Modified TAMDI‡	Assumption of fixed amount of foods and beverages that could contain the flavour (portion sizes per food categories)	Normal use levels (industry)	1†
	SPET§	Assumption of daily consumption of a single food category containing the flavouring agent of interest (highest dietary exposure based on a 'standard portion' size)	Normal use levels (industry)	
Data at individual level using EPIC- Soft	Food group level	Data aggregated in food groups that **MAY** contain the flavour	'Refined' concentration**	2
	Food item level	Data disaggregated: food items that **DO** contain the flavour – without alterations in the list of descriptors in the EPIC-Soft database	'Refined' concentration	3a
		Data disaggregated: food items that **DO** contain the flavour with alterations in the list of descriptors in the EPIC-Soft database	'Refined' concentration	3b††
		Data disaggregated: food items that **DO** contain the flavour plus foods that **MAY** contain the flavour – Same as 3a/3b plus use of descriptor 'unknown' in the facet flavour	'Refined' concentration	3c

Level of Refinement (vertical label at left)

* Maximized Survey-Derived Daily Intake
† Dietary exposure is expressed in mg kg^{-1} bw day^{-1}, considering an individual weighing 60 kg
‡ Modified Theoretical Added Maximum Daily Intake
§ Single-Portion Exposure Technique
** Concentration values from industry (normal use levels) and analytical determinations found in the literature
†† For raspberry ketone only

Screening assessment of coumarin

Because the literature has shown that observed levels of coumarin in food products containing cinnamon can be in fact higher than the maximum permitted level[27], further screening calculations were made to assess the dietary exposure to coumarin by considering the observed coumarin content in cinnamon products. Therefore, extra calculations of SPET and mTAMDI were done with use levels of cinnamon as reported by the Flavour and Extract Manufacturers' Association (FEMA)[17] and assuming a constant of coumarin amounts in two types of cinnamon (cassia cinnamon: 0.3%; Ceylon cinnamon: 0.008%[27]).

Step 2 – Use of food consumption data aggregated in food groups.
At this step, food consumption data at the individual level (Dutch EFCOVAL sample) were grouped in food categories based on the EPIC-Soft grouping system[34]. It was assumed that all foods within a given 'flavourable food category' contained the flavouring of interest (see **Appendix 1**). For instance, raspberry ketone may be added to some foods in the dairy food group (e.g., yogurts). Thus, in the assessment of step 2, all foods belonging to the yogurt category, a subgroup category of dairy products, were assumed to contain raspberry ketone, even though some foods are known not to contain it.

Concentration levels used in step 2 were called 'refined concentrations' (see **Appendix 2**). First choice for the concentration data was normal use levels reported by industry. An exception was made for caffeine contents in non-alcoholic beverages since reported industry levels (0.13 mg/kg) were clearly underestimated as compared to the analytical determinations gathered in the literature (see **Appendix 2**). For glycyrrhizinic acid, upper use levels were used. Analytical determinations from literature were also used in the cases where the flavouring was known to occur in its natural form or when levels of added flavourings were not reported by industry. For instance, glycyrrhizinic acid is known to be added to soy sauce, but use levels in sauces have not been reported. List of references used to collect the flavouring concentration data in foods can be provided upon request.

Step 3 - Use of food consumption data at the level of foods items.
Within this step, three models were created based on the consumption of food items from the Dutch EFCOVAL sample. The first two models (**3a and 3b in Table 1**) considered the consumption of foods that, according to the name of the product or to the use of facets and descriptors available in EPIC-Soft, do contain the flavouring.

The difference between the two steps was that step 3b included information from flavourings after the descriptors of the facet flavour had been extended (for the assessment of raspberry ketone only) in the EFCOVAL study, while step 3a gave information that would have been available before the extension. In the last of the three models (3c), foods were identified in the same way as for step 3a (for glycyrrhizinic acid, caffeine and coumarin) and 3b (for raspberry ketone), but the descriptor 'unknown flavour' was assumed to include the flavouring of interest. Concentration levels used in step 3 were the same as used in step 2: 'refined concentration'.

Data analysis

To estimate dietary exposure to the flavourings, food consumption was multiplied by the concentration of the chemical in the food and then divided by the body weight to be expressed in mg kg-1 bw day-1. In step 1, a body weight of 60 kg was assumed whereas for steps 2 and 3, individually measured body weights were used. Food consumption data in steps 2 and 3 were based, for each individual, on the average of the two 24-h dietary recalls. In these two steps, potential dietary exposure to the flavourings was estimated for each subject. Besides the mean and the median intake of the total group, the 95th percentile of the population distribution was used to characterize highly exposed subjects. As stated by EFSA[41], the 95th percentile can be assessed with approximately 130 subjects when using a binominal distribution[42]. Furthermore, the average contribution of the different food groups to the overall dietary exposure in steps 2 and 3 was estimated in percentages. Data processing and descriptive statistical analyses were performed using SAS software (version 9.1; SAS Institute, Inc., Cary, NC, USA).

Results

All three screening methods (step 1) provided exposure estimates at or over the safety limit for raspberry ketone and, therefore, further refinement of the dietary exposure assessment was needed for this flavouring (**Table 2**). For glycyrrhizinic acid, the MSDI method indicated a dietary exposure at least 16 times lower than the safety limit, whereas the two other methods (SPET and mTAMDI) provided estimates above the safety limit, indicating the need of refinement. Caffeine presented an estimate above the safety limit based on the MSDI method. Additional refinement of coumarin dietary exposure was not necessary based on the screening methods.

However, dietary exposure assessment of coumarin using models of steps 2 and 3 was carried out given that the four selected flavourings were meant to be examples for practical testing of dietary exposure assessment through the use of EPIC-Soft.

Descriptive analyses of dietary exposure assessment using highly aggregated consumption data at the food group level are presented in **Table 3** (step 2). Average and high (95th percentile) levels of exposure to raspberry ketone, glycyrrhizinic acid and caffeine were above the safety limit, indicating the need for more detailed assessment of these three flavourings. Average dietary exposure to coumarin was below the safety limit, despite of the conservative model on food consumption used in Step 2, but above the safety limit at the 95th percentile. Therefore, additional investigation of dietary exposure to coumarin was necessary within the stepwise approach.

Table 3 also presents results of the dietary exposure done at the food item level (step 3). When identifying foods by the name of the product and without using the extended facets and descriptors (step 3a), the mean dietary exposure was under the safety limits, except for caffeine. At the 95th percentile, the dietary exposure to glycyrrhizinic acid and caffeine were three times higher than their safety limit. In the case of raspberry ketone, if considering the adjustments made in the database for facets and descriptors (step 3b), dietary exposure was higher than values obtained in step 3a. In the next step (3c), when not only foods that surely contained the flavouring substance were included in the model but also those for which the flavour was not specified (use of descriptor 'unknown'), mean dietary exposure to raspberry ketone and glycyrrhizinic acid was slightly higher as compared to steps 3a and 3b. In the case of coumarin and caffeine, dietary exposure was the same in all of these steps. By comparing **Tables 2** and **3** it appears that in some cases the screening techniques lead to a dietary exposure lower than that of the refined exposure assessment. It was the case for raspberry ketone where SPET was 0.03 mg kg^{-1} versus 0.04 and 0.05 mg kg^{-1} at the 95th percentile at steps 3b and 3c, respectively.

In the investigation of food groups contributing to the exposure in each step of the assessment (**Figure 1**), it can be seen that for raspberry ketone the main sources of the flavourings were the same in almost all steps: 'dairy products' and 'non-alcoholic beverages'. Yet, while 'non-alcoholic beverages' and 'cakes' were the food groups most contributing to the dietary exposure of raspberry flavouring in step 3a, dietary exposure to raspberry contained in 'dairy products' became an important source with the use of facets and descriptors in step 3b.

Table 2 – Dietary exposure assessment to flavourings (mg kg^{-1} body weight day^{-1}) using screening methods

Flavouring	Safety limit*	MSDI		SPET		mTAMDI	
		Production volume	Dietary Exposure	Concentration source	Dietary Exposure	Concentration source	Dietary Exposure
Raspberry ketone	0.03	IOFI‡	0.05	FEMA§	0.03	FEMA	0.06
Glycyrrhizinic acid	0.16	EFFA**	<0.01	EFFA	3.3	EFFA	3.5
Coumarin	0.10	-	-	EU legislation††	0.05	EU legislation	0.02
-from *Cassia* cinnamon		-	-	FEMA	0.03	FEMA	0.04
-from *Ceylon* cinnamon		-	-	FEMA	<0.01	FEMA	<0.01
Caffeine	6.7	FEMA	7.3°	FEMA	<0.01	FEMA	<0.01

* Raspberry ketone: Threshold of Toxicological Concern in relation to structural class I[19]; Glycyrrhizinic acid: Provisional LOAEL[23]; Coumarin: Tolerable Daily Intake[30]; Caffeine: Tolerable daily intake[33].

† MSDI: Maximized Survey-Derived Daily Intake; SPET: Single-Portion Exposure Technique; mTAMDI: modified Theoretical Added Maximum Daily Intake.

‡ International Organisation of Flavour Industry.

§ Flavour and Extract Manufacturers' Association (US).

** European Flavour and Fragrance Association.

†† Use of maximum permitted levels instead of absent use levels.

° In the absence of EU production volumes for caffeine, per capita dietary exposure in the USA (based on USA production volumes) was used as a proxy for per capita dietary exposure in the EU

Table 3 – Estimated dietary exposure to flavourings (expressed as mg kg^{-1} body weight day^{-1}) in a sample of 121 adults from the Netherlands*(food consumption combined with refined concentration data†)

Flavourings	Safety limits‡	Dietary exposure assessment (Steps 2 and 3)											
		2: Food group level			3: Food item level								
					3a: without improvement of descriptors in the EPIC-Soft			3b: with improvements of descriptors in the EPIC-Soft			3c: all foods from step 3a and 3b plus the foods where 'unknown' descriptor in the facet flavour was reported		
		Mean	Median	P95th	Mean	Median	P95th	Mean	Median	P95th	Mean	Median	P95th
Raspberry Ketone	0.03	0.08	0.07	0.18	<0.01	0	<0.01	<0.01	<0.01	0.04	0.01	<0.01	0.05
Glycyrrhizinic acid	0.16	0.46	0.34	1.37	0.11	<0.01	0.46	-	-	-	0.13	<0.01	0.52
Coumarin	0.10	0.07	0.06	0.12	<0.01	<0.01	0.02	-	-	-	<0.01	<0.01	0.02
Caffeine	6.7	18.6	17.0	43.5	6.81	5.47	16.8	-	-	-	6.82	5.47	16.8

* Sample population is part of the European Food Consumption Validation (EFCOVAL) project

† Concentration values from normal use levels and analytical determinations

‡ Raspberry ketone: Threshold of Toxicological Concern in relation to structural class I[19]; Glycyrrhizinic acid: Provisional LOAEL[23]; Coumarin: Tolerable Daily Intake[30]; Caffeine: Tolerable daily intake[33].
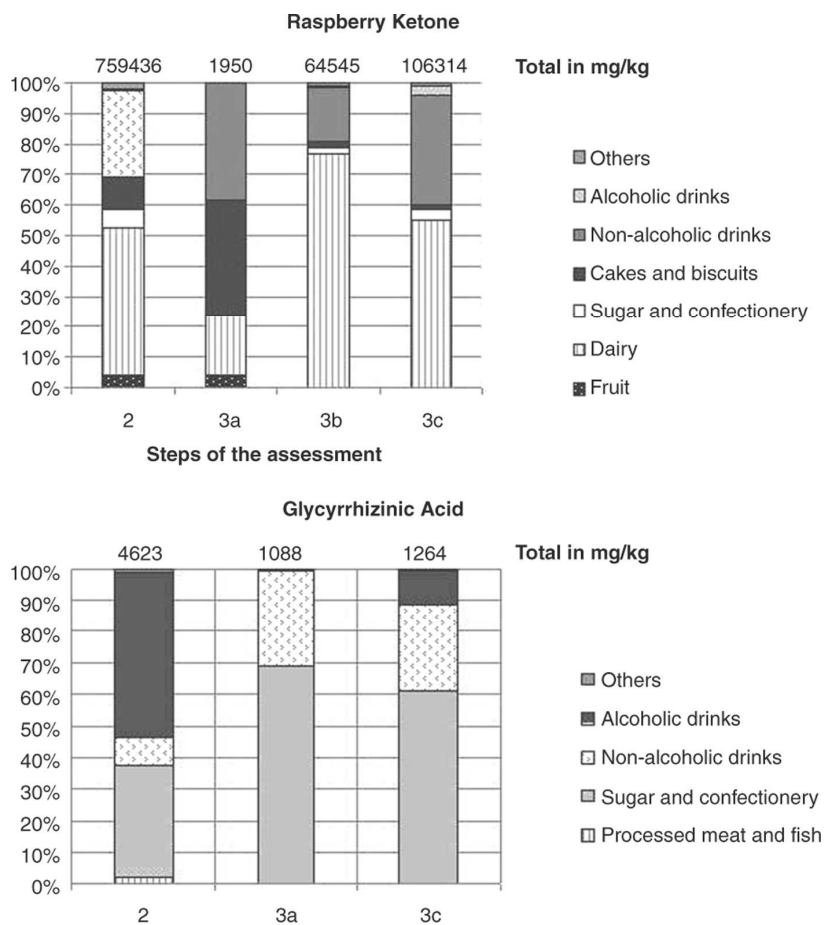
Figure 1 - Dietary exposure to raspberry ketone and glycyrrhizinic acid and their food group sources in each step (2, 3a, 3b and 3c) of the assessment. 2-Food group level: All foods belonging to a flovourable food group are included in the model. 3a-Food item without modifications in EPIC-Soft, 3b-Food item with modifications in EPIC-Soft, 3c-Steps 3a and 3b plus foods which were reported as non-specified flavour.

In the case of glycyrrhizinic acid, 'sugar and confectionery' and 'non-alcoholic beverages' (most herbal teas) were the bigger contributors of the substance in all steps, with a probable overestimation of the contribution from 'sugar and confectionery' in step 2 (group level) as compared with the other steps. The same pattern of overestimation at food group level is seen in the assessment of coumarin and caffeine (figures not shown). Main contributors to dietary exposure were 'cakes', 'biscuits' and 'tea' for coumarin and 'non-alcoholic beverages' for caffeine at all steps.

Discussion

The dietary exposure to raspberry ketone, glycyrrhizinic acid, coumarin and caffeine was estimated in this study using a stepwise approach. It has been shown that the refinement of food consumption data in the assessment of dietary exposure to flavourings might be necessary, but dependent of the chosen screening method for the assessment. When using data from the 24-h dietary recall by means of EPIC-Soft software, the dietary exposure to raspberry ketone was higher in the model where descriptors were extended as compared with the model where no adjustments were considered.

The dietary exposure calculated using the screening methods exceeded the safety limits and therefore implied the need of more refined assessment for raspberry ketone, glycyrrhizinic acid and caffeine, but with somewhat different results depending on the method used and on the flavouring under assessment. In particular, variation in outcomes using different screening methods was observed; whereas by the MSDI method the exposure to glycyrrhizinic acid was evaluated to be of no safety concern, the dietary exposure assessed by SPET and mTAMDI indicated the need of further refinement. On the other hand, dietary exposure to caffeine assessed by MSDI indicated the need of refined assessment while the other two methods did not indicate it. One of the reasons for variation in the results from the screening methods is probably the difference in assumptions between them (e.g., the percentage of consumers in the dietary exposure and how conservative they are, i.e., whether individuals, who consume large quantities of flavoured foods, are considered in the dietary exposure assessed by the different methods). Although it is beyond the scope of this study to investigate the accuracy of such estimates, this topic deserves further attention. In fact, many of the conservative assumptions and default values that are currently used in screening assessments were established some time ago and in some cases they were originally based on subjective or arbitrary estimates[12]. In the case of the MSDI, which until recently was the unique method used by JECFA to assess dietary exposure within the safety evaluation of flavourings, the insufficient conservativeness of the method has been discussed in a number of scientific publications[43-47]. Most recently, JECFA has acknowledged the likely underestimation of the MSDI method in the assessment of some flavourings and developed a new method (SPET), which takes into account different food patterns of consumers and the uneven distribution of dietary exposure in consumers of flavourings[38]. Furthermore, according to EFSA, the appropriateness of the conservative assumptions and default values that are used in screening assessments of chemicals,

including flavourings, may require further investigation. Analysis of uncertainty in the screening assessment may not be required, provided they include proper conservative assumptions to take account of uncertainty[12].

Once the need of further refinement in the dietary exposure is identified, other limitations might be encountered in the assessment of exposure to chemicals in the diet. For instance, the knowledge of chemical concentration data in foods is limited and the ability of dietary methods to assess dietary exposure to chemicals can be uncertain. In our study, this last issue was explored through the different models created to assess the dietary exposure to flavourings in the Dutch population.

In the first model created (step 2), the dietary exposure was characterized by investigating the consumption of flavourings at the food group level. As noted in Table 3, the dietary exposures of the four flavourings were high as compared with all other steps of the assessment. Considering that in this model foods that do not contain the flavouring may have been quantified as part of the dietary exposure, we recognize a certain degree of overestimation in the estimate. This should be, however, an indication of safe dietary exposure in case the estimate would be below the safety limit. However, the need for further refinement of the food consumption data collected at the individual level appeared necessary for the four flavourings under assessment.

With the data on food items collected with the 24-h dietary recall, it has been noted that the adjustments made in the software databases for the raspberry ketone case resulted in a higher dietary exposure to this flavouring. The number of consumers in step 3b, where the new raspberry ketone descriptors have been included, was eight times higher as compared with the step with no modifications in EPIC-Soft (data not shown). This is the result of food consumption data collected at a lower aggregation level and with more details. Assuming that the 24-h dietary recall provided an accurate estimate of the intake of flavoured foods, the high dietary exposure to raspberry ketone in step 3b suggests that such adjustments, which characterize the consumption of foods in more detail, are useful when assessing dietary exposure to flavourings. Nonetheless, 24-h dietary recalls are known to underestimate the dietary intake of some individuals[48; 49] and because of the lack of proper validated biomarkers for these flavourings, the accuracy of the estimate of dietary exposure to flavourings, as assessed, cannot be ensured without further research.

No alterations in the Dutch EPIC-Soft version were implemented for glycyrrhizinic acid, coumarin and caffeine, and evaluation of such alterations was therefore not

possible. However, for some types of flavourings, such as glycyrrhizinic acid, the use of facets and descriptors might not be that important for an accurate dietary exposure assessment given that the food name itself often indicates the presence of the flavouring, which would be enough for the food identification. Nevertheless, additional exploration is needed for this conclusion. Moreover, we do not know to what extent the consumption of cinnamon was correctly identified. First, the use of spices, including cinnamon, during home cooking is not collected during the 24-h dietary recall using EPIC-Soft software. Second, this spice in particular may not be easily identified by the name of the product and neither by the use of descriptors since it does not seem to be clear to the population whether a certain food would contain cinnamon or not. The dietitians of this study reported that for the food group most expected to contain cinnamon (cereals and biscuits), subjects were not able to provide this kind of detail and that they as interviewers had no experience in collecting information about flavourings. Third, the authors of this study may not have correctly identified the presence of cinnamon in certain culinary products such as soups since the presence of cinnamon is not always evident. Because of these reasons, the dietary exposure to coumarin may have been underestimated in this assessment. As a check whether the descriptors of the four flavourings might have been sufficiently identified, the potential flavoured foods with descriptor 'non-specified' were assumed to include the flavouring of interest (step 3c). The dietary exposure did not considerably change for any of the four flavourings in this step.

The assessment of dietary exposure by the different steps and their food group sources gives an indication that such changes in the database (in facets and descriptors) may be food group dependent. Most probable, some degree of uncertainty was present in the assessment at food group level, which tends to overestimate the dietary exposure to flavourings. When the need of more detailed dietary exposure assessment to a specific flavour is identified, the more detailed approach could be limited to a number of food groups, for which it is known that the flavouring can be present and descriptors should be added.

It is important to mention that the estimates presented in steps 2 and 3 of the assessment are not representative of the usual Dutch food consumption. Because of the lack of representativeness of the sample and the limited number of survey days, chronic dietary exposure may not have been correctly estimated in this assessment. In fact, the collection of only two days of 24-h recalls does not allow one to assess chronic exposure but only short-term exposure. This is probably the reason why the refined exposure assessment performed for raspberry ketone leads to higher values

than that obtained with the SPET technique. An improved refined assessment could be performed by using additional information on usual intake of flavoured foods, such as a food propensity questionnaire. Subar et al.[50] have shown that food propensity questionnaires may offer important covariate information in supplementing 24-h recalls for estimating the usual intake of food groups. This is possibly true for assessing chemicals in the diet as well. Furthermore, only dietary exposure has been considered in our assessment and contribution from other sources (e.g., medicines) may lead to an additional exposure. In fact, the safety limits we have used in this assessment should refer to the total exposure to the flavourings, but with the study performed we can only conclude on exposure from the diet. In addition, the small number of evaluated flavourings limits the possibilities to extrapolate the results of our study to other types of flavourings. Another limitation is the scarce availability of concentration data on chemicals. These are relatively seldom published in open literature and therefore difficult to retrieve[7]. In the case of flavouring substances, few analytical data are currently available and little is known about the influence of storage and processing on the residues of these substances in food[12]. Consequently, a high variability in the available concentration data is expected. This study, however, was a first exploration of the possibilities of assessing dietary exposure to food chemicals by using data collected at the individual level with the standardised 24-h recall.

In summary, this study showed that the collection of detailed food consumption data at the individual level is useful and should be further explored for other flavourings. In addition, the possibility of further adaptations of the databases used in EPIC-Soft software seemed to provide a higher dietary exposure to raspberry ketone as compared with the non-modified databases, which may also be true for other flavourings. Yet, the need for alterations may still differ depending on the nature of the flavouring under assessment. To study further the usefulness of detailed food consumption data in the dietary exposure assessment of flavourings and other chemicals, research should include biological markers and analytical determination in flavoured foods, which would warrant a check of the accuracy of such estimates. Finally, the benefit of assessing the usual intake of chemicals in the diet by combining 24-h recalls and food propensity questionnaires is a topic that deserves more exploration.

Acknowledgments

References

1. European Commission (1999) Commission Decision 1999/217/EC of 23 February 1999 adopting a register of flavourings substances used in or on foodstuffs drawn up in application of Regulation (EC) No 2232/96 of European Parliament and of the Council of 28 October 1996. In *Off. J. Eur. Comm.*, pp. 1-137.
2. European Commission (2002) Commission Decision 2002/113/EC of 23 January 2002 amending Commission Decision 1999/217/EC as regards the register of flavouring substances used in or on foodstuffs. In *Off. J. Eur. Comm.*, pp. 1-160.
3. European Commission (2004) Commission Decision of 7 April 2004 amending Decision 1999/217/EC as regards the register of flavouring substances. In *Off. J. Eur. Comm.*, pp. 28-36.
4. European Commission (2005) Commission Decision of 18 May 2005 amending Decision 1999/217/EC as regards the register of flavouring substances used in or on foodstuffs. In *Off. J. Eur. Comm.*, pp. 73-76.
5. EFSA (2009) Flavouring Group Evaluation 18, Revision 1 (FGE. 18 Rev1):Aliphatic, alicyclic and aromatic saturated and unsaturated tertiary alcohols, aromatic tertiary alcohols and their esters from chemical groups 6 and 8. European Food Safety Authority. EFSA-Q-2003-150B. *The EFSA Journal* 978, 1-85.
6. WHO (2009) *Evaluation of certain food additives. Prepared by the sixty-ninth meeting of the Joint FAO/WHO Expert Committee on Food Additives (JECFA). WHO technical report series (952) 2008.* Rome (Italy): WHO.
7. EFSA (2005) Opinion of the Scientific Committee on a request from EFSA related to Exposure Assessments. European Food Safety Authority. EFSA-Q-2003-107. *The EFSA Journal* 249, 1-26.
8. Lawrie CA & Rees NM (1996) The approach adopted in the UK for the estimation of the intake of food additives. *Food Addit Contam* 13, 411-416.
9. WHO (1997) *Food Consumption and Exposure Assessment of Chemicals. WHO/FSF/FOS/97 5.* Geneva (Switzerland): FAO/WHO.
10. Gibney MJ & Lambe J (1996) Estimation of food additive intake: methodology overview. *Food Addit Contam* 13, 405-410.
11. FAO/WHO (2005) *Consultations and workshops : dietary exposure assessment of chemicals in food : report of a joint FAO/WHO consultation, 2-6 May 2005.* Annapolis, Maryland (USA): FAO/WHO.
12. EFSA (2006) Guidance of the Scientific Committee on a request from EFSA related to Uncertainties in Dietary Exposure Assessment. European Food Safety Authority. EFSA-Q-2004-019. *The EFSA Journal* 438, 54.
13. Kroes R, Muller D, Lambe J *et al.* (2002) Assessment of intake from the diet. *Food Chem Toxicol* 40, 327-385.
14. Verger P, Ireland J, Møller A *et al.* (2002) Improvement of comparability of dietary intake assessment using currently available individual food consumption surveys. *Eur J Clin Nutr* 56, S18-S24.
15. Borejsza-Wysocki W (1994) Biosynthesis of p-hydroxyphenylbutan-2-one in raspberry fruits and tissue cultures. *Phytochemistry (Oxford)* 35, 623-628.
16. Gallois A (1982) Rapid determination of p-hydroxyphenyl-1-butanone-3 in raspberries by TLC. *Sciences des Aliments* 2.
17. Burdock GA (2005) *Fenaroli's handbook of flavor ingredients.* 5th ed. Florida (US): CRC Press.
18. Gerasimov AV (2001) Application of planar chromatography with computer processing of chromatograms to the analysis of flavoring materials, using the determination of 4-(n-hydroxyphenyl)-2-butanone (raspberry ketone) as an example. *Zh Anal Khim* 56.
19. Cramer GA, Ford RA & Hall RL (1978) Estimation of toxic hazard—a decision tree approach. *Food Cosmet Toxicol* 16, 255–276.
20. WHO (2001) *Safety evaulation of certain food additives and contaminants. Prepared by the fifty five meeting of the Joint FAO/WHO Committee on Food Additives (JECFA). WHO Technical report series (901).* Geneva (Switzerland): WHO.
21. Stormer FC, Reistad R & Alexander J (1993) Glycyrrhizic acid in liquorice - evaluation of health hazard. *Food Chem Toxicol* 31.
22. Fenwick GR, Lutomski J & Nieman C (1990) Liquorice, Glycyrrhiza glabra L. - composition, uses and analysis. *Food Chem* 38.

23. SCF (2003) Opinion of the Scientific Committee on Food on glycyrrhizinic acid and its ammonium salt. SCF/CS/ADD/EDUL/225, pp. 41 [CotE Communities, editor]. Luxembourg: Office of Official Publications of the EC.

24. WHO (2006) *Safety evaluation of certain food additives and contaminants. Prepared by the sixty-three meeting of the Joint FAO/WHO Expert Committee on Food Additives (JECFA). WHO technical report series (54).* Geneva (Switzerland): WHO.

25. SCF (1991) *Scientific Committee for Food. Report (29th series). Commission of the European Communities, Food Science and Techniques. SCF/CS/ADD/FLAV/61final.* Luxembourg: Office of Official Publications of the EC.

26. Rychlik M (2008) Quantification of Free Coumarin and Its Liberation from Glucosylated Precursors by Stable Isotope Dilution Assays Based on Liquid Chromatography-Tandem Mass Spectrometric Detection. *J. agric. food chem.* 56, 796-801.

27. BfR (2006) *Consumers, who eat a lot of cinnamon, currently have an overly high exposure to coumarin. Federal Institute for Risk Assessment. BfR Health Assessment No. 043/2006.* Berlin (Germany): Federal Institute for Risk Assessment.

28. Felter SP, Vassalo JD, Carlton BD *et al.* (2006) A safety assessment of coumarin taking into account species-specificity of toxicokinetics. *Food Chem Toxicol* 44, 462-475.

29. European Commission (2008) Regulation (EC) No 1334/2008 of the European Parliament and of the council of 16 December 2008 on flavourings and certain food ingredients with flavouring properties for use in and on foods and amending Council Regulation (EEC) No 1601/91, Regulations (EC) No 2232/96 and (EC) No 110/2008 and Directive 2000/13/EC. In *Off. J. Eur. Comm.*, pp. 34-50.

30. EFSA (2004) Opinion of the Scientific Panel on food additives, flavourings, processing aids and materials in contact with food (AFC) related to Coumarin. European Food Safety Authority. EFSA-Q-2003-118. *The EFSA Journal* 104, 1-36.

31. EFSA (2008) Coumarin in flavourings and other food ingredients with flavouring properties Scientific Opinion of the Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food (AFC)(Question No EFSA-Q-2008-677) Adopted on 8 July 2008. *The EFSA Journal* 793, 1-15.

32. Gilbert RM (1984) Caffeine consumption. In *The Methylxanthine beverages and foods: Chemistry, consumption, and health effects.*, pp. 185-213 [GA Spiller, editor]. Berlin (GE): Spring Verlag.

33. Nawrot P, Jordan S, Eastwood J *et al.* (2003) Effects of caffeine on human health. *Food Addit Contam* 20, 1-30.

34. Slimani N, Deharveng G, Charrondiere RU *et al.* (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Methods Programs Biomed* 58, 251-266.

35. Slimani N, Ferrari P, Ocke M *et al.* (2000) 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* 54, 900-917.

36. FDA/CFSAN (1993) LanguaL User's Manual. version 1993.

37. Ocké M, Hulshof K & van Rossum C (2005) The Dutch national food consumption survey 2003. Methodological issues. *Arch Public Health* 63, 227-241.

38. WHO (2007) *Safety evaluation of certain food additives and contaminants. Prepared by the sixty-seventh meeting of the Joint FAO/WHO Expert Committee on Food Additives (JECFA). WHO technical report series (940) 2006.* Rome (Italy): WHO.

39. EFFA (2003) *European Flavour and Fragrance Association (EFFA). Submission to the European Commission, February 20th, 2003 [document number unknown].*

40. EFSA (2004) Opinion of the Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in contact with Food (AFC) on a request from the Commission related to Flavouring Group Evaluation 3 (FGE.03): Acetals of branched- and straight-chain aliphatic saturated primary alcohols and branched- and straight-chain saturated aldehydes, and an orthoester of formic acid, from chemical groups 1 and 2. Question number EFSA-Q-2003-146. *The EFSA Journal* 107, 1-59.

41. EFSA (2008) Guidance document for the use of the Concise European Food Consumption Database in Exposure Assessment, Parma, EFSA/DATEX/2008/01. *The EFSA Journal*, 1-8.

42. Conover WJ (1971) *Practical Nonparametric Statistics.* New York: Wiley.

43. Lambe J, Cadby P & Gibney M (2002) Comparison of stochastic modelling of the intakes of intentionally added flavouring substances with theoretical added maximum daily intakes (TAMDI) and maximized survey-derived daily intakes (MSDI). *Food Addit Contam* 19, 2-14.

44. Arcella D & Leclercq C (2005) Assessment of dietary intake of flavouring substances within the procedure for their safety evaluation: advantages and limitations of estimates obtained by means of a per capita method. *Food Chem Toxicol* 43, 105-116.

45. Leclercq C (2007) Issues arising when methods used to assess dietary exposure to flavouring substances are compared. *Food Chem Toxicol* 45, 2336-2337.

46. Munro IC & Danielewska-Nikiel B (2006) Comparison of estimated daily intakes of flavouring substances with no-observed-effect levels. *Food Chem Toxicol* 44, 758-809.

47. Hall RL & Ford RA (1999) Comparison of two methods to assess the intake of flavouring substances. *Food Addit Contam* 16, 481-495.

48. Willett W (1998) *Nutritional epidemiology*. 2nd ed. New York: Oxford University Press.
49. Bingham S (1987) The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutr Abstr Rev* 9, 705-742.
50. Subar AF, Dodd KW, Guenther PM *et al.* (2006) The Food Propensity Questionnaire: Concept, Development, and Validation for Use as a Covariate in a Model to Estimate Usual Food Intake. *J Am Diet Assoc* 106, 1556-1563.

Appendix 1 – Flavourable food groups considered in the dietary exposure assessment of step 2

Raspberry Ketone	Glycyrrhizinic Acid
• Alcoholic beverages (liqueurs, brandies, gin subgroups) • Biscuits • Breakfast cereals • Cakes • Dairy products (yogurt, milk beverages, cream desserts, puddings subgroups) • Dressing and dessert sauces • Fruits • Non-alcoholic beverages • Sugar and confectionery (jams, non-chocolate, ice cream, sorbet, water ice subgroups)	• Alcoholic beverages (spirit, aniseed drinks and liqueur subgroups) • Dressing sauces • Fish products • Liquorice confectionery (non-confectionery chocolate and ice cream subgroups) • Non-alcoholic beverages (herbal tea subgroup) • Processed meats
Coumarin	Caffeine
• Alcoholic beverages (wine, beer subgroups) • Biscuits • Breads • Breakfast cereals • Cakes • Dairy products (desserts subgroup) • Dessert sauces • Fruits • Non-alcoholic beverages • Root vegetables • Sugar confectionary	• Alcoholic beverages (liqueur subgroup) • Breakfast cereals • Cakes and biscuits • Dairy products (milk beverages and cream desserts subgroups) • Dessert sauces • Non-alcoholic beverages (carbonated drinks, coffee, tea subgroups) • Sugar and confectionery (syrup, chocolate bar, ice cream subgroups)

Appendix 2 – Refined concentrations[*] used in the dietary exposure assessment of steps 2 (food group level) and 3 (food item level)

Raspberry Ketone		*Glycyrrhizinic Acid*	
Foods	mg kg^{-1}	Foods	mg kg^{-1}
Baked goods	13.1	Alcoholic beverages	135
Chocolate (n=2)	9.3 ± 9.1	Liquorice confectionery	1500
Ice cream	2.6	Non-alcoholic beverages	50
Jam	0.3	Soy sauce (n=5)	37 ± 19.4
Non-alcoholic beverages	2.8		
Raspberry (n=39)	1.3 ± 1.2		
Sauce	0.9		
Yogurt	20.2		

Coumarin[†]		*Caffeine*	
Foods	mg kg^{-1}	Foods	mg kg^{-1}
Baked goods	16.6	Baked goods	0.06
Bilberry	0.005	Brewed Coffee	680
Breakfast cereal	7.5	Cocoa powder	340
Celery	16.6	Chocolate milk	60
Cinnamon powder	3000	Chocolate syrup	106
Dairy products	1.1	Cola drinks	125
Frozen dairy	1.1	Dark chocolate	700
Jam	2.9	Espresso	2473
Non-alcoholic beverages	0.06	Energy drinks	240
Pudding	3.8	Puddings	0.3
		Frozen dairy	0.3
		Liquor	170
		Milk chocolate	220
		Tea	205
		White chocolate	14

[*] Number of samples used = 1 unless otherwise specified; in that case mean +/- SD are reported. See methods section.

[†] Based on percentage of coumarin in cinnamon products (cassia cinnamon: 0.3%[27]) and in the use levels of cinnamon reported by FEMA (Flavour and Extract Manufacturers' Association) at Fenaroli's Handbook of flavour ingredients[17], except for bilberry and celery, which were collected from the literature.

futurepaneuropeanfoodconsu
mptionsurvey24hrecallepicso
ftefovald   iasunce
rtaint   data
portic   nvers
ionfa   suali
ntakesbiomarkersofdietaryint
akemultilevelassessmentexpo
suretochemicalsgeneralisatio
nimplicationsfindingswayfor
wardconcludingremarkssuffic
ientlyvalidmethodsuitableco
mparisondatamonitoringdiet

Chapter 7

Discussion

In view of a future pan-European food consumption survey, the aim of this thesis was to evaluate the data obtained with two non-consecutive 24-h recalls using EPIC-Soft for comparisons of dietary intake between European populations. The following paragraphs summarize and discuss the main findings of this thesis. In addition, recommendations for future research and implications for a future pan-European survey are given.

Main findings

Initially (chapter 2), we observed that protein intake, estimated by two non-consecutive 24-h recalls, was underestimated by 8% in both genders and potassium intake by 7% in men and 4% in women. The 24-h recall method appeared to provide sufficiently valid data for comparing mean protein and potassium intake between European populations although variability in bias of 4-7% can be expected. Additionally, the method appeared to be sufficiently valid for assessing and comparing the protein and potassium adequacy of populations, based on cut-off points at the ends of the intake distribution and less appropriate to assess other parts of the distribution. Next, we substantiated the results obtained in chapter 2 by combining the data collected in EFCOVAL with data from the European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study. Consequently, data from 14 centres were used to perform linear multilevel analysis (Chapter 3). Based on that analysis, we observed that the mean bias in protein intake for both genders and in potassium intake for women, collected with a single 24-h recall using EPIC-Soft, did not vary across centres and to a certain extent varied for potassium intake in men (coefficient of variation = 9.5%). In both analyses (Chapters 2 and 3), BMI was the factor mostly explaining the variation in bias in the observed protein and potassium intake across European populations. Subsequently, more evidence about the performance of the method was obtained (Chapters 4-6). Using data from the EFCOVAL validation study (Chapter 4), we observed that two standardised 24-h recalls using EPIC-Soft and a food propensity questionnaire (FPQ) appeared to be appropriate to rank individuals according to their fish and fruit and vegetable intake in a comparable way in five European centres. Then, we evaluated the impact of design aspects of 24-h recall assessments on bias in protein and potassium intake within the European centres (Chapter 5). We observed that protein intake reported by face-to-face interviews at the study site was less accurate than by telephone interviews in some centres (~10% difference). In addition, we concluded that second 24-h recall

assessments seemed to be less accurate than first recalls (~10% difference) and that, in one out of five centres, protein intake estimated during weekends and potassium intake estimated during weekdays were less accurate than during other days of the week (~12% difference). Finally, we observed that detailed dietary data collected at the individual level may be necessary to assess the dietary exposure to flavourings. In addition, the possibility of adaptations of the databases used in EPIC-Soft software seemed to provide more detailed information for the dietary exposure to the flavouring raspberry ketone, as compared to non-modified databases (Chapter 6).

Uncertainties in our findings

The following paragraphs address uncertainties in our evaluation about the use of food composition data, portion sizes, conversion factors, usual intake modelling, biomarkers, multilevel assessments, and about dietary exposure assessments to chemicals.

Food composition data

> *"The limitations of food composition tables or databases are often not sufficiently understood"*
>
> *(Greenfield and Southgate, 2003)* [1]

Food composition data are used to convert the intake of foods into nutrients, as done for calculation of protein and potassium intake using data of 24 hour recalls in this thesis. It is well known that the use of food composition tables (FCTs) may result in errors in nutrient intake estimates and that these will invariably contribute to the total error in dietary intake assessments[2-5]. Consequently, the observed variations in bias in estimates of protein and potassium intake between-centres in our analysis could have been either smaller or larger. We expect that the recent efforts of the European Food Information Resource Network (EuroFIR) will result in harmonised and sustained food composition data across Europe[6], thus contributing to the minimization of differences in the accuracy of estimated nutrients between European countries.

Another important aspect about FCTs is the fact that EPIC-Soft software does not automatically link food intakes to their nutrient composition. In addition, the number of unique foods that are reported in EPIC-Soft is generally much larger than the

number of codes available in FCTs[7], largely because of the high level of detail that is asked with the method. Consequently, subjective decisions can be expected in coding during the linkage process and invariably errors can also occur. To minimize such errors, an automated and transparent procedure, with country-specific considerations to link intake data to food composition databases in EPIC-Soft, is an option to be considered during future developments of this computerised 24-h recall. In addition to this, non-specified food descriptions can be used when an interviewee is not able to provide the level of detail that is requested for the foods. Procedures to deal with these non-specified foods are therefore needed. For example, an interviewee may not be able to determine if a consumed food was fortified or of organic production. As suggested in EFCOVAL, procedures to deal with these non-specified reports could be based on the use of probabilistic approaches, techniques for imputation of missing values, or a facility to collect a posteriori missing information from the participant[7].

Portion size estimation

> *"Where portion size is to be assessed in several groups whose dietary habits differ, a single set of standard portions may not reflect the true variation in portion size"*
>
> *(Nelson, 1997)[8]*

To estimate food portions in EPIC-Soft, household measures, weight/volume estimates, standard units, bread shapes and photographs are used by the interviewees. While estimating food portions, systematic and random errors may occur. In a study of errors in portion size estimation using drawings of breads from EPIC-Soft, a sample from Belgium (n=111) overestimated their mean bread intake by ~11%[9]. This type of error may have affected the bias in protein and potassium intake differently across the centres (chapters 2, 3, and 5) as well as the ranking of individuals according to their usual fish and fruit & vegetable intake (chapter 4).

Furthermore, some studies suggested that errors in the assessment of portion sizes may be dependent on the BMI of the subjects[10]. We also found that BMI was an important contributor to possible differences in protein and potassium bias across European populations. Therefore, further development and evaluation of tools to estimate portion sizes in EPIC-Soft may enhance comparability of dietary data between countries in future surveys. For example, it would be worthwhile to gain more knowledge on the reasons why specific groups of interviewees, such as obese subjects, have difficulties in estimating consumed portion sizes. Additionally, it would

be valuable to further develop country-specific portion sizes of foods, especially photographs, across Europe. Furthermore, the EFSA EGFCD* group recommended that the suitability and accuracy of the method used for food portion estimations in a future pan-European survey needs to be validated and examined periodically[11].

Conversion factors (yields, weight changes, fat absorption)

"A long-term policy could be to compile yields or weight changes […] to compare the various national values"
(Bergström, 1994)[12]

In EPIC-Soft, algorithms are used to systematically convert the amounts of reported foods to their final mode of consumption, e.g., cooked without edible part. Accordingly, the conversions from volume to weight, from raw to cooked weights, from foods with to without inedible parts and considering fat absorption are calculated automatically by the software using standardised country- and food-specific conversion factors. Whilst these algorithms contribute to the standardisation of EPIC-Soft results, the quality of used conversion factors remains uncertain. Thus, it is recognised that food and nutrient intake in this thesis can be prone to errors due to inaccurate conversion factors. Unfortunately, the actual contribution and direction of possible errors in conversion factors is not known. In EFCOVAL, Ocké and colleagues[7] pointed out that the quality and transparency of used conversion factors in EPIC-Soft can still be improved. For example, this can be done by documenting the sources of these coefficients and sharing them between countries.

Assessment of usual intakes

"The general idea […] was to go on with improving dietary assessment methods until we can put little camera's in the mouth of our subjects and we accurately can measure all the people eat during a longer period"

(van Staveren, 2003)

With the use of two replicates of 24-h recall, usual dietary intakes were modelled by the multiple-source-method (MSM). MSM was developed within the EFCOVAL

---

* Expert Group on Food Consumption Data

consortium[13]. Yet, other statistical methods are available, e.g., National Cancer Institute (NCI)[14], NUSSER[15] and AGEMODE[16]. Such statistical methods provide means to remove the day-to-day variance from daily intakes and obtain an estimate of the usual intake distribution with a more precise variance. Nevertheless, it seems that smaller sample sizes, high within-person variations and pronounced departures from normality can lead to uncertain usual intake distributions[17; 18]. Additionally, the recommended minimum number of two days for estimating usual intakes with those statistical methods refers to the estimation of group distributions[19], while it has not been investigated for estimating individual usual intakes. Thus, the relatively small sample sizes of our analyses (Chapters 2 and 4) and the rather skewed distributions observed for fish intake in our samples together with the estimation of individual's usual intake using only two days (Chapter 4) may have resulted in imprecise estimates of the estimated usual intakes. According to a recent EFSA workshop that discussed different available methods to estimate usual dietary intake, there is apparently not one optimal method for estimating usual intakes for all cases, and the choice of the most appropriate method needs to be fine-tuned case by case[18]. Among others, this choice may depend on the need to consider additional food frequency data, left-censored data, and brand loyalty[18]. An advantage of using a method such as MSM in our analyses was that information of whether a person was a true consumer, as based on an FPQ, could be integrated into the estimation procedure. This seemed to be essential for the evaluation of fish intake - an episodically consumed food group in our assessment - because most centres presented a high proportion of zero values based on 24-h recall assessments.

Biomarkers of dietary intake

"Dietary intake cannot be estimated without error and probably never will be"

(Beaton, 1994)[20]

Generally, there are three main requirements for using biomarkers as references of dietary intake. First, the biomarker has to have a clear dose-response relationship with the intake – sensitivity. Second, the biomarker should reflect the cumulative effect of diet over the desired period of time – time integration. Third, the biomarker should not (or little) be influenced by non-dietary determinants, such as genetic and lifestyle factors[21]. Recovery biomarkers in general fulfil these requirements provided the proper collection of samples and handling of analytical techniques. For concentration biomarkers, on the other hand, the fulfilment of such requirements is more complex

because this depends on the degree of homeostatic control of the nutrient in available biologic samples, on the range of intake in the population, and on the existence of determinants other than intake[21]. Consequently, firm conclusions about the evaluation of food group assessments collected with 24-h recalls, using serum carotenoids and fatty acids in phospholipids as reference method, across European populations are not fully warranted.

Especially, the use of serum carotenoids to evaluate the estimates of fruit and vegetables intake across different populations was uncertain. This was because of the variety of fruits and vegetables consumed across European populations and of the different contents of carotenoids in specific fruits or vegetables. As most of the evidence for using serum carotenoids as reference method for fruit and vegetable intake in comparisons of populations comes from observational studies[22; 23], further investigations based on controlled feeding trials are needed. Such investigations for evaluating the use of fatty acids as biomarkers of fish intake may also be necessary because n-3 fatty acids are mainly present in fatty fish and differences in type of fish consumed across European populations seem to exist[24].

Multilevel assessment

"While there is an increasing literature surrounding the use of multilevel modelling, it is almost all […] in a language that is difficult for many epidemiologists to understand"

(Bowen, 2007) [25]

The multilevel analysis carried out in the third chapter of this thesis deserves some further consideration. The modelling approach used to explain the variability of bias across European populations may not have been necessary, if judged by testing whether the intercepts varied across centres (data not shown). However, the multilevel analysis was assumed to be more informative than a fixed analysis, because our a priori hypothesis was that the validity of the 24-h recall methodology may truly be different between countries. By using multilevel analysis, we were able to demonstrate whether characteristics at the individual- or centre-level could explain the postulated heterogeneity in bias across the centres when disentangling the within- and between-centre variance components. Likewise, such separation of variance components may be of interest for the analysis of upcoming results of a future pan-European food consumption survey or any other nutritional survey that has collected data from individuals being clustered within groups.

Furthermore, multilevel analyses were not performed to evaluate the ranking of European centres according to their fish and fruit & vegetable intake compared to their respective biomarkers. However, the multilevel assessment for food groups may be of special importance since we expect a larger variability in errors in those assessments than in those of the nutrients. The main reason why multilevel analyses for the food groups are not presented in this thesis is that taking the difference or ratio of intake and biomarker values, as done for protein and potassium, does not make sense when concentration markers are used as reference; further thoughts about incorporating the concentration biomarker in a multilevel assessment are needed.

Gaps in the evaluation of the dietary exposure to chemicals

"Methods, databanks, as well as statistical tools that improve the comparability of the [chemical] exposure assessment in European countries are becoming more important"
(Kroes and colleagues, 2002)[26]

Within EFCOVAL only one chemical category, flavourings, was chosen to narrow down the evaluation of usefulness of EPIC-Soft for the chemical exposure assessment (Chapter 6). The assessment of flavourings in the diet was limited by not including an independent estimate (i.e., biomarker) for assessing the accuracy of the estimated flavourings in the diet. Although the inclusion of biomarkers for assessing chemical exposure assessment was considered and discussed among experts in EFCOVAL, we did not succeed in finding a suitable biomarker for this type of assessment. Other than for flavouring substances, a few biomarkers for chemical exposure assessment are available (e.g, mycotoxin, aflatoxin, dialkylphthalates) and their use to evaluate chemical exposure assessment could be explored in future.

Concentration and occurrence data of chemicals in foods, which may also be recognised as a type of food composition data, are a major source of error in the assessment of chemicals in the diet. For many chemicals, the availability of concentration data is scarce[27] and many assumptions are needed on the contents of chemicals in the diet when assessing dietary exposure to chemicals. That is the case not only when using dietary intake data estimated from 24-h recalls, but also when using aggregated data in worse case scenarios. In the case of dietary exposure to flavouring substances estimated in this thesis, it was sometimes assumed that the concentration data were the same for all foods within a food group. Consequently,

this may have led to an over- or underestimated intake of flavourings. In this context, initiatives like the FACET[*] project[28] and the EFSA expert group for chemical occurrence data[29] can be helpful. Both initiatives are committed to improving the quality of concentration and occurrence data in foods, which will hopefully lead to better dietary exposure assessment of chemicals in Europe.

Generalisation

In this section, the arguments for generalisation of the presented results to other nutrients and food groups, flavourings and chemicals, countries and food patterns, and population groups are described.

Generalisation to other nutrients and food groups. In EFCOSUM, the most relevant dietary components identified for monitoring in a pan-European food consumption survey were the food groups vegetable, fruit, bread, and fish, and the nutrients saturated fatty acids, total fat and alcohol. In the EFCOVAL study and this thesis, however, we only evaluated the accuracy of the intakes of fruit, vegetable, fish, protein and potassium. This was done because of the limited availability of biomarkers of intake and the uncertainty in the food composition values for some other nutrients, such as n-3 fatty acids and carotenoids, in some of the centres. Any generalisation from our results to other foods or nutrients needs caution because evidence exists that some foods and related nutrients might be selectively misreported[30-32]. For instance, the OPEN study in the US showed that underreporting of energy was somewhat greater than of protein, suggesting a bias toward more underreporting of fat, carbohydrates, or alcohol[32]. Likewise, our results showed that bias in observed protein intake was different from bias in potassium intake, especially for men based on the multilevel analysis. Additionally, we did not consider nutrients that can be ingested from dietary supplements or fortified foods in the analyses of this thesis. As the contribution of supplements or fortified foods to the total nutrient intake can be substantial, extra caution when generalizing our results to such nutrients is necessary because smaller or larger bias in intake collected with the 24-h recalls may be observed.

---

[*]Acronym for: Flavourings, additives and food contact material exposure task

Generalisation to other flavourings and chemicals. Similarly, we have only included the flavouring category to explore the assessment of chemicals in the diet. Although the results of this exploration were promising, yet more than 2700 flavourings are currently available to be added to foods and beverages in Europe and only the assessment of four flavourings were evaluated in Chapter 6. However, because some of these hundreds of flavourings share similarities in the way that they can be identified in EPIC-Soft, extrapolation of our results to other flavourings may be warranted. Any further extrapolation to the variety of chemicals in the diet needs caution. Particularly, the data for assessing other chemicals may require a different level of detail than was studied for flavourings. For example, the dietary exposure to residues of contact materials are mainly retrieved from recalling the material of packing from the food that was consumed whereas the presence of acrylamides in the diet is mainly identified by the type of cooking method.

Generalisation from centres to countries and to other food patterns. It should be considered whether the findings of this thesis can be extrapolated from centres to their respective countries or even to other European countries that will be part of a pan-European survey as well as to large countries with subpopulations of different socioeconomic classes and food patterns, such as the U.S. population. Especially, the EFCOVAL centres are not representative of their country and a selective sample may have been included with motivated and health-conscious participants, who could have learned about the type of assessment because urine collections and therefore also 24-h recalls over the same days had to be planned. Nevertheless, the centres involved in the assessments of this thesis included subjects with different socioeconomic backgrounds and a large heterogeneity in food patterns, with representations from the diverse European patterns[24; 33; 34]. In fact, our results (Chapters 2 and 3) showed that the educational level of individuals, the human development index (a proxy for socioeconomic status of the centres) and the food pattern[*] of the centres did not affect the accuracy of protein and potassium intake estimated across different European populations. Instead, individual BMI and design aspects of the assessment were seen as determinants of mis-reporting of food consumption. Thus, we may expect that the learned lessons from this thesis can also be useful to the application of the 24-h recall method in other European countries and other large countries with subpopulations with a diverse food pattern.

---

[*] See chapter 3 for the definition of food pattern.

Generalisation to other population groups. Moreover, the generalisation of our results to age groups other than adults cannot be totally granted. In particular, the unique use of 24-h recalls in children of different age groups has not been recommended by EFCOSUM, EFCOVAL or EFSA[11; 35; 36] mainly because the help from parents or guardians is necessary to estimate their intake. In EFCOVAL, a repeated one-day food record was recommended for preschoolers (4-6 y) and a repeated 24-h recall method together with a registration booklet to record foods eaten out of home for school children (7-14 y)[36]. Nonetheless, the EFSA EFGCD group recommended the sole use of a dietary record method for children (36 m to 10 y)[11]. It has also been suggested by the EGFCD[*] group to use EPIC-Soft as data-entry system for the dietary records collected in children, as done in the Dutch food consumption survey[37]. This use was tested within the 'pilot study for the assessment of food consumption and nutrient intake among kids in Europe' (PANCAKE) and was considered to be a useful approach[38]. In the case of elderly people, the use of 24-h recalls may not be a problem in healthy populations[35; 39]. However, the fact that the method is influenced by the cognitive skills and other functional abilities of elderly subjects[39] does not assure that our results from adults would be similar in elderly populations.

Some implications of our findings

This thesis showed that some important determinants of mis-reporting of food consumption can affect the quality of data and their comparability across diverse European populations. BMI was shown to be one of the most important determinants of misreporting of foods across Europe, not only predicting the bias in protein and potassium intake within the centres, but also explaining differences between them. This adds evidence to literature in the area[40; 41] and suggests that any future comparison of 24-h recall data needs to take into account the BMI of individuals.

Furthermore, careful choices have to be made for the application of 24-h assessments in future pan-European food consumption surveys. For example, as our results suggest different performance of the method depending on the mode of application (face-to-face vs. telephone interviews). If countries opt to use different modes of administration for applying 24-h recalls in future pan-European surveys, the data collected needs to be analysed and interpreted considering these differences.

---

[*] Expert Group on Food Consumption Data

Lastly, it is of interest to compare the validity results of the 24-h recall under evaluation in this thesis to the results of other methodologies used in national food consumption surveys. However, it appears that no comparable evaluation studies have been performed in other populations. An exception to that is the OPEN study that evaluated the 24-h recall method used in the US-American and Canadian food consumption surveys and showed that protein intake collected with two days of the 24-h recall had a similar percentage of underestimation (11-15%) as observed for the 24-h recall using EPIC-Soft in this thesis (Chapter 2).

The way forward

In view of a future pan-European food consumption survey and the further development of dietary assessments, some suggestions to move these fields forward are given below.

First, the results about the quality of the data collected with 24-h recalls using EPIC-Soft and their comparability across different European populations are promising. However, one can question whether the comparability of these data can be warranted from now on, especially if potential rapid changes in dietary habits and food supply are considered. It is difficult to be certain about it and we suggest that would be worthwhile to collect biomarker information in subsamples in future surveys. This would require ongoing attention for quality control of the 24-h recall method and further evaluations that have not been possible yet, as for instance the accuracy of assessing dietary exposure to chemical substances.

Second, considering that few biomarkers are available to evaluate the intake obtained with dietary methods, it is certainly valuable to explore the development of new biomarkers, especially the recovery ones. Suitable biomarkers for evaluating the assessment of other nutrients, foods and chemical substances have to be searched. A possible option is the further development of the new class of predictive biomarkers, as has been recently proposed for assessing sugar intake[42]. In addition, a possible useful field for developing biomarkers may be of the nutritional metabonomics for evaluating dietary patterns[43].

Third, although statistical methods to model the usual intake distributions from 24-h recalls have greatly developed in the past years, there is actually not one best method for estimating usual intakes. In particular, data on usual intakes to be estimated in future surveys may be used for different purposes and we suggest that currently available methods to estimate usual intakes from short-term intakes can still be

improved. For instance, we should have a more clear understanding of the implications of applying such methods in data that may not have a normal distribution even after transformation, or in replicate data for which independent errors cannot be assumed. Of utmost importance, those methods should be able to incorporate the different types of information that are needed for the assessment of foods, nutrients or chemicals in the diet, e.g., brand loyalty and additional frequency data.

Furthermore, FCT and portion size quantification are considered important sources of errors in the application of 24-h recalls, or any other dietary method. Although considerable improvements in the quality of FCTs in Europe were achieved by EuroFIR, the continuous evaluation of FCTs and portion size estimations in future surveys is necessary. A possible way to evaluate these sources of errors is the quantification of uncertainties that are related to them, as also addressed in EFCOVAL[44]. However, because those analyses require knowledge on the amount of uncertainty to be expected (e.g., coefficients of variation in portion size), further experimental studies are first desired. Moreover, some technological advances in dietary assessments could also be useful to avoid errors from FCTs and food portion quantification[45]. For example, to improve the quantification of portion sizes, 3-D pictures could be used in computerised 24-h recalls.

We also recommend further studies to elucidate the differences in the misreport of foods and related nutrients between men and women, as observed for potassium intake in our study but not for protein. Together with the fact that potassium is a nutrient present in a larger variety of foods/food groups and more equally distributed among different food groups than protein, it could be that some specific foods or food groups that contain potassium are differently misreported by men and women.

At last, partially related to the former, we recommend further research to appreciate the roles of cognitive and communicative aspects toward the data quality of standardised 24-h recalls, which at the moment are not clearly demonstrated. For instance, cognitive research suggests that individuals remember in different ways and that different interpretations may be given for a question or response, especially in the presence of an interviewer[46; 47]. Thus, sources of errors in 24-h recall assessment, such as from memory or interviewer, may be better elucidated. Results from this type of research may eventually be helpful to account for differences in the performance of the method between men and women.

Concluding remarks

The work described in this thesis gives evidence to conclude that the use of 24-h recalls (EPIC-Soft) for estimating and comparing protein, potassium, fish and fruit and vegetable intake across European populations is a sufficiently valid method, and thus, suitable to be used in future pan-European food consumption surveys. The combination of 24-h recalls with a food frequency or propensity questionnaire is necessary for estimating and comparing the intake of non-episodically consumed foods. It is also necessary that the data from 24-h recalls will be properly analysed and interpreted considering characteristics that may influence the report of dietary intake across countries, especially BMI and mode of administration. Additionally, it can be concluded that 24-h recalls using EPIC-Soft provide useful dietary data for the dietary exposure assessment to flavouring substances.

Furthermore, difficulties in the application of the 24-h recall method in some settings are acknowledged (e.g., collection of Saturday's intake in some countries) and the developments for improving the performance of the method are encouraged, but this should not prevent the use of the method. Over and above, it is essential that recruited interviewers are properly qualified and extensively trained, as was the case in our evaluations.

Lastly, advances in the field of dietary assessment of populations are attained by providing evidence that the 24-h recall can deliver sufficiently comparable data for monitoring dietary intake of European countries. From a public health point of view, this is very important for assessing the adequacy and safety of dietary intake, the development of food and nutrition policies and strategies, and the development and evaluation of prevention programs in Europe.

Note

The work in this thesis is part of the European Food Consumption Validation (EFCOVAL) study. Complete evaluations and recommendations from EFCOVAL, other than previously presented, are published elsewhere[7; 17; 36; 44; 48-56].

Reference

1. Greenfield H & Southgate DAT (2003) *Food composition data : production, management and use*. 2nd ed. Rome: FAO Publishing Management Service.
2. West CE & van Staveren WA (1997) Food consumption, nutrient intake, and the use of food composition tables. In *Design concepts in nutritional epidemiology*, pp. 107-122 [BM Margetts and M Nelson, editors]. Oxford etc.: Oxford University Press.
3. Slimani N, Ruth Charrondiere U, van Staveren W *et al.* (2000) Standardization of Food Composition Databases for the EPIC: General Theoretical Concept. *J Food Comp Anal* 13, 567.
4. Slimani N, Deharveng G, Unwin I *et al.* (2007) Standardisation of an European end-user nutrient database for nutritional epidemiology: what can we learn from the EPIC Nutrient Database (ENDB) Project? *Trends Food Sci Technol* 18, 407-419.
5. Deharveng G, Charrondiere UR, Slimani N *et al.* (1999) Comparison of nutrients in the food composition tables available in the nine European countries participating in EPIC. European Prospective Investigation into Cancer and Nutrition. *Eur J Clin Nutr* 53, 60-79.
6. Egan MB, Fragodt A, Raats MM *et al.* (2007) The importance of harmonizing food composition data across Europe. *Eur J Clin Nutr* 61, 813-821.
7. Ocke MC, Slimani N, Brants HA *et al.* (2011) Potential and requirements for a standardized pan-European food consumption survey using the EPIC-Soft software. *Eur J Clin Nutr (in press)*.
8. Nelson M (1997) The validation of dietary assessment. In *Design concepts in nutritional epidemiology*, pp. 241-272 [M Nelson, editor]. Oxford etc.: Oxford University Press.
9. De Keyzer W, Huybrechts I, De Maeyer M *et al.* (2011) Food photographs in nutritional surveillance: errors in portion size estimation using drawings of bread and photographs of margarine and beverages consumption. *Br J Nutr* 105, 1073-1083.
10. Nelson M, Atkinson M & Darbyshire S (1994) Food photography. I: The perception of food portion size from photographs. *Br J Nutr* 72, 649-663.
11. EFSA (2009) General principles for the collection of national food consumption data in the view of a pan-European dietary survey. *The EFSA Journal* 7, 1-51.
12. Bergström B (1994) Nutrient losses and gains in the preparation of foods., available at http://www.slv.se/upload/dokument/rapporter/mat_naring/1994_32_Livsmedelsverket_nutrient_losses_and_gains.pdf available on 14 Mar 2011.
13. German Institute of Human Nutrition (2009) The Multiple Source Method (MSM). https://nugo.dife.de/msm/
14. Tooze JA, Midthune D, Dodd KW *et al.* (2006) A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc* 106, 1575-1587.
15. Nusser SM, Carriquiry AL, Dodd KW *et al.* (1996) A Semiparametric Transformation Approach to Estimating Usual Daily Intake Distributions. *J Amer Statistical Assoc* 91, 1440-1449.
16. Waijers PM, Dekkers AL, Boer JM *et al.* (2006) The potential of AGE MODE, an age-dependent model, to estimate usual intakes and prevalences of inadequate intakes in a population. *J Nutr* 136, 2916-2920.
17. Souverein OW, Dekkers A, Geelen A *et al.* (2011) Comparing four different methods to estimate usual intake distributions. *Eur J Clin Nutr (in press)*.
18. EFSA (2010) Statistical modelling of usual intake. Scientific report. EFSA-Q-2009-00841. 1-23.
19. Hoffmann K, Boeing H, Dufour A *et al.* (2002) Estimating the distribution of usual dietary intake by short-term measurements. *Eur J Clin Nutr* 56 Suppl 2, S53-62.
20. Beaton GH (1994) Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology. *Am J Clin Nutr* 59, 253S-261S.
21. Hunter D (1998) Biochemical indicators of dietary intake. In *Nutritional epidemiology*, 2nd ed., pp. 174-243 [W Willett, editor]. New York etc.: Oxford University Press.
22. Al-Delaimy WK, Ferrari P, Slimani N *et al.* (2005) Plasma carotenoids as biomarkers of intake of fruits and vegetables: individual-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Eur J Clin Nutr* 59, 1387-1396.
23. Al-Delaimy WK, Slimani N, Ferrari P *et al.* (2005) Plasma carotenoids as biomarkers of intake of fruits and vegetables: ecological-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Eur J Clin Nutr* 59, 1397-1408.
24. Welch AA, Lund E, Amiano P *et al.* (2002) Variability of fish consumption within the 10 European countries participating in the EPIC study. *Public Health Nutr* 5, 1273-1285.
25. Bowen L (2007) Applied Multilevel Analysis: A Practical Guide. Jos WR Twisk. *Int J Epidemiol* 36, 934-935.
26. Kroes R, Müller D, Lambe J *et al.* (2002) Assessment of intake from the diet. *Food Chem Toxicol* 40, 327-385.
27. EFSA (2005) Opinion of the Scientific Committee on a request from EFSA related to Exposure Assessments. European Food Safety Authority. EFSA-Q-2003-107. *The EFSA Journal* 249, 1-26.

28. Raffo A (2010) The new "FACET" (Flavourings, Additives and food Contact material Exposure Task) project: Towards a more reliable assessment of dietary exposure to flavourings. *MoniQA International Conference*, available at http://www.ucd.ie/facet/publications/facetpubtitle,24900,en.html

29. EFSA (2010) Minutes of the 3rd meeting of expert group for chemical occurence data. available at http://www.efsa.europa.eu/en/datexegs/documents/datexchemicaloccurence.pdf on 18 Jan 2011

30. Pryer JA, Vrijheid M, Nichols R *et al.* (1997) Who are the 'low energy reporters' in the dietary and nutritional survey of British adults? *Int J Epidemiol* 26, 146-154.

31. Rumpler WV, Kramer M, Rhodes DG *et al.* (2008) Identifying sources of reporting error using measured food intake. *Eur J Clin Nutr* 62, 544-552.

32. Subar AF, Kipnis V, Troiano RP *et al.* (2003) Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol* 158, 1-13.

33. Welch AA, Fransen H, Jenab M *et al.* (2009) Variation in intakes of calcium, phosphorus, magnesium, iron and potassium in 10 countries in the EPIC study. *Eur J Clin Nutr* 63, S101-S121.

34. Agudo A, Slimani N, Ocké M *et al.* (2002) Consumption of vegetables, fruit and other plant foods in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts from 10 European countries. *Public Health Nutr* 5, 1179-1196.

35. Biro G, Hulshof KF, Ovesen L *et al.* (2002) Selection of methodology to assess food intake. *Eur J Clin Nutr* 56 Suppl 2, S25-32.

36. Trolle E, Amiano P, Ege M *et al.* (2011) Evaluation of 2x24h recalls combined with a food recording booklet, against a 7 day food record method among school children. *Eur J Clin Nutr (in press)*.

37. Ocke MC, van Rossum CTM, Fransen HP *et al.* (2007) *Dutch National Food Consumption Survey-Young Children 2005/2006.* Bilthoven: RIVM.

38. Huybrechts I, Casagrande C, Nicolas G *et al.* (2011) Inventory of experiences from national/regional dietary monitoring surveys using EPIC-Soft *Eur J Clin Nutr (in press)*.

39. de Vries JH, de Groot LC & van Staveren WA (2009) Dietary assessment in elderly people: experiences gained from studies in the Netherlands. *Eur J Clin Nutr* 63 Suppl 1, S69-74.

40. Heerstrass DW, Ocke MC, Bueno-de-Mesquita HB *et al.* (1998) Underreporting of energy, protein and potassium intake in relation to body mass index. *Int J Epidemiol* 27, 186-193.

41. Lissner L, Troiano RP, Midthune D *et al.* (2007) OPEN about obesity: recovery biomarkers, dietary reporting errors and BMI. *Int J Obes (Lond)* 31, 956-961.

42. Tasevska N, Midthune D, Potischman N *et al.* (2011) Use of the predictive sugars biomarker to evaluate self-reported total sugars intake in the Observing Protein and Energy Nutrition (OPEN) study. *Cancer Epidemiol Biomarkers Prev* 20, 490-500.

43. Jenab M, Slimani N, Bictash M *et al.* (2009) Biomarkers in nutritional epidemiology: applications, needs and new horizons. *Hum Genet* 125, 507-525.

44. Souverein OW, de boer WJ, Geelen A *et al.* (2011) Quantifying uncertainty intake due to portion size estimation in 24-hour recalls for dietary surveys. *J Nutr (in press)*.

45. Thompson FE, Subar AF, Loria CM *et al.* (2010) Need for technological innovation in dietary assessment. *J Am Diet Assoc* 110, 48-51.

46. Suchman L & Jordan B (1990) Interactional troubles in face-to-face survey interviews. *J Am Statistical Assoc* 85, 232-241.

47. Menon G (1993) The effects of accessibility of Information in Memory on Judgments of Behavioral Frequencies. *J Consum Res* 20, 431-440.

48. Huybrechts I, Geelen A, de Vries JH *et al.* (2011) A respondents' evaluation of the 24-h dietary recall method (EPIC-Soft) in the EFCOVAL project. *submitted*.

49. Harttig U, Haubrock J & Boeing H (2011) The MSM program; The web-based statistics package for estimating usually dietary intake using the Multiple Source Method. *Eur J Clin Nutr (in press)*.

50. Andersen LF, Lioret S, Brants H *et al.* (2011) Recommendations for a trans European dietary assessment method among children and adolescents. *Eur J Clin Nutr (in press)*.

51. Trolle E, Amiano P, Ege M *et al.* (2011) Feasibility of 2x24h recall combined with a food recording booklet, using EPIC-Soft, among school children. *Eur J Clin Nutr (in press)*.

52. Slimani N, Casagrande C, Deharveng G *et al.* (2011) The standardized computerized 24-hour dietary recall method EPIC-Soft adapted for pan-European dietary monitoring. *Eur J Clin Nutr (in press)*.

53. de Boer E (2011) Rationale and methods of the EFCOVAL project. *Eur J Clin Nutr (in press)*.

54. Trolle E, Amiano P, Ege M *et al.* (2011) Feasibility of repeated 24h recall combined with a food recording booklet, using Epic-soft, among preschoolers. *Eur J Clin Nutr (in press)*.

55. Haubrock J, Nothlings U, Volatier JL *et al.* (2011) Estimating usual food intake distributions by using the multiple source method in the EPIC-Potsdam calibration study. *J Nutr* 141, 914-920.

56. Le Donne C, Piccinelli R, Sette S *et al.* (2011) Overview of existing European food consumption databases: critical aspects in relation to their use for the assessment of dietary exposure to additives, flavourings and residues of food contact materials. *Int J Food Sci Nutr* 62, 121-132.

Samenvatting and Resumo

(summaries in Dutch and Portuguese)

Samenvatting

Momenteel worden verschillende methodieken gebruikt voor het verzamelen van voedingsgegevens van volwassenen bij de diverse voedselconsumptiepeilingen in Europa. Niet alleen de methode van dataverzameling is verschillend, maar ook de mate van detail en de kwaliteit van de verzamelde data verschilt in deze Europese peilingen tussen de landen. Daarom is het niet mogelijk om de verschillen in inname van voedingsmiddelen en nutriënten tussen de Europese populaties nauwkeurig vast te stellen. In deze context hebben experts in het 'European Food Consumption Survey Methods' (EFCOSUM) consortium voorgesteld om dezelfde methode te gebruiken voor het gestandaardiseerd monitoren van de voedselconsumptie in Europese landen, namelijk een tweedaagse niet-opeenvolgende 24-uur's navraagmethode (recall). Naar aanleiding van de aanbevelingen van EFCOSUM werkte het 'European Food Consumption Validation' (EFCOVAL) consortium verder aan de ontwikkeling en validatie van de 24-uur's navraagmethode met EPIC-Soft software voor het bepalen van de inname van voedingsmiddelen, nutriënten en potentieel gevaarlijke stoffen in toekomstige Europese voedselconsumptiepeilingen. Het onderzoek beschreven in dit proefschrift is uitgevoerd binnen het EFCOVAL consortium en had als doel het evalueren van de kwaliteit van de data verzameld met de gestandaardiseerde 24-uur's navraagmethode met EPIC-Soft voor het vergelijken van de inname van volwassenen tussen landen voor toekomstige Europese voedselconsumptiepeilingen.

We begonnen onze studie met het vergelijken van de validiteit tussen vijf Europese centra voor het schatten van de gebruikelijke eiwit en kalium inname, op basis van twee niet-opeenvolgende, gestandaardiseerde 24-uur recalls met EPIC-Soft (**hoofdstuk 2**). Hiervoor maakten we gebruik van twee zogenaamde 'recovery' biomerkers, namelijk stikstof en kalium in de urine, als referentiemethoden voor eiwit en kalium inname. We hebben gegevens verzameld van 600 volwassenen tussen de 45 en 65 jaar oud in vijf Europese landen: België, Tjechië, Frankrijk, Nederland en Noorwegen. Van elke deelnemer zijn twee 24-uur recalls en twee 24-uur's urines verzameld. De resultaten lieten zien dat mannen en vrouwen hun eiwit inname gemiddeld onderrapporteerden met 8%. Voor kalium inname gold dat mannen met 7% onderrapporteerden en vrouwen met 4%. De variabiliteit van de bias in eiwit en kalium inname liep op tot 7% in de verschillende Europese centra. Verder zagen we dat de prevalentie van deelnemers met een adequate eiwit en kalium inname bepaald met de geobserveerde gegevens vrij goed overeen kwam (<10% verschil) met de prevalentie volgens de biomerker gegevens. Tenslotte zagen we matige correlaties

voor de rangorde van individuen tussen hun eiwit en kalium inname en de respectievelijke biomerkers. Deze waren vergelijkbaar tussen de centra. De resultaten uit de validatiestudie laten zien dat de 24-uur's navraagmethode een voldoende valide methode is om eiwit en kalium inname te schatten en te vergelijken tussen Europese populaties, ondanks dat variabiliteit in bias verwacht kan worden.

Om verder inzicht te verkrijgen in de nauwkeurigheidsdeterminanten van de 24-uur's navraagmethode, hebben we een lineaire multilevel analyse uitgevoerd waarbij de gegevens van de EFCOVAL studie gecombineerd zijn met gegevens van negen andere Europese centra uit de 'European Prospective Investigation into Cancer and Nutrition' (EPIC) calibratiestudie (**hoofdstuk 3**). In deze studie, zijn eiwit en kalium inname verzameld met een enkele 24-uurs recall geëvalueerd met behulp van de 'recovery' biomerker in de urine. Bij de uiteindelijke modellen zagen we geen verschil in de gemiddelde bias in eiwit inname voor mannen en vrouwen en in kalium inname voor vrouwen tussen de centra. Een kleine variatie in de bias werd gezien in de kalium inname van mannen (coëfficiënt van variatie = 9.5%). Verklarende variabelen op individueel niveau, namelijk BMI, dag van de week en de manier van afname van het interview, voorspelden de bias in eiwit en kalium inname en verklaarden de tussen-centra variatie. Variabelen op het niveau van het centrum (bijvoorbeeld 'human development index') hadden geen invloed op de resultaten. In beide analyses die gepresenteerd worden in hoofdstukken 2 en 3 was BMI de factor die het grootste deel van de variatie in bias in de geobserveerde eiwit en kalium inname tussen Europese populaties verklaarde.

Aanvullend hebben we bewijs verzameld over de prestatie van de methode voor het vaststellen van de inname van voedingsmiddelen. Hiertoe evalueerden we de rangorde van de individuen volgens hun gebruikelijk vis en fruit & groente inname met concentratie biomerkers in het bloed, respectievelijk vetzuren in fosfolipiden en carotenoïden (**hoofdstuk 4**). We gebruikten gegevens van de EFCOVAL validatiestudie, inclusief de gegevens van biomerkers gemeten in het bloed en van de frequentie van gebruik van voedingsmiddelen verzameld om mensen te identificeren die de geselecteerde voedingsmiddelengroepen nooit consumeren. We zagen zwakke tot matige correlaties tussen vis en fruit & groente inname en hun respectievelijk biomerkers. De samengevoegde ('pooled') correlatie tussen gebruikelijk visinname en EPA plus DHA in fosfolipiden was 0.19 voor mannen en 0.31 voor vrouwen. Dit leek niet verschillend te zijn tussen de centra (p>0.50). Voor gebruikelijke fruit & groente inname was de samengevoegde correlatie met de som van carotenoïden 0.31 voor mannen en 0.40 voor vrouwen. Wederom waren er geen aanwijzingen voor

verschillen tussen de centra (p>0.10). Deze resultaten suggereren dat twee gestandaardiseerde 24-uur recalls gebruikmakend van EPIC-Soft in combinatie met een vragenlijst om de frequentie van inname van voedingsmiddelen te bepalen, een geschikte methode is voor het rangschikken van personen volgens hun gebruikelijke vis en fruit & groente inname in de vijf Europese centra.

Vervolgens hebben we de invloed geëvalueerd van verschillende wijzen van afnemen van het interview (persoonlijk vs. telefonisch interview), de dag van de recall (1e vs. 2e), dagen van de week (weekdag vs. weekend) en dag van de interview (1 of 2 dagen later) op de bias in eiwit en kalium inname binnen de EFCOVAL studie (**hoofdstuk 5**). We zagen dat de bias in eiwit en kalium inname vergelijkbaar was voor de verschillende wijzen van afnemen van het interview, de dag van de recall, dagen van de week en tijd van het interview in sommige, maar niet alle centra. De resultaten suggereren dat de 24-uur's navraagmethode met een telefonisch interview betere resultaten opleverde dan een persoonlijk interview in het onderzoekscentrum in sommige Europese centra en dat de tweede recall minder goede resultaten opleverde dan de eerste. Bovendien bleek dat de dagen van de week gelijk vertegenwoordigd moeten zijn in voedselconsumptiepeilingen of dat, indien dat niet het geval is, dat moet worden meegenomen in de analyse van de gegevens.

In **hoofdstuk 6** presenteren we een verkennend onderzoek met als doel om de bruikbaarheid te bepalen van de voedingsgegevens verzameld met behulp van de 24-uur's navraagmethode voor het schatten van de chemische blootstelling gebruikmakend van een stapsgewijze aanpak. Smaakstoffen waren de chemische categorie die gebruikt is als een voorbeeld en vier smaakstoffen zijn geïncludeerd in dit onderzoek: raspberry ketone, cafeïne, glycyrrhizinezuur en cumarine. Allereerst zijn drie screeningsmethoden gebruikt en als de blootstelling de veiligheidslimiet overschreed voor één van deze drie methoden werd een meer gedetailleerde bepaling nodig geacht. Dit was het geval voor raspberry ketone, glycyrrhizinezuur en cafeïne, maar niet voor cumarine. Vervolgens zijn de individuele voedingsgegevens van Nederland, verzameld als onderdeel van de EFCOVAL studie, gebruikt om de blootstelling aan de smaakstoffen te schatten. De resultaten lieten zien dat de verzameling van gedetailleerde voedingsgegevens op individueel niveau nuttig was bij het bepalen van de blootstelling aan deze smaakstoffen. Tevens zagen we dat aanpassingen aan de databases die gebruikt worden in EPIC-Soft de mogelijkheid boden om meer gedetailleerde informatie te verzamelen over raspberry ketone.

In **hoofdstuk 7** worden de belangrijkste bevindingen van dit proefschrift in een breder perspectief bediscussieerd. De onzekerheden in onze evaluaties over het

gebruik van de gegevens van de samenstelling van voedingsmiddelen, portiegroottes, omzettingsfactoren, modellering van gebruikelijke inname, biomerkers, multilevel analyse en over de blootstellingsanalyse worden behandeld. Verder wordt de mogelijkheid beschreven voor generalisatie van de resultaten naar andere nutriënten, voedingsmiddelengroepen, chemische stoffen, landen, voedingspatronen en populatiegroepen. Een aantal mogelijkheden voor toekomstig onderzoek wordt aangestipt, bijvoorbeeld de noodzaak voor een voortdurende kwaliteitscontrole van de 24-uur's navraagmethode.

Tenslotte wordt geconcludeerd dat twee niet-opeenvolgende 24-uur recalls met EPIC-Soft voldoende valide en bruikbare gegevens opleveren voor het vergelijken van de voedselconsumptie tussen Europese populaties in toekomstige pan-Europese voedselconsumptiepeilingen.

Resumo

Atualmente, inquéritos nutricionais na Europa utilizam metodologias diferentes para coletar os dados dietéticos de populações adultas. Não só os métodos dietéticos são diferentes, mas também o detalhamento e a qualidade dos dados coletados em inquéritos europeus diferem muito entre os países. Portanto, no momento não é possível avaliar as diferenças no consumo alimentar em populações adultas na Europa. Neste contexto, especialistas participantes do consórcio 'European Food Consumption Survey Methods' (EFCOSUM) propuseram utilizar o mesmo método para o monitoramento padronizado da ingestão dietética em países europeus, que é o recordatório de 24-h aplicado duas vezes de maneira não-consecutiva. Seguindo as recomendações do EFCOSUM, o consórcio 'European Food Consumption Validation' (EFCOVAL) trabalhou no sentido de desenvolver e validar a metodologia do recordatório de 24-h, usando o software EPIC-Soft, para avaliação de alimentos, nutrientes e substâncias potencialmente perigosas em futuros inquéritos nutricionais na Europa. O presente trabalho foi realizado no âmbito do consórcio EFCOVAL e teve como objetivo avaliar a qualidade dos dados coletados com o método recordatório de 24-h, usando EPIC-Soft, para comparações de consumo alimentar de adultos entre países no futuro inquérito pan-europeu em nutrição.

Começamos nosso estudo comparando a validade da ingestão habitual de proteína e potássio estimada a partir de dois recordatórios de 24-h não-consecutivos, usando EPIC-Soft, entre cinco centros europeus (**capítulo 2**). Foram utilizados dois biomarcadores ('recoveries'), nitrogênio e potássio na urina, como os métodos de referência para a ingestão de proteína e potássio. Dados de 600 adultos, entre 45 e 65 anos, foram coletados em cinco países europeus: Bélgica, República Checa, França, Holanda e Noruega. De cada participante, dois recordatórios de 24-h e duas urinas de 24-h foram coletadas. Em média, observamos que homens e mulheres subestimaram a ingestão de proteínas estimadas pelos dois recordatórios de 24-h em 8%. Quanto a ingestão de potássio na dieta, os homens subestimaram em 7% e as mulheres em 4%. A variabilidade de viés ('bias') na ingestão de proteína e potássio estimada foi de até 7% entre os centros europeus. Além disso, observamos que a prevalência de indivíduos com quantidade adequada de ingestão de proteína e potássio concordaram razoavelmente bem com a prevalência obtida com os biomarcadores (diferença <10%). Finalmente, correlações moderadas foram observadas na classificação dos indivíduos de acordo com suas ingestões de proteína e potássio com seus respectivos marcadores, parecendo ser comparáveis entre os cinco centros. Esses resultados sugerem que o recordatório de 24-h é um método suficientemente válido para avaliar

e comparar o consumo de proteínas e potássio entre as populações européias, apesar de que variabilidade em viés pode ser esperada.

Para explorar os determinantes da validade do método recordatório de 24-h foi realizada uma análise multinível linear, combinando os dados do estudo EFCOVAL com dados de outros nove centros europeus que participam no estudo de calibração 'European Prospective Investigation into Cancer and Nutrition' (EPIC) (**capítulo 3**). Neste estudo, as ingestões de proteína e potássio coletados com um recordatório de 24-h foram avaliados por biomarcadores na urina. No modelo final, observou-se que a média do viés da ingestão de proteína em ambos os sexos e da ingestão de potássio em mulheres não variou entre os centros e, em certa medida variou na ingestão de potássio em homens (coeficiente de variação=9,5%). Enquanto as variáveis explicativas em nível individual, ou seja, IMC, dias da semana e modo de administração, previram o viés da ingestão de proteínas e potássio e explicaram a variação entre os centros, as variáveis relativas ao centro (por exemplo, o índice de desenvolvimento humano) não influenciaram os resultados. Em ambas as análises apresentadas nos capítulos 2 e 3, o IMC foi o fator que mais explicou a variação em viés observada na ingestão de proteína e potássio em populações européias.

Em seguida, mais evidencias sobre o desempenho do método recordatório de 24-h foram obtidas por meio da avaliação da classificação dos indivíduos de acordo com o seu consumo de peixes, frutas e vegetais utilizando biomarcadores de concentração, respectivamente, ácidos graxos em fosfolipídios e carotenóides (**capítulo 4**). Dados do EFCOVAL foram utilizados, incluindo os dados de biomarcadores obtidos com uma coleta de sangue e os dados sobre propensão alimentar, que foram coletados com um questionário para identificar os não-consumidores dos grupos de alimentos estudados. Em geral, observou-se correlações fracas a moderadas entre os peixes e frutas e vegetais e seus respectivos biomarcadores. A ('pooled') correlação entre a ingestão habitual de peixes e EPA+DHA em fosfolipídios foi de 0,19 nos homens e 0,31 nas mulheres, parecendo não haver diferenças entre os centros (p>0,50). Para a ingestão habitual de frutas e verduras, a ('pooled') correlação com a soma dos carotenóides foi de 0,31 nos homens e 0,40 nas mulheres, também parecendo não haver diferenças entre os centros (p>0,10). Esses resultados sugerem que os dois recordatórios de 24-h não-consecutivos, usando EPIC-Soft, em combinação com o questionário de propensão alimentar parece ser apropriado para classificar indivíduos de acordo com o seu consumo habitual de peixes, frutas e vegetais entre os cinco centros europeus de uma maneira comparável.

Depois disso, avaliamos o impacto de diferentes modos de administração (presencial vs entrevista por telefone), dia do recordatório (1º vs 2º), dias da semana (segunda à

sexta vs final de semana) e o dia da entrevista (1 ou 2 dias depois) sobre o viés da ingestão de proteína e potássio avaliado no estudo EFCOVAL (**capítulo 5**). Em geral, observamos que o viés no consumo de proteínas e potássio foi comparável entre os modos de administração, os dias do recordatório, os dias da semana e os dias da entrevista, em alguns, mas não todos, os centros. Os resultados sugerem que recordatórios de 24-h coletados por meio de entrevistas telefônicas proporcionaram uma avaliação mais precisa do que por meio de entrevistas presenciais no centro de pesquisa em alguns centros europeus, e que as avaliações do segundo recordatório de 24-h são menos precisas que o primeiro recordatório Além disso, os resultados sugeriram que os dias da semana devam ser representados igualmente em inquéritos alimentares ou devidamente ajustado durante a análise dos dados.

No **capítulo 6**, apresentamos um estudo exploratório, que objetivou investigar a utilidade dos dados dietéticos coletados com o recordatório de 24-h para estimar a exposição química, usando uma abordagem 'passo-a-passo'. As substâncias aromatizantes foram a categoria de produto químico usado como um exemplo e quatro aromas foram incluídos na avaliação: cetona da framboesa, cafeína, ácido glicirrízico e cumarina. Primeiramente, três métodos de triagem foram aplicados. Quando a exposição por meio da dieta excedeu os limites de segurança para um desses três métodos, julgou-se necessário a utilização de uma avaliação mais detalhada com menos dados agregados. Esse foi o caso de cetona da framboesa, ácido glicirrízico e cafeína, mas não da cumarina. Em seguida, dados de consumo alimentar coletados de indivíduos na Holanda, que fizeram parte do estudo EFOVAL, foram utilizados para estimar a exposição alimentar aos aromas. Os resultados sugeriram que a coleta detalhada de dados sobre o consumo alimentar a nível individual foi útil na avaliação da exposição alimentar dos aromas estudados. Além disso, observou-se que a possibilidade de adaptação dos bancos de dados utilizados no EPIC-Soft forneceram informações mais detalhadas para a exposição alimentar a cetona da framboesa, em relação às bases de dados não-modificados.

No **capítulo 7** as principais conclusões desta tese são discutidas com uma perspectiva mais ampla. As incertezas em nossas avaliações sobre o uso de dados de composição de alimentos, tamanho de porções, fatores de conversão, modelagem estatística do consumo alimentar habitual, biomarcadores, avaliações de multiníveis, e sobre as avaliações da exposição química na dieta são abordadas. Além disso, a generalização dos resultados apresentados a outros nutrientes, grupos de alimentos, aromas e substâncias químicas; países e padrões alimentares; grupos de população são descritas. Algumas indicações para futuras pesquisas na área também são mencionadas, como a necessidade de controle contínuo da qualidade do método recordatório de 24-h.

Por fim, conclui-se que o recordatório de 24-h aplicados duas vezes de maneira não-consecutiva, usando EPIC-Soft, fornece dados suficientemente válidos e adequados para comparar o consumo alimentar entre as populações européias em futuros inquéritos pan-europeus em Nutrição.

Affiliation of the co-authors

Lene F. Andersen
Department of Nutrition
Institute of Basic Medical Sciences
University of Oslo
Oslo, Norway

Joline WJ Beulens
Julius Center for Health Sciences and
Primary Care, University Medical Center
Utrecht
Utrecht, The Netherlands.

Heiner Boeing
Department of Epidemiology, German
Institute of Human Nutrition,
Potsdam-Rehbrücke, Germany

Evelien de Boer
National Institute for Public Health and
the Environment (RIVM)
Bilthoven, The Netherlands

Hendriek Boshuizen
Division of Human Nutrition
Wageningen University
Wageningen, The Netherlands

Marie-Christine Boutron-Ruault
INSERM
Centre for Research in Epidemiology and
Population Health
Villejuif, France

H. Bas de Bueno-de-Mesquita
National Institute for Public Health and
the Environment (RIVM)
Bilthoven, The Netherlands

Francesca L Crowe
Cancer Epidemiology Unit
Nuffield Department of Clinical Medicine
University of Oxford
Oxford, United Kingdom

Marcela Dofkova
National Institute of Public Health
Department for Food Safety and Nutrition
Brno, Czech Republic

Cinzia Le Donne
National Research Institute for Food and
Nutrition (INRAN)
Rome, Italy

Heinz Freisling
Dietary exposure assessment group
International Agency for Research on
Cancer (IARC)
Lyon, France

Anouk Geelen
Division of Human Nutrition
Wageningen University
Wageningen, The Netherlands

Paul J. M. Hulshof
Division of Human Nutrition
Wageningen University
Wageningen, The Netherlands

Inge Huybrechts
Department of Public Health
Faculty of Medicine and Health Sciences
Ghent University
Ghent, Belgium

Rudolf Kaaks
Division of Cancer Epidemiology,
German Cancer Research Center
Heidelberg, Germany

Willem De Keyzer
Department of Nutrition and Dietetics
University College Ghent
Ghent, Belgium

Vittorio Krogh
Nutritional Epidemiology Unit,
Fondazione IRCCS Istituto Nazionale
Tumori
Milan, Italy

Lionel Lafay
Food Safety Department
Food Intake-Nutritional Epidemiology
Unit, French Agency for Food,
Environmental and Occupational Health
Safety (ANSES), Maisons-Alfort, France

Catherine Leclercq
National Research Institute for Food and
Nutrition (INRAN)
Rome, Italy

Inger T. L. Lillegaard
Department of Nutrition
Institute of Basic Medical Sciences
University of Oslo
Oslo, Norway

Maria Santucci de Magistris
Department of Public Clinical and
Experimental Medicine, Federico II
Naples, Italy

Irene Margaritis
Food Safety Department
Food Intake-Nutritional Epidemiology
Unit - French Agency for Food,
Environmental and Occupational Health
Safety (ANSES)
Maisons-Alfort, France

Alison McTaggart
Department of Public Health and Primary
Care, University of Cambridge
Cambridge, United Kingdom.

Androniki Naska
WHO Collaborating Center for Food and
Nutrition Policies, Department of Hygiene
Epidemiology and Medical Statistics
University of Athens Medical School
Athens, Greece

Marga C. Ocke
National Institute for Public Health and
the Environment (RIVM)
Bilthoven, The Netherlands

Antonio Raffo
National Research Institute for Food and
Nutrition (INRAN)
Rome, Italy

Irena Rehurkova
National Institute of Public Health
Department for Food Safety and Nutrition
Brno, Czech Republic

Fulvio Ricceri
Human Genetics Foundation (HUGEF)
Turin, Italy

Jiri Ruprich
National Institute of Public Health
Department for Food Safety and Nutrition
Brno, Czech Republic

Stefania Sette
National Research Institute for Food and
Nutrition (INRAN)
Rome, Italy

Els Siebelink
Division of Human Nutrition
Wageningen University
Wageningen, The Netherlands
Nadia Slimani
Dietary exposure assessment group
International Agency for Research on
Cancer (IARC)
Lyon, France

Anne-Sophie Rousseau
INSERM U907 (Laboratoire Dysfonctions
Métaboliques et Diabetes, Mécanismes et
Approches Thérapeutiques)
Université de Nice Sophia –Antipolis,
Nice, France

Olga W. Souverein
Division of Human Nutrition
Wageningen University
Wageningen, The Netherlands

Rosario Tumino
Cancer Registry and Histopathology Unit
"Civile M.P.Arezzo" Hospital
Ragusa, Italy

Pieter van't Veer
Division of Human Nutrition
Wageningen University
Wageningen, The Netherlands

Jeanne H. M. de Vries
Division of Human Nutrition
Wageningen University
Wageningen, The Netherlands

Acknowledgement

Four years of my life are partly translated in this thesis. Fortunately, I was not alone. There were many people that helped me to go through this intense and gratifying period...

At the beginning, there was **Ronaldo Oliveira**, my dear friend, who forwarded me the vacancy announcement of my PhD position. He also wanted to apply for the position, but he was kind enough to share the opportunity with me. Ronaldo, thank you for the friendship.

Right after, there were my supervisors, **Pieter van't Veer, Jeanne de Vries** and **Anouk Geelen**, who gave me the opportunity to have my PhD degree. I will always be grateful to you for believing that I could do the job. From you **Pieter**, I will remember the inspiring meetings that we had every month. I must confess that I was always very anxious to meet my promotor. But also, I always left the meetings less anxious about the work I had to accomplish and with the feeling that you knew that the papers and thesis would turn out fine. From you **Jeanne**, I will always remember the care (almost like a mother) towards my work. Your annual poems around Sinterklaas day were so thoughtful. From you **Anouk**, I will always remember the prompt help. You were indeed a daily supervisor that was there to support and encourage my development. Hartelijk bedankt!

Then, there were a number of European scientists (outside Wageningen), who were part of the EFCOVAL project: **Antoinette Kaic-Rak, Arnold Dekkers, Caroline Wilson, Catherine Leclercq, Cinzia Le Donne, Corinne Casagrande, Geneviève Deharveng, Elaine Purse, Ellen Trolle, Elly Buurma, Evelien de Boer, Heiner Boeing, Heinz Freisling, Henny Brants, Inge Huybrechts, Irena Rehurkova, Jan van der Laan, Jean-Luc Volatier, Jennifer Haubrock, Jiri Ruprich, Lene Andersen, Lionel Lafay, Marcela Dofkova, Marga Ocke, Maryse Niekerk, Max Feinberg, Mia Bellemans, Mieke de Maeyer, Nadia Slimani, Pilar Amiano, Sandrine Lioret, Therese Lillegaard, Ulrich Harttig, Tue Christensen, Wanda Vos and Willem de Keizer** (and I am probably forgetting someone). I thank you all so much for welcoming a non-European into a European project. I enjoyed a lot the scientific and non-scientific talks we had during our EFCOVAL meetings. I also appreciate your patience with me, who was one of the only students in the group. I learned a lot from you!

There were also other members of EFCOVAL (from Wageningen) for which I am very grateful: **Els Siebelink, Olga Souverein** and **Paul Hulshof**. **Els**, thanks for

your support to my many questions during the collection and management of the Dutch data. It was a pleasure to collaborate with you during that period. I also appreciate your kind words during coffee breaks or occasional meetings. **Olga**, it is for a good reason that I choose you as one of my paranimphs: I don't think I would have made through my thesis without your support. You were there to cheer me up when things were not going okay; and you were also there to call my attention when I was losing track of my work. Plus, I enjoyed very much having you as roommate at the office. I will never forget the great laughs we had about propositions from my thesis that were never published or the silly things from daily life. **Paul**, thanks for the many opportunities to discuss the laboratory analyses of the EFCOVAL study.

During the second year of my PhD, I had the opportunity to work with the research group of **Catherine Leclercq** in Italy. It was one of the most beautiful periods of my PhD life. I thank you Catherine for the great teaching into the field of dietary exposure assessment to chemicals. **Cinzia Le Donne, Deborah Martone, Stefannia Sette** (who taught me how to use SAS), **Antonio Raffo, Rafaela Piccineli, Pasquale Buonocore, Marika Ferrari, Davide Arcella** (for your advices), I am very thankful for your support while I was in Rome and for teaching me on how to appreciate a nice cup of Italian coffee. Grazie mille!! Spero anche di vedervi in tempi brevi.

During the last year of my PhD, I worked with the research group of **Nadia Slimani** in France. Nadia, it was an honour to be part of your team and it was inspiring to cooperate with you. Needless to say that I learned a lot from you all: **Aurelie Moskal, Beatrice Vozar, Bertrand Hemon, Carine Biessy, Corinne Casagrande, Geneviève Deharveng, Graham Byrnes, Heinz Freisling, Jerome Vignat, Sabina Rinaldi, Veronique Chajes,** and **Thomas Cler. Nicole Suty** and **Eve Elakroud,** thanks for the all administrative work. Merci beaucoup! Je suis très hereuse, parce que je serais de retour à Lyon pour travailler de nouveau avec vous.

A number of people from the Division of Human Nutrition cannot be forgotten. **Lidwien van der Heyden**, you were the first one to welcome into the division. Thanks for allowing time to answer all my administrative questions - I recognize they were too many :). **Eric van Munster and Riekie Janssen**, thanks for processing all my financial requests, especially during the PhD tour 2009. Special thanks to the secretaries of our division, **Cornelia van Bree-Evers, Gabriëlle van den Broek, Gea Brussen, Karen Zweers, Lyda Jong, Lous Duym, Marie Jansen,** who helped during these years. My dear colleagues from the lab, **Betty van der Struijs, Ly Nhien, Pieter Versloot, Tineke van Roekel-Jansen**, thanks for all the work during

the lab analysis. **Dione Bouchaut and Jan Harryman**, thanks for your essential IT-related work. **Saskya Meybom**, thank you for your help and smiles during these years. **Hendriek Boshuizen**, thanks for arriving at the right time :) and helping me out with finalizing the statistical analyses. **Peter Hollman,** thanks for advices related to EuroFir. **Monica Mars**, thanks for the help in the analysis of my last paper. **Anneleen Kuijsten and Rosalie Dhonukshe-Rutten,** thanks for answering my random questions and having the nice talks at the corridor.

Roelfina Mihalj-IJken from the human resources division, I do not know what I would have done without your assistance to arrange all the visa issues. Thanks for everything.

Dear **Fre Pepping** from VLAG, thanks for the nice talks along the years, including the performance of Felipe Massa at the Formula 1. ☺

Of course, there were many PhD buddies who shared their moments with me in the past four years. **Gerda Pot**, I undoubtedly will start with you. You were my PhD-mate, 'responsible' to help me out around Wageningen during my first days in the Netherlands. I am so grateful to have had you around these years and became your friend. Even living in Cambridge, you were always present. The memories of our weekend afternoons with cappuccino, your visit to Rome, your presence during cocktail parties, or walking around Wageningen...will be always with me! **Marja Molag** and **Akke Botma**, it was a pleasure to have you working 'next door'. **Marja**, I am very honoured for your invitation to be your paranimph and to your friendship even after you left the division. Plus, it was a lot of fun 'to work' with you. **Akke**, I appreciate a lot your kindness towards me and advices during all these years. You are a talented nutritionist and no doubt you will go far with your career. My dear sandwich-PhD friends, **Ana Carla** and **Maria Elisa**, the time in the Netherlands was extremely enjoyable when you were around. I thank you for allow me to be your friend and I hope to see you soon! Many other names are to be mentioned: **Anand**, **Anne, Anthonie, Catherine, Cecile, Cora, Coraline, Dieuwertje, Du, Elise, Elske, Esmee, Eveline, Gertrude, Janneke, Jannete, Kevin, Linda, Linde, Martinette, Michael, Michiel, Michel, Mirre, Nadia, Nicole, Nidhi, Noortje, Ohid, Ondine, Pascalle, Pauline, Pleunie, Por, Rani, Renate B., Renate W., Rianne, Rina, Romana, Sandra, Sanne, Simone (buddy), Siyazi, Sophie, Truus, Vera, Victoire, Wieke, Yara**. My PhD years were much more pleasing because of the time we spent during the PhD tour, Newtrition committee, PhD meetings, WE-day, NWO days, supervision of exams or at the canteen and corridors. Especially, the PhD tour

committee 2009, **Cecile, Rianne, Michel** and **Sanne**, it was a great journey to organise the tour to the Nordic countries with you.

Then, there was the Brazilian community in Wageningen! Muitos amigos permitiram que eu estivesse um pouco mais próxima do Brasil por meio das suas amizades em Wageningen. Especialmente**, Anabele, Carol Mosca, Carol Silva, Cinara, Larissa,** e agregados (**Wiebe, Denis, Jânio**). Obrigada pela amizade, por abrirem seus lares para mim e para o Omar nestes quatro anos! Foi bem mais especial estar na Holanda, pois tive a companhia de vocês. Outros tantos brasileiros tornaram a minha jornada em Wageningen mais agradável. Não vai ser possível nomear todos, mas agradeço **Carla, Davi, Carlos, Felipe, Flavinha, Glaciela, Isabela, Luciana, Manoel, Marina, Murilo, Priscilla, Rafael, Rita, Odair, Saulito e Vanja** por estarem presentes na minha vida aqui. E claro que vou falar mais sobre você Carol. **Carol Mosca**, there was also for a good reason that I chose you as my other paranimph. First, someone needed to represent the Brazilians ☺. Second and most important, you have become such a good friend supporting my work until the last minute before submitting my thesis. I know our friendship is for a lifetime. Having you and **Olga** sitting next to me at the big day means a lot to me.

Thanks to good friends that shared their lives with me in Wageningen: **Arturo, Edward, Iemke, Marjolein, Pablo** and **Victor Hugo**.

I must also thank great friends that I met while in Italy. **Silvia Filosa** and your loving family, thanks for welcoming Omar and I at your home.

A special thank you goes to **Fred del Percio** and **Warren Burdock** from Wee Burn Country Club, Darien-CT-US, who once did me a great favor by allowing me to learn English. Those months of work were essential for me to arrive here. Thank you!

Then, my family. **Pai** e **mãe**, impossível colocar no papel a gratidão que sinto por vocês. De qualquer forma, obrigada pelo apoio, pelas visitas e por entender que a sua filha queria muito obter o doutorado um tanto distante da sua terra natal. Obrigada também aos meus irmãos, **Márcio, Patrick e Lucas**; cunhadas **Eliane, Adriana e Taise**; sobrinhos **Beatriz, Márcio, Yasmin e Yan**; por estarem presentes na minha vida de doutoranda, ainda que a distância. Obrigada a minha tia amada, **Neves**, que com muitas visitas alegrou minha vida aqui fora. Também, as tias **Lurdinha e Cidinha** e as primas **Ana, Camila, Mariana e Marina** que constantemente me mandaram mensagens de incentivo. E a algumas amigas-irmãs que sempre me

apoiaram: **Ingrid, Josi, Michele, Lina, Luciane, Vanessa, Viviana**. Gracias también por el apoyo de mi familia mexicana, especialmente mi suegra **Oralia**. No hay espacio para mencionar los nombres de todos mis nueve cuñados y agregados más los tantos >25 sobrinos ☺, pero los quiero mucho.

Finally, you my love. **Omar**, thank you por hacer parte de mi vita! ☺ I have no words to express how grateful I am for you support during all these years. It was not easy, you know. But it would have been much harder if you were not here with me. Gracias!

About the author

Curriculum vitae

Sandra Patricia Crispim was born on February 12, 1979 in Itajaí, Santa Catarina, Brazil. After completing middle school at 'Colégio Valério Gomes' and secondary school at 'Colégio Cenecista Pedro Antônio Fayal', she obtained her bachelor degree in Nutrition at the University Vale do Itajaí in 2001. Within the bachelor program, she was involved in several research projects, mostly related to nutritional assessment of individuals and populations. In 2001, Sandra completed a 6-month program as nutritionist trainee at the Santa Catarina Hospital - Blumenau. In the year of 2004, she obtained her master's degree in Nutrition Science at the Federal University of Viçosa. Her MSc thesis was on "The relative validity of a food frequency questionnaire". In 2007, Sandra was appointed as a PhD student at the Division of Human Nutrition of Wageningen University, the Netherlands. Her PhD research was part of the European Food Consumption Validation (EFCOVAL) consortium. Sandra performed part of her PhD research while hosted at the Italian National Institute for Food and Nutrition Research (INRAN), Rome, Italy, in 2008 and at the International Agency for Research on Cancer (IARC), Lyon, France, in 2010. She was a member of the organizing committee of PhD study tour to Denmark, Finland and Sweden in 2009. In 2011, she was selected for the European Nutrition Leadership Programme (ENLP). Currently, Sandra continues her collaboration with researchers from the dietary exposure assessment group at IARC, where she will be appointed as Post-Doc after her PhD defence.

List of publications (full papers)

1. Crispim, SP; Geelen, A; de Vries, JHM; Freisling, H; Souverein, OW; Hulshof, PJM et al. Bias in protein and potassium intake collected with 24-h recalls (EPIC-Soft) is rather comparable across European populations. *Submitted*
2. Crispim, SP; Geelen, A; Siebelink, E; Huybrechts, I; Lillegaard, ITL; Margaritis, I et al. Design aspects of 24-h recall assessments may affect the estimates of protein and potassium intake in dietary surveys. *Submitted*
3. Crispim, SP; Geelen, A; Souverein, OW; Hulshof, PJM; Ruprich, J; Dofkova, M et al. Biomarker-based evaluation of two 24-h recalls (EPIC-Soft) for estimating and comparing fish, fruit and vegetable intakes across five European centres: results from the European Food Consumption Validation (EFCOVAL) study. *Accepted European Journal of Clinical Nutrition 2011.*
4. Huybrechts, I; Casagrande, C; Deharveng, G; Geelen, A; Crispim, SP, De Keyzer, W et al. Inventory of experiences from national/regional dietary monitoring surveys using EPIC-Soft. *Accepted European Journal of Clinical Nutrition 2011.*
5. Crispim, SP; Geelen, A; de Vries, JHM; Souverein, OW; Hulshof, PJM; Lafay, L et al. Two non-consecutive 24-h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres – Results from the European Food Consumption Validation (EFCOVAL) study. *British Journal of Nutrition, 2011; 105: 447-458.*
6. Crispim, SP; Geelen, A; Le Donne, C; de Vries, JHM; Sette, S; Raffo, A. et al. Dietary exposure to flavouring substances: from screening methods to detailed assessments using food consumption data collected with EPIC-Soft software. *Food Additives and Contaminants, 2010, 27(4): 433-446.*
7. Crispim, SP; Silva, MMS; Panato, E; Rosado, LEP; Rosado, GP; Ribeiro, RLC. Relative validity of a food frequency questionnaire for use in adults. *Journal of Nutrition, 2009, 22: 81-96 (in Portuguese)*
8. Crispim, SP; Ribeiro, RLC; Silva, MMS; Rosado, LEP; Rosado, GP. Educational influence in the relative validation of a food frequency questionnaire, Viçosa, Minas Gerais, Brazil. *European Journal of Clinical Nutrition, 2006, 60(11): 1311-1316.*
9. Grillo, LP; Crispim, SP; Siebert, AN; Andrade, ATW; Rossi, A; Campos, IC. Lipid profile and obesity in schoolchildren of low income. *Brazilian Journal of Epidemiology, Sao Paulo, 2005, 8: 75-81 (in Portuguese)*
10. Venturi, I; Sant'Anna, LC; Crispim, SP; Bramorski, A; Mello, RMV. Training for conservation and hygiene of food: a proposal of educative practice. *Higiene Alimentar, 2004, 8(125): 32-35, (in Portuguese)*
11. Crispim, SP; Grillo, LP; Felipe, MR; Santos, PF. Nutritional situation of a population in a city with low index of human development, Castro Alves - Bahia. *Nutrition Brazil, 2004, 3(5): 297-303 (in Portuguese)*
12. Crispim, SP; Lima, ES; Calil, J; Felipe, MR; Grillo, LP. Prevalence of anaemia in schoolchildren and adolescents belonging to schools of Bombinhas City - SC. *Nutrition Brazil, 2003, 2(4): 196-202 (in Portuguese)*
13. Crispim, SP; Franceschini, SC; Priore, SE; Fisberg, RM. Validation of dietary measurements–a review. *Nutrire, 2003, 2: 127-141 (in Portuguese)*
14. Crispim, SP; Ribeiro, RCL; Silva, MMS. Validation of food frequency questionnaires. *Nutrition Brazil, 2003, 2(5): 286-290. (in Portuguese)*

Overview of completed training activities

Name of the course	Organizer/ Institution	Year
a. Discipline specific activities		
8th Nutritional and lifestyle epidemiology, Wageningen, NL	VLAG	2007
8th Production and use of food composition data in nutrition, Wageningen, NL	WUR	2007
Workshop 'Food consumption data needs for the assessment of dietary exposure to flavourings in the EU', Rome, IT	INRAN	2007
Workshop 'Recommendations on a trans-European dietary assessment method among children', Copenhagen, DK	DTU	2007
Workshop 'Methodology of validation studies' with oral presentation, Ghent, BE	EFCOVAL	2008
Workshop 'Past, current and future use of the EPIC-Soft programme', Paris, FR	EFCOVAL	2008
2009 Annual meeting NWO nutrition with oral presentation, Deurne, NL	NWO	2009
7th International conference on diet and activity measurements (ICDAM) with oral presentation, Washington, US	US-NCI	2009
EFCOVAL conference with oral presentation, Utrecht, NL	EFCOVAL	2009
Wageningen nutritional sciences forum with poster presentation, Division Human Nutrition, Wageningen, NL	WUR	2009
2nd World congress of Public Health Nutrition with oral and poster presentation, Porto, PT	WPHNA	2010
b. General courses		
PhD competence assessment, Wageningen, NL	WGS	2007
VLAG PhD week, Bilthoven, NL	VLAG	2007
Academic writing, Wageningen, NL	CENTA	2008
Communications and media training, Prague, CZ	EUROFIR	2008
Regression analysis, Erasmus Summer Programme, Rotterdam, NL	NIHES	2008
Scientific writing, Wageningen, NL	CENTA	2009
Career perspectives, Wageningen, NL	WGS	2010
Grants writing workshop, Lyon, FR	WHO/IARC	2010
c. Optional courses and activities		
Preparation research proposal, Wageningen, NL	WUR	2007
PhD guest (4 months) at the National Institute for Food and Nutrition Research, Research group food safety-exposure analysis, Rome, IT	INRAN	2008
MSc course Exposure assessment in nutrition and health research, Wageningen, NL	WUR	2008
Organising and participating PhD tour to Nordic countries (Denmark, Finland and Sweden)	WUR	2009
PhD guest (6 months) at the International Agency for Research on Cancer, Dietary exposure assessment group, Lyon, FR	WHO/IARC	2010
EFCOVAL research meetings, twice per year, different locations in Europe	EFCOVAL	2007-2009
Research presentations Division Human Nutrition, Wageningen, NL	WUR	2007-2010
Methodology and epidemiology research meetings, Division Human Nutrition, Wageningen, NL	WUR	2008-2011
17th European Nutrition Leadership Programme. Luxembourg, LU	ENLP	2011