ORIGINAL COMMUNICATION

Plasma carotenoids as biomarkers of intake of fruits and vegetables: individual-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC)

WK Al-Delaimy[1,30], P Ferrari[1], N Slimani[1], V Pala[2], I Johansson[3,4], S Nilsson[4], I Mattisson[5], E Wirfalt[5], R Galasso[6], D Palli[7], P Vineis[8], R Tumino[9], M Dorronsoro[10], G Pera[11], MC Ocké[12], HB Bueno-de-Mesquita[12], K Overvad[13], MªD Chirlaque[14], A Trichopoulou[15], A Naska[15], A Tjønneland[16], A Olsen[16], E Lund[17], EHR Alsaker[17], A Barricarte[18], E Kesse[19], M-C Boutron-Ruault[19,20], F Clavel-Chapelon[19], TJ Key[21], E Spencer[21], S Bingham[22], AA Welch[23], M-J Sanchez-Perez[24], G Nagel[25], J Linseisen[25], JR Quirós[26], PHM Peeters[27], CH van Gils[27], H Boeing[28], AL van Kappel[1,31], J-P Steghens[29] and E Riboli[1]*

[1]International Agency for Research on Cancer (IARC-WHO), Lyon, France; [2]Epidemiology Unit, National Cancer Institute, Milan, Italy; [3]Department of Odontology and Cariology, Umeå University, Umeå, Sweden; [4]Department of Nutritional Research, Umeå University, Umeå, Sweden; [5]Department of Medicine, Surgery & Orthopaedics, Lund University, Malmö, Sweden; [6]Ospedale Oncologico Regionale, Rionero, Italy; [7]Molecular & Nutritional Epidemiology Unit, CSPO, Scientific Institute of Tuscany, Florence, Italy; [8]Cancer Epidemiology Unit, University of Turin and CPO-Piemonte, Turin, Italy; [9]Cancer Registry, Azienda Ospedaliera 'Civile—M.P.Arezzo', Ragusa, Italy; [10]Public Health Office of Guipuzkoa, Health Council of the Basque Country, San Sebastian, Spain; [11]Department of Epidemiology, Catalan Institute of Oncology (ICO), L'Hospitalet de Llobregat, Barcelona, Spain; [12]National Institute of Public Health and the Environment (RIVM), Bilthoven, Netherlands; [13]Department of Clinical Epidemiology, Aalborg Hospital and Aarhus University Hospital, and Department of Epidemiology and Social Medicine, University of Aarhus, Denmark; [14]Department of Epidemiology, Council for Health and Social Affairs of Murcia, Spain; [15]University of Athens Medical School, Athens, Greece; [16]Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark; [17]Institute of Community Medicine, University of Tromsø, Norway; [18]Institute of Public Health of Navarra, Pamplona, Spain; [19]INSERM, U XR 521, Institut Gustave Roussy, Villejuif, France; [20]INSERM, U 557, ISTNA, CNAM, Paris, France; [21]Cancer Research UK, Radcliffe Infirmary, University of Oxford, UK; [22]MRC Dunn Human Nutrition Unit, Cambridge, UK; [23]Strangeways Research Laboratory, University of Cambridge, UK; [24]Andalusian School of Public Health, Granada, Spain; [25]Division of Clinical Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany; [26]Regional Office of Public Health of Asturias, Oviedo, Spain; [27]Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht, Netherlands; [28]German Institute of Human Nutrition, Potsdam-Rehbrücke, Germany; and [29]Biochemistry Laboratory, Hôpital Edouard Herriot, Lyon, France

Objective: The aim in this study was to assess the association between individual plasma carotenoid levels (α-carotene, β-carotene, lycopene, β-cryptoxanthin, lutein, zeaxanthin) and fruit and vegetable intakes recorded by a calibrated food questionnaire (FQ) and 24-h dietary recall records (24HDR) in nine different European countries with diverse populations and widely varying intakes of plant foods.

Design: A stratified random subsample of 3089 men and women from nine countries participating in the European Prospective Investigation into Cancer and Nutrition (EPIC), who had provided blood samples and dietary and other lifestyle information between 1992 and 2000, were included.

Results: β-Cryptoxanthin was most strongly correlated with total fruits (FQ $r = 0.52$, 24HDR $r = 0.39$), lycopene with tomato and tomato products (FQ $r = 0.38$, 24HDR $r = 0.25$), and α-carotene with intake of root vegetables ($r = 0.39$) and of total carrots ($r = 0.38$) for FQ only. Based on diet measured by FQ and adjusting for possible confounding by body mass index (BMI), age, gender, smoking status, alcohol intake, and energy intake, the strongest predictors of individual plasma carotenoid levels were fruits ($R^2_{partial} = 17.2\%$) for β-cryptoxanthin, total carrots ($R^2_{partial} = 13.4\%$) and root vegetables ($R^2_{partial} = 13.3\%$) for α-carotene, and tomato products ($R^2_{partial} = 13.8\%$) for lycopene. For 24HDR, the highest $R^2_{partial}$ was for fruits in relation to β-cryptoxanthin (7.9%).

Conclusions: Intakes of specific fruits and vegetables as measured by food questionnaires are good predictors of certain individual plasma carotenoid levels in our multicentre European study. At individual subject levels, FQ measurements of fruits, root vegetables and carrots, and tomato products are, respectively, good predictors of β-cryptoxanthin, α-carotene, and lycopene in plasma.

European Journal of Clinical Nutrition (2005) **59**, 1387–1396. doi:10.1038/sj.ejcn.1602252; published online 14 September 2005

Introduction

It has been concluded in several reviews that increasing the consumption of fruits and vegetables is likely to reduce cancer risk (Block *et al*, 1992; WCRF/AICR, 1997; IARC, 2003). The possible cancer-preventing effects may be via carotenoids or via other compounds in fruits and vegetables (Ziegler, 1989; Steinmetz & Potter, 1991; Wattenberg, 1992; IARC, 2003). Carotenoids are phytochemicals that are present in a variety of plant foods. More than 40 carotenoids have been identified, but six of them are found in humans at higher levels than the rest and have therefore received most attention from researchers. These are α-carotene, β-carotene, lutein, β-cryptoxanthin, lycopene, and zeaxanthin. Different fruits and vegetables contain different amounts of these carotenoids. This variability in content may differ even for the same fruit or vegetable depending on the size, growing and harvesting conditions, degree of maturity, processing, storage, and cooking (Willett, 1998). Different populations may therefore have different carotenoid profiles in their plasma according to the type and amount of fruits and vegetables consumed and other bioavailability and demographic factors. Measurement of fruit and vegetable intake by dietary questionnaires and records is prone to measurement error due to day-to-day variation in intake and difficulty in accurately quantifying the amount of intake (Kaaks *et al*, 1994; Byers, 2001; Subar *et al*, 2001). Furthermore, determination of carotenoid intake from dietary questionnaires relies heavily on the accuracy of food composition tables, which are also prone to limitations (Deharveng *et al*, 1999).

If carotenoid plasma levels could estimate the average amount (usually over a period of a few weeks) of total or specific fruit and vegetable groups consumed by an individual, this would provide good biomarker(s) of dietary intake of these foods. The carotenoid plasma level may also be directly related to the risk of these diseases and therefore provide a quantitatively objective measure of the average dose of the exposure of interest. Several earlier studies have found an association between plasma carotenoid levels and intake of fruits and vegetables (Granado *et al*, 1992, 1996; Campbell *et al*, 1994; Michaud *et al*, 1998; Tucker *et al*, 1999; Resnicow *et al*, 2000; Block *et al*, 2001; van Kappel *et al*, 2001a, b). However, these studies usually included small number of subjects (Campbell *et al*, 1994; Michaud *et al*, 1998; Block *et al*, 2001) or did not adjust for other possible confounders in this association (Granado *et al*, 1996). Furthermore, there is limited data from more than one country (van Kappel *et al*, 2001a), making it difficult to assess whether plasma carotenoid levels could predict fruit and vegetable intake in populations with diverse dietary habits.

We report here the results from a large cross-sectional study on plasma levels of six carotenoids within the EPIC cohorts, including men and women from 16 geographical areas in nine European countries. In this analysis, we aim to assess the overall association between individual plasma levels of carotenoids and intake of total and specific fruits and vegetables in a subsample of the EPIC study subjects.

Subjects and methods

EPIC is a multicentre prospective cohort study investigating the relation between diet, nutritional and metabolic characteristics, various lifestyle and environmental factors, and

the risk of cancer and other chronic diseases among 521 483 subjects (Riboli & Kaaks, 1997). In all, 23 research centres in 10 European countries are participating in the EPIC study, which is coordinated by the International Agency for Research on Cancer (IARC) in Lyon, France. Collection of data and blood samples started in 1992. A subsample of 37 000 participants also provided dietary information by 24-h dietary recall (24HDR) records for calibration purposes (Slimani *et al*, 2002). EPIC is unique in that it is one of the largest studies that combine information on diet and lifestyle obtained by means of questionnaires, as well as a biorepository of blood samples collected from 386 080 subjects.

Study population

For the subsample of subjects included in this study, 16 geographical areas (regions) were designated by grouping centres and/or geographical areas within the EPIC study: France (Paris and surroundings), Florence (central Italy), Varese/Turin (northern Italy), Ragusa/Naples (southern Italy), northern Spain (San Sebastian, Pamplona, Oviedo), Granada (southern Spain), Murcia (south-eastern Spain), Cambridge (subjects living in Norfolk), Oxford study centre (vegetarians living throughout the UK), the Netherlands (including subjects from Utrecht and Bilthoven), Athens in Greece, Heidelberg (south-west Germany), Potsdam (former East-Germany), Malmö (southern Sweden), Umeå (northern Sweden), and Denmark (including subjects from Aarhus and Copenhagen). In each of these regions, 100 women and 100 men (except in France where only women had been recruited to the cohort) were randomly selected among the subjects participating in the EPIC calibration substudy involving 5–12% of the EPIC cohort (Slimani *et al*, 2002), with the exception of the Oxford centre, where the subjects selected were all vegetarians and included all available vegans. The selection followed a stratified sampling scheme with 50 subjects (25 men and 25 women) in each of four age-strata (45–49, 50–54, 55–59, and 60–64 y of age at the time of blood sampling). Participants in Denmark, Greece, and Umeå had different age distributions from the rest of the sample, but excluding them from analyses did not modify the age-adjusted carotenoid levels (Al-Delaimy *et al*, 2004). Among the vegetarians from the UK, 65 of the 100 men selected were vegans (eating no animal products) and 35 were lacto-vegetarians (eating no meat or meat products); 88 of the 99 women selected were vegans and 11 were lacto-vegetarians.

In total, 3089 subjects were selected for participation in the study. Aliquots were missing for four subjects, and 42 subjects were excluded because of laboratory and other technical reasons (including 23 subjects who were run in one batch with incorrect identification labels). Values of individual carotenoids from nine subjects were missing and they were therefore excluded. A further 125 subjects were excluded from the analyses because of incomplete informa-

tion (65 for FQ and 60 for 24HDR). Thus, the current analyses include 2969 subjects for FQ and 2974 subjects for 24HDR. Individual food items and smoking information were missing for some subjects, and they were not included in the regression. In some analyses, therefore, the number of subjects may be less than that indicated above.

Dietary data

Information on individual dietary intake is described in detail elsewhere (Riboli *et al*, 2002; Slimani *et al*, 2002). In brief, information on usual individual dietary intakes was assessed by FQ at baseline for each subject entering the EPIC cohort. Countries differed in the type of validated questionnaire they used; some used extensive dietary questionnaires, others semiquantitative questionnaires, and others a diet history methodology combining a food questionnaire with a 7- or 14-day menu record. Some of these were self-reported while in others the information was obtained through interview. The total intakes of fruits and vegetables were calculated as continuous variables by multiplying the frequency of intake by the portion size for each dietary item divided by the same unit of time using a common classification and food definition across countries (Riboli *et al*, 2002).

Although a single 24HDR is not expected to be representative of average fruit and vegetable intake for an individual, we wanted to assess the association of such a measure with biomarkers of carotenoids regardless of the timing of blood drawing (only 45.5% of 24HDR were collected in the same season as collection of blood samples, while up to 87% of questionnaires were completed in the same season that the corresponding blood sample was taken). A single 24HDR measurement, by means of highly standardized interviews, was taken across all EPIC countries using a computerized program (EPIC-SOFT) developed for this purpose (Slimani *et al*, 1999, 2002). Information on all of the foods and beverages consumed during the previous day was collected, entered, and coded automatically according to common rules. Individual food portion sizes were estimated using a common picture book containing sets of photographs of 140 foods and recipes (van Kappel *et al*, 1994), and other available methods such as standard units and household measurements. Trained dieticians conducted all of the interviews face-to-face. More details on the concept of standardization and structure of EPIC-SOFT and the distribution of fruit and vegetable intake among participating countries are described in detail elsewhere (Slimani *et al*, 1999, 2000; Agudo *et al*, 2002).

Blood collection and laboratory analyses

When study participants visited the local study centre for completion of questionnaires and anthropometric measurements, blood samples were taken and then separated and aliquoted for storage in plastic straws: plasma (12 straws);

serum (eight straws); leukocytes (four straws); and erythrocytes (four straws). For each subject, half the aliquots were stored locally in the study centre while the other half was shipped in dry ice or liquid nitrogen to the central biorepository at IARC in Lyon, France. Carotenoids are little affected by short-term storage and transport (Hankinson et al, 1989; Key et al, 1996), and there was no effect of storage time or laboratory methods on the levels of carotenoids. More details about blood collection and storage are described elsewhere (Al-Delaimy et al, 2004).

The laboratory analytical method is described in an earlier study (Steghens et al, 1997). In brief, samples (200 μl) were analyzed for carotenoids by reversed-phase high-performance liquid chromatography (HPLC-1100 system, Hewlett Packard). Samples were analyzed in batches homogenous for sex and age category and in randomized order for region of residence of subjects. Peaks for carotenoids that were under the quantification limits were set to zero. There were two samples for lutein, 14 for zeaxanthin, three samples for lycopene, five samples for β-carotene, and 47 samples for α-carotene that were below the limit of quantification. The 47 samples below the limit of quantification for α-carotene (0.2 μmol/l) were due to the relatively low levels of α-carotene compared to other carotenoids and the limited sensitivity of the assay to quantify such low levels with precision, and they were therefore set to zero. Peaks that could not be detected because of technical problems were excluded. Results for six plasma carotenoids (α-carotene, β-carotene, lycopene, β-cryptoxanthin, lutein, and zeaxanthin) are presented in this paper. Between-day coefficients of variation in concentrations of individual levels and over the entire period of analysis (11 months) were less than 7.6% except for zeaxanthin (16.5%). No significant between-day drift was observed.

Statistical methods

The errors in 24 HDR measurements are expected to be random, distributed around individuals' true mean intake estimates. With one replicate per subject available, 24 HDR measurements are expected to provide correct estimates of intake at the aggregate levels, and have been shown to be highly correlated to biomarkers of absolute intake in our earlier analyses (Slimani et al, 2003). FQ values were therefore rescaled on centre-specific 24HDR means using an additive correction factor to compute calibrated questionnaire measurements, FQ*, as

$$FQ_{ij}^* = FQ_{ij} + (\overline{24HDR_j} - \overline{FQ_j})$$

where $i = 1, \ldots n_j$ refers to the sample size in each centre, $j = 1, \ldots, J$ indexes the EPIC centres and $\overline{24HDR_j}$ and $\overline{FQ_j}$ are the centre-specific sample means. This solution provides an ecological calibration in order to improve comparability of FQ measurements across EPIC centres. In the present work, we aimed at comparing the overall association between biomar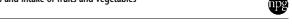ker levels and dietary questionnaire intakes. In order to provide a comparison at the individual level, while exploiting the overall heterogeneity of exposure, we removed the centre-specific bias in FQ measurements, while the within-centre variability of FQ measurements is not altered.

Spearman's rank correlation coefficient was used to assess the association between dietary variables and plasma carotenoid levels. The dietary variables were chosen based on initial crude scatter plots screening for different fruits and vegetables and on previously existing data. The dietary variables considered in our analyses are total fruits and vegetables, total vegetables, total fruits, green leafy vegetables, root vegetables, cabbage, citrus fruits, fruits other than citrus, total tomato products, and total carrots. Multivariate regression analysis using GLM procedures in SAS (SAS Institute, Cary, NC: SAS version 8.0) was used to assess the contribution of dietary variables to plasma carotenoid levels while adjusting for other possible confounders. Based on earlier analyses (Al-Delaimy et al, 2004), the model included the following variables as covariates: gender, season of blood collection, BMI, age, alcohol intake, smoking status (non-smoker, past smoker, current smoker), and energy intake levels. Region was not adjusted for in the model because calibrated FQ data were used in the statistical analyses. To evaluate the overall correlation coefficients, no adjustment by region was performed, otherwise the results would refer to the within-center component of the association rather than the overall correlation of interest to us. Calculation of partial R^2 ($R^2_{partial}$) (using the type III sum of squares analyses) was used to assess the degree of variability each dietary variable contributes to plasma carotenoids (Kleinbaum et al, 1988). $R^2_{partial}$ expresses the degree of variability explained by an independent variable, given other independent variables in the model, divided by the residual sum of squares of the model excluding that independent variable and then multiplied by 100 to express it as a percentage. This percentage represents the amount of variability of the individual carotenoid levels (dependent variable) explained by the specific dietary variable.

Results

Relevant characteristics of the study subjects are presented in Table 1. The mean plasma carotenoid levels are shown according to age, gender, smoking status, season of blood collection, BMI, and alcohol intake. Lycopene levels were lower with increased age for both men and women, while α- and β-carotene were higher with age for men. Women had higher levels of carotenoids except for lycopene, which was slightly higher for men, and levels were similar for zeaxanthin. Lutein, β-cryptoxanthin, α- and β-carotene levels were lower among past and current smokers compared to never smokers. As expected, there were seasonal effects on levels of specific carotenoids. β-Cryptoxanthin levels were higher in winter when citrus fruits are abundant, while

Table 1 Mean levels (with standard errors) of plasma carotenoids (μmol/l) according to categories of the participants' age and season at the time of blood collection, smoking, alcohol intake, gender and BMI

	Lutein		Zeaxanthin		β-Crypto xanthin		Lycopene		α-Carotene		β-Carotene		Total carotenoids	
	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.	Mean	s.e.
Age at blood collection (y)														
Men														
45–59	0.36	0.011	0.07	0.002	0.24	0.010	0.79	0.023	0.10	0.006	0.33	0.014	1.91	0.045
50–54	0.36	0.010	0.09	0.003	0.22	0.010	0.76	0.021	0.12	0.007	0.36	0.019	1.93	0.045
55–59	0.41	0.012	0.10	0.003	0.23	0.011	0.74	0.023	0.13	0.007	0.35	0.013	1.98	0.047
60–64	0.39	0.011	0.09	0.003	0.24	0.010	0.67	0.023	0.15	0.008	0.39	0.016	1.95	0.044
Women														
45–49	0.42	0.012	0.09	0.002	0.35	0.014	0.74	0.021	0.20	0.011	0.54	0.019	2.36	0.049
50–54	0.43	0.011	0.09	0.003	0.33	0.012	0.73	0.020	0.20	0.009	0.54	0.027	2.34	0.054
55–59	0.48	0.013	0.10	0.002	0.37	0.016	0.70	0.020	0.21	0.009	0.55	0.024	2.42	0.053
60–64	0.42	0.013	0.09	0.002	0.32	0.013	0.68	0.023	0.20	0.010	0.54	0.024	2.27	0.056
Gender														
Male	0.38	0.005	0.09	0.001	0.23	0.005	0.74	0.011	0.12	0.004	0.36	0.008	1.94	0.023
Female	0.44	0.006	0.09	0.001	0.34	0.007	0.71	0.011	0.20	0.005	0.54	0.012	2.35	0.027
Smoking status														
Former smoker	0.41	0.007	0.09	0.002	0.26	0.007	0.74	0.014	0.16	0.006	0.44	0.013	2.12	0.031
Current smoker	0.36	0.009	0.08	0.002	0.22	0.008	0.71	0.017	0.11	0.004	0.34	0.011	1.85	0.036
Unknown	0.46	0.036	0.10	0.008	0.37	0.049	0.83	0.074	0.20	0.048	0.57	0.070	2.55	0.190
Season of blood collection														
Spring	0.44	0.009	0.10	0.002	0.29	0.008	0.75	0.016	0.15	0.005	0.48	0.015	2.24	0.037
Summer	0.37	0.008	0.09	0.002	0.22	0.006	0.81	0.016	0.17	0.007	0.49	0.015	2.15	0.037
Autumn	0.42	0.008	0.08	0.002	0.25	0.008	0.74	0.015	0.18	0.007	0.46	0.015	2.15	0.036
Winter	0.41	0.008	0.09	0.002	0.37	0.011	0.63	0.014	0.15	0.006	0.41	0.015	2.08	0.034
BMI (kg/m²)														
< 21	0.48	0.020	0.09	0.004	0.29	0.016	0.81	0.033	0.29	0.020	0.71	0.042	2.69	0.087
21–23.9	0.43	0.009	0.09	0.002	0.27	0.009	0.76	0.017	0.21	0.008	0.55	0.021	2.32	0.041
24–26.9	0.41	0.007	0.09	0.002	0.29	0.008	0.74	0.013	0.16	0.005	0.45	0.011	2.16	0.029
27–29.9	0.41	0.009	0.09	0.002	0.31	0.011	0.69	0.017	0.14	0.006	0.40	0.014	2.06	0.037
30–33	0.38	0.012	0.09	0.003	0.29	0.013	0.68	0.023	0.10	0.006	0.32	0.012	1.90	0.046
>33	0.37	0.015	0.09	0.003	0.28	0.016	0.66	0.027	0.09	0.006	0.33	0.021	1.84	0.057
Alcohol intake (g/day)														
0	0.42	0.011	0.09	0.002	0.37	0.015	0.71	0.022	0.15	0.007	0.41	0.014	2.18	0.047
0–15	0.41	0.005	0.09	0.001	0.29	0.006	0.74	0.010	0.18	0.005	0.51	0.011	2.24	0.025
15.1–30	0.42	0.011	0.09	0.002	0.25	0.009	0.73	0.018	0.15	0.007	0.41	0.016	2.08	0.041
>30	0.39	0.011	0.09	0.003	0.23	0.012	0.69	0.020	0.11	0.005	0.31	0.011	1.85	0.043

lycopene levels were higher in summer when tomatoes are available. Carotenoids were generally lower with higher BMI, but zeaxanthin and β-cryptoxanthin levels do not seem to be affected by BMI. β-Cryptoxanthin levels were negatively correlated with alcohol intake levels.

Spearman correlation coefficients between individual carotenoids are shown in Table 2. The highest correlation was between lutein and zeaxanthin ($r = 0.75$), while the lowest was between α-carotene and lycopene and β-cryptoxanthin ($r = 0.11$). The correlations between vegetable and fruit intake and plasma carotenoids are presented in Table 3. FQ dietary variables were more strongly and significantly correlated with plasma carotenoid levels than 24HDR variables, although some of the 24HDR variables followed a similar trend of correlations to that of FQ with plasma

carotenoid levels. The strongest correlations were between total 'fruits' and plasma β-cryptoxanthin (FQ $r = 0.52$, 24HDR $r = 0.39$). The variable 'fruits and vegetables' was also strongly correlated with plasma β-cryptoxanthin (FQ $r = 0.46$, 24HDR $r = 0.39$) and to a lesser extent with lutein (FQ $r = 0.38$, 24HDR $r = 0.31$) and zeaxanthin (FQ $r = 0.36$, 24HDR $r = 0.32$). Lutein plasma levels were slightly more correlated with total 'vegetables' (FQ $r = 0.30$, 24HDR $r = 0.23$) than the other carotenoids. Root vegetables and total carrots from FQ were specifically correlated to α-carotene ($r = 0.39$ and $r = 0.38$, respectively). Lycopene correlated most with tomato products (FQ $r = 0.38$). Lutein was correlated most with fruits other than citrus (FQ $r = 0.42$).

The results of multivariate regression analyses for the association between dietary variables and plasma carotenoid

Table 2 Spearman correlation coefficients ($P<0.0001$) between individual log normal plasma carotenoids levels (μmol/l)

	Lycopene	β-Cryptoxanthin	Zeaxanthin	α-Carotene	β-Carotene
Lutein	0.39	0.36	0.73	0.22	0.34
Lycopene		0.19	0.26	0.11	0.30
β-Cryptoxanthin			0.48	0.11	0.25
Zeaxanthin				0.13	0.24
α-Carotene					0.71

Table 3 Spearman correlation coefficients for the correlation between log-transformed carotenoid plasma levels (μmol/l) and log-transformed fruit and vegetable variables recorded by FQ ($n=2969^a$) and 24HDR ($n=2974$)

	Dietary method	Lutein	Zeaxanthin	β-Cryptoxanthin	Lycopene	α-Carotene	β-Carotene	Total carotenoids
Fruits and vegetables	FQ	0.38	0.36	0.46	0.24	0.09	0.17	0.38
	24HDR	0.31	0.32	0.39	0.14	0.10	0.15	0.30
Vegetables	FQ	0.30	0.26	0.26	0.19	0.16	0.21	0.31
	24HDR	0.23	0.22	0.19	0.12	0.13	0.16	0.21
Leafy vegetables	FQ	0.24	0.27	0.25	0.06	0.06	0.12	0.20
	24HDR	0.14	0.16	0.14	*	*	0.07	0.11
Root vegetables	FQ	*	−0.08	−0.07	−0.06	0.39	0.18	0.05
	24HDR	*	−0.04	*	*	0.20	0.09	*
Cabbages	FQ	*	−0.06	−0.07	*	0.18	0.09	*
	24HDR	*	*	−0.05	*	0.10	0.05	*
Tomato and tomato products	FQ ($n=2298$)	0.36	0.25	0.25	0.38	0.08	0.19	0.39
	24HDR	0.18	0.14	0.10	0.25	0.06	0.12	0.22
Total carrots	FQ	*	−0.05	*	*	0.38	0.18	0.08
	24HDR	*	*	*	*	0.19	0.08	0.04
Fruit	FQ	0.37	0.37	0.52	0.22	*	0.11	0.36
	24HDR	0.25	0.26	0.39	0.10	0.06	0.11	0.25
Fruit other than citrus	FQ ($n=2390$)	0.42	0.32	0.40	0.32	0.18	0.24	0.43
	24HDR	0.21	0.21	0.27	0.09	0.08	0.12	0.21
Citrus fruit	FQ ($n=2390$)	0.16	0.14	0.41	0.10	0.06	0.09	0.20
	24HDR	0.09	0.10	0.26	*	*	*	0.09

[a]Unless indicated in the table.
*Nonsignificant *P*-values, $P \geqslant 0.05$.

levels, adjusted for age, gender, alcohol intake, smoking status, BMI, season of collection of blood sample, and energy intake are summarized in Table 4. $R^2_{partial}$ values are presented in the table for each dietary variable in the FQ and 24HDR in relation to plasma carotenoids. As expected, FQ data provided a better prediction of individual plasma carotenoid levels. After adjustment for possible confounders, there were weaker associations for some FQ fruit and vegetable variables and most 24HDR fruit and vegetable variables with individual plasma carotenoid levels. The variables most strongly correlated with total carotenoid levels were 'fruits and vegetables' (15%), 'tomato and tomato products' (13.3%), and 'fruits other than citrus' (14.1%). The strongest dietary variable predictors of individual carotenoids were fruits (17.2%) for β-cryptoxanthin, total carrots (13.4%) and root vegetables (13.3%) for α-carotene, and tomato products

(13.8%) for lycopene. For 24HDR, the highest $R^2_{partial}$ was for fruits in relation to β cryptoxanthin (7.9%).

Discussion

In our study, we found that carotenoid plasma levels are related to variables of dietary fruit and vegetable intake. The relationship, however, varies. Plasma β-cryptoxanthin levels were strongly predicted by fruit intake, while α-carotene levels were strongly predicted by total carrot intake and lycopene by tomato products even after adjustment for other confounding factors in a diverse European population. As expected, FQ data were better able to predict plasma carotenoid levels at the individual level than a single 24HDR because FQ represents a longer duration of past

Table 4 $R^2_{partial}$ values that reached statistical significance for FQ and 24HDR log-transformed dietary variables in relation to individual log-transformed plasma carotenoid levels in a multivariate regression model adjusted for age, gender, alcohol intake, smoking status, BMI, season of blood sample collection, and energy intake

	Dietary method	Lutein (%)	Zeaxanthin (%)	β-Cryptoxanthin (%)	Lycopene (%)	α-Carotene (%)	β-Carotene (%)	Total carotenoids (%)
Fruits and vegetables	FQ	**11.4**	**11.3**	**15.8**	6.1	1.6	2.1	15.0
	24HDR	5.5	5.4	7.3	1.8	0.8	1.2	6.5
Vegetables	FQ	7.3	5.8	3.7	3.9	3.5	3.7	9.8
	24HDR	3.1	2.7	1.7	1.1	0.9	1.13	3.3
Leafy vegetables	FQ	4.5	6.9	3.7	0.5	0.8	1.3	4.1
	24HDR	1.7	1.8	0.8	0.2	*	0.3	1.1
Root vegetables	FQ	0.4	0.7	0.9	0.9	**13.3**	1.8	*
	24HDR	0.3	0.3	0.3	0.2	4.6	0.6	*
Cabbages	FQ	0.1	0.4	0.8	0.4	1.5	*	*
	24HDR	*	*	0.4	*	0.7	*	*
Tomato and tomato products	FQ	9.5	5.7	3.8	**13.8**	0.8	2.1	13.3
	24HDR	2.8	1.8	0.8	5.2	0.2	0.5	4.0
Total carrots	FQ	*	0.2	0.3	0.4	**13.4**	1.9	0.3
	24HDR	0.1	*	*	*	4.7	0.6	*
Fruit	FQ	9.7	9.6	**17.2**	5.0	0.2	0.6	11.2
	24HDR	3.8	3.5	7.9	1.2	0.2	0.5	4.4
Fruit other than citrus	FQ	**12.0**	7.8	**10.1**	9.5	1.4	1.8	14.1
	24HDR	2.6	2.4	3.9	0.9	0.4	0.7	3.3
Citrus fruit	FQ	1.6	1.8	**12.9**	0.9	*	*	2.8
	24HDR	0.6	0.9	5.1	*	*	*	0.8

*P-value $\geqslant 0.05$, P-values of F-test on type III sum of squares estimate. Values in bold type indicate the carotenoid values that can be explained most strongly by a specific dietary variable.

exposure, and most questionnaires were collected in the same season as blood collection. However, some 24HDR dietary variables were also significantly correlated with individual plasma carotenoid levels, which suggests that a single 24HDR record for assessment of intake of certain fruits and vegetables in Europe has some relevance for actual average intake.

The country-specific dietary questionnaires used in this study were not standardized across countries and might contain measurement errors varying both in nature and magnitude. However, these systematic measurement errors were minimized by using the calibration method described in the Methods section. 24HDR measurements provide unbiased estimates at the population level. Although it is recognized that all self-reported dietary intakes, including 24-HDR, contain measurement errors, if the direction and magnitude of systematic dietary measurements are approximately constant across study populations, the reference method can be used for between-centre calibration. The questionnaire is therefore calibrated against a dietary method with only a relative validity but which is comparable across the study population. There were strong intercorrelations between the individual carotenoids, and this would

make it difficult to differentiate completely the associations between specific carotenoids and fruits and vegetables. No single fruit or vegetable variable is completely specific to an individual carotenoid plasma level because carotenoids are widely distributed in most plants. Nevertheless, it was clear from the results that the carotenoids that were better markers of dietary intake were those that were more specific to certain fruits or vegetables. Therefore the higher specificity of a plasma carotenoid as a biomarker is gauged by the relatively higher quantity that a certain fruit or vegetable contributes to that specific carotenoid compared to all the others.

Few earlier studies have assessed the association between plasma levels of carotenoids as biomarkers of fruit and vegetable intake. Campbell et al (1994) found that all carotenoids were related to total fruit and vegetable intake among 99 participants from Minnesota in the US, while in our study we found α- and β-carotene to be weakly correlated to this specific variable compared to the other carotenoids. In their study, β-carotene was well correlated to total fruits and vegetables ($r = 0.45$) and total fruits ($r = 0.46$), while none of the dietary variables in our study were strongly correlated with β-carotene. The differences in results can be

attributed to the type and number of study subjects. At the country level, the correlations of plasma β-carotene levels with total fruits and vegetables varied between 0.11 in Germany and 0.35 in Denmark. The use of β-carotene as a colorant and in fortification of certain foods in some European countries may also explain the weaker correlation in our study. Interestingly, the one correlation that did not change in both studies was that for β-cryptoxanthin and total fruit intake. As in our study, Campbell *et al* found that β-cryptoxanthin was strongly correlated to total fruit intake ($r = 0.56$).

A study in New York involving 302 women (van Kappel *et al*, 2001b) also found that β-cryptoxanthin was most strongly correlated with total fruit intake ($r = 0.40$) compared to four other carotenoids, while all four carotenoids (β-cryptoxanthin, lutein, α- and β-carotene) were similarly correlated to total vegetable intake ($r = 0.25$–0.29). The authors adjusted for BMI, cholesterol and triglyceride levels but the correlations did not change. In our study, which included results from nine countries with varying intakes, α-carotene and lutein were moderately correlated to total vegetable intake. Two more recent Dutch studies also found β-cryptoxanthin to be the carotenoid most strongly correlated to fruits ($r = 0.42$) and fruits and vegetables ($r = 0.41$) (Bogers *et al*, 2004; Jansen *et al*, 2004). The latter study was carried out by scientists participating in the Dutch component of the EPIC study. They found somewhat weaker correlations between carotenoids and fruit and vegetable intake. For β-cryptoxanthin and fruit intake, the correlation was 0.37. These lower correlations were to be expected, since their population was a homogeneous Dutch population, unlike our randomly selected heterogenous European population, which would lead to better correlations.

Other investigators generally reported lower correlations between plasma carotenoids and fruit and vegetable intake (Michaud *et al*, 1998; Tucker *et al*, 1999; Resnicow *et al*, 2000). The results for β-cryptoxanthin were significant for women ($r = 0.34$) in one study (Tucker *et al*, 1999) and for men ($r = 0.36$) in another (Michaud *et al*, 1998) after adjustment for other confounders. In the study by Tucker *et al*, the correlation between β-cryptoxanthin and fruit and vegetable intake was very low for men ($r = 0.16$). Block *et al* (2001) attribute their higher correlations ($r = 0.50$) and those of Campbell *et al* (1994) to the fact that they used only data of extreme fruit and vegetable intake while the other studies included the full range of frequencies. This could be true for studies with small numbers of subjects, but for larger numbers of subjects, as in our study, the results were consistent even though the whole range of fruit and vegetable intake was included.

In a detailed study of the contribution of fruits and vegetables to individual serum carotenoid levels in Spain, Granado *et al* (1996) found that close to 70% of α-carotene came from carrots, and 76% of β-cryptoxanthin from oranges. Two other studies from the US and Spain found similar percentages for these two specific food groups

(Granado *et al*, 1992; Tucker *et al*, 1999). This supports our findings that these two dietary variables had the strongest prediction of α-carotene and β-cryptoxanthin carotenoid levels. However, the variability in intake between countries may lead to variability in the associations between individual fruits and vegetables and specific carotenoids in plasma. For example, in a study comparing fruit and vegetable intake in Malmö (Sweden) and Granada (Spain), it was found that carrot intake was strongly correlated with α-carotene levels in the Swedish population ($r = 0.61$) but less strongly correlated in the Spanish population ($r = 0.47$) (van Kappel *et al*, 2001a). This was explained by the much higher α-carotene levels in Sweden (mean = 0.16 mg/l) compared to Spain (0.09 mg/l).

In conclusion, our study supports the earlier findings that citrus fruits, and fruits in general, can be correlated with individual plasma β-cryptoxanthin levels. Furthermore, this association is consistent even after adjustment for other possible confounders across different European regions. β-Cryptoxanthin in plasma is an acceptable biomarker of average citrus fruit and total fruit intake among Europeans. Similarly, in our study, α-carotene levels were well predicted by total carrot intake, and lycopene by tomato intake. Therefore, specific carotenoids that have been consistently shown to be appropriate biomarkers of specific fruits and vegetables should be measured to complement questionnaires and to better estimate associations between fruit and vegetable or carotenoid intake and disease outcomes.

References

Agudo A, Slimani N, Ocké MC, Naska A, Miller AB, Kroke A, Bamia C, Karalis D, Vineis P, Palli D, Bueno-de-Mesquita HB, Peeters PHM, Engeset D, Hjartåker A, Navarro C, Martínez García C, Wallström P, Zhang JX, Welch AA, Spencer E, Stripp C, Overvad K, Clavel-Chapelon F, Casagrande C & Riboli E (2002): Consumption of vegetables, fruit and other plant foods in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts from 10 European countries. *Publ. Health Nutr.* 5 (Suppl), S1179–S1196.

Al-Delaimy WK, van Kappel AL, Ferrari P, Slimani N, Steghens JP, Bingham S, Johansson I, Wallström P, Overvad K, Tjønneland A, Key TJ, Welch AA, Bueno-de-Mesquita HB, Peeters PHM, Boeing H, Linseisen J, Clavel-Chapelon F, Guibout C, Navarro C, Quirós JR, Palli D, Celentano E, Trichopoulou A, Benetou V, Kaaks R & Riboli E (2004): Plasma levels of six carotenoids in nine European countries: report from the European Prospective Investigation into Cancer and Nutrition (EPIC). *Publ. Health Nutr.* 7, 713–722.

Block G, Patterson B & Subar A (1992): Fruit, vegetables, and cancer prevention: a review of the epidemiologic evidence. *Nutr. Cancer* 18, 1–29.

Block G, Norkus E, Hudes M, Mandel S & Helzlsouer K (2001): Which plasma antioxidants are most related to fruit and vegetable consumption? *Am. J. Epidemiol.* 154, 1113–1118.

Bogers RP, van Assema P, Kester AD, Westerterp KR & Dagnelie PC (2004): Reproducibility, validity, and responsiveness to change of a short questionnaire for measuring fruit and vegetable intake. *Am. J. Epidemiol.* 159, 900–909.

Byers T (2001): Food frequency dietary assessment: how bad is good enough? *Am. J. Epidemiol.* 154, 1087–1088.

Campbell DR, Gross MD, Martini MFC, Grandits GA, Slavin JL & Potter JD (1994): Plasma carotenoids as biomarkers of vegetable and fruit intake. *Cancer Epidemiol. Biomarkers Prev.* 3, 493–500.

Deharveng G, Charrondière UR, Slimani N, Southgate DAT & Riboli E (1999): Comparison of nutrients in the food composition tables available in the nine European countries participating in EPIC. *Eur. J. Clin. Nutr.* 53, 60–79.

Granado F, Olmedilla B, Blanco I & Rojas-Hidalgo E (1992): Carotenoid composition in raw and cooked Spanish vegetables. *J. Agric. Food Chem.* 40, 2135–2140.

Granado F, Olmedilla B, Blanco I & Rojas-Hidalgo E (1996): Major fruits and vegetable contributors to the main serum carotenoids in the Spanish diet. *Eur. J. Clin. Nutr.* 50, 246–250.

Hankinson SE, London SJ, Chute CG, Barbieri RL, Jones L, Kaplan LA, Sacks FM & Stampfer MJ (1989): Effect of transport conditions on the stability of biochemical markers in blood. *Clin. Chem.* 35, 2313–2316.

IARC Handbooks of Cancer Prevention, Vol 8 (2003): *Fruit and Vegetables*. Lyon: International Agency for Research on Cancer.

Jansen MCJF, van Kappel AL, Ocké MC, Van't Veer P, Boshuizen HC, Riboli E & Bueno-de-Mesquita HB (2004): Plasma carotenoid levels in Dutch men and women, and the relation with vegetable and fruit consumption. *Eur. J. Clin. Nutr.* 58, 1386–1395.

Kaaks R, Riboli E, Estève J, van Kappel AL & van Staveren WA (1994): Estimating the accuracy of dietary questionnaire assessments: validation in terms of structural equation models. *Stat. Med.* 13, 127–142.

Key T, Oakes S, Davey G, Moore J, Edmond LM, McLoone UJ & Thurnham DI (1996): Stability of vitamins A, C, and E, carotenoids, lipids, and testosterone in whole blood stored at 4 degrees C for 6 and 24 h before separation of serum and plasma. *Cancer Epidemiol. Biomarkers Prev.* 5, 811–814.

Kleinbaum DG, Kupper LL & Muller KE (1988): *Applied Regression Analysis And Other Multivariable Methods*, 2nd edn. Belmont, CA: Duxbury Press. p 151.

Michaud DS, Giovannucci EL, Ascherio A, Rimm EB, Forman MR, Sampson L & Willett WC (1998): Associations of plasma carotenoid concentrations and dietary intake of specific carote-noids in samples of two prospective cohort studies using a new carotenoid database. *Cancer Epidemiol. Biomarkers Prev.* 7, 283–290.

Resnicow K, Odom E, Wang T, Dudley WN, Mitchell D, Vaughan R, Jackson A & Baranowski T (2000): Validation of three food frequency questionnaires and 24-h recalls with serum carotenoid levels in a sample of African-American adults. *Am. J. Epidemiol.* 152, 1072–1080.

Riboli E & Kaaks R (1997): The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int. J. Epidemiol.* 26 (Suppl 1), S6–S14.

Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondière UR, Hémon B, Casagrande C, Vignat J, Overvad K, Tjønneland A, Clavel-Chapelon F, Thiébaut A, Wahrendorf J, Boeing H, Trichopoulos D, Trichopoulou A, Vineis P, Palli D, Bueno-de-Mesquita HB, Peeters PHM, Lund E, Engeset D, González CA, Barricarte A, Berglund G, Hallmans G, Day NE, Key TJ, Kaaks R & Saracci R (2002): European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Publ. Health Nutr.* 5 (Suppl), S1113–S1124.

Slimani N, Deharveng G, Charrondiere RU, van Kappel AL, Ocké MC, Welch A, Lagiou A, van Liere M, Agudo A, Pala V, Brandstetter B, Andrén C, Stripp C, van Staveren W & Riboli E (1999): Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centres participating in the EPIC project. *Comput. Meth. Programs Biomed.* 58, 251–266.

Slimani N, Ferrari P, Ocké M, Welch A, Boeing H, van Liere M, Pala V, Amiano P, Lagiou A, Mattisson I, Stripp C, Engeset D, Charron-dière R, Buzzard M, van Staveren W & Riboli E (2000): Standardization of the 24-h 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.

Slimani N, Kaaks R, Ferrari P, Casagrande C, Clavel-Chapelon F, Lotze G, Kroke A, Trichopoulos D, Trichopoulou A, Lauria C, Bellegotti M, Ocké MC, Peeters PHM, Engeset D, Lund E, Agudo A, Larrañaga N, Mattisson I, Andren CJ, Johansson I, Davey G, Welch AA, Overvad K, Tjønneland A, van Staveren WA, Saracci R & Riboli E (2002): European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study: rationale, design and popula-tion characteristics. *Publ. Health Nutr.* 5 (Suppl), S1125–S1145.

Slimani N, Bingham S, Runswick S, Ferrari P, Day NE, Welch AA, Key TJ, Miller AB, Boeing H, Sieri S, Veglia F, Palli D, Panico S, Tumino R, Bueno-de-Mesquita B, Ocké MC, Clavel-Chapelon F, Tricho-poulou A, van Staveren WA & Riboli E (2003): Group level validation of protein intakes estimated by 24-h diet recall and dietary questionnaires against 24-h urinary nitrogen in the European Prospective Investigation into Cancer and Nutrition (EPIC) Calibration Study. *Cancer Epidemiol. Biomarkers Prev.* 12, 784–795.

Steghens JP, van Kappel AL, Riboli E & Collombel C (1997): Simultaneous measurement of seven carotenoids, retinol and alpha-tocopherol in serum by high-performance liquid chromato-graphy. *J. Chromatogr. B* 694, 71–81.

Steinmetz KA & Potter JD (1991): Vegetables, fruit, and cancer. II. Mechanisms. *Cancer Causes Control* 2, 427–442.

Subar AF, Thompson FE, Kipnis V, Midthune D, Hurwitz P, McNutt S, McIntosh A & Rosenfeld S (2001): Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires; the Eating at America's Table Study. *Am. J. Epidemiol.* 154, 1089–1099.

Tucker KL, Chen H, Vogel S, Wilson PW, Schaefer EJ & Lammi-Keefe CJ (1999): Carotenoid intakes, assessed by dietary questionnaire, are associated with plasma carotenoid concentrations in an elderly population. *J. Nutr.* 129, 438–445.

van Kappel AL, Amoyel J, Slimani N, Vozar B & Riboli E (1994): *EPIC-SOFT Picture Book for Estimation of Food Portion Sizes*. Lyon: International Agency for Research on Cancer.

van Kappel AL, Martinez-Garcia C, Elmstahl S, Steghens J-P, Chajes V, Bianchini F, Kaaks R & Riboli E (2001a): Plasma carotenoids in

relation to food composition in Granada (Southern Spain) and Malmo (Southern Sweden). *Int. J. Vitam. Nutr. Res.* **71**, 97–102.

van Kappel AL, Steghens JP, Zeleniuch-Jacquotte A, Chajes V, Toniolo P & Riboli E (2001b): Serum carotenoids as biomarkers of fruit and vegetable consumption in the New York Women's Health Study. *Publ. Health. Nutr.* **4**, 829–835.

Wattenberg LW (1992): Inhibition of carcinogenesis by minor dietary constituents. *Cancer Res.* **52** (Suppl 7), S2085–S2091.

Willett WC (1998): *Nutritional Epidemiology* 2nd edn. New York: Oxford University Press.

World Cancer Research Fund/American Institute for Cancer Research (1997): *Food, Nutrition and the Prevention of Cancer: A Global Perspective*. Washington DC: AICR.

Ziegler RG (1989): A review of the epidemiologic evidence that carotenoids reduce the risk of cancer. *J. Nutr.* **119**, 116–122.