Background

Methods for assessing diet are prone to measurement error, which may be substantial in large cohort investigations. Biomarkers can be used as objective measures with which to compare estimates of nutritional exposure using different methods.

Methods

Cross sectional comparisons in 12,474 men and women of regression between biomarkers for vitamin C, sodium, potassium, fibre, carbohydrate, fat and phytoestrogens with intakes derived from food diaries and food frequency questionnaires (FFQ), and odds ratios for risk of ischaemic heart disease (IHD) by dietary and plasma vitamin C.

Results

There were strong ($P < 0.001$) associations between biomarkers and intakes as assessed by food diary. Coefficients were markedly attenuated for data obtained from the FFQ, especially so for vitamin C, potassium and phytoestrogens ($Z P < 0.05$). Risk of IHD was associated with plasma vitamin C ($P < 0.001$) and intake of vitamin C and fruit and vegetables assessed by food diary ($P$ quintile trends <0.001, 0.001) but not by the FFQ ($P$ quintile trends 0.923, 0.186).

Conclusions

Nutritional data that reflect the findings from biomarkers reduce measurement error and will thus improve statistical power in studies of gene nutrient interactions in cohort studies.

Keywords

Diet, food frequency questionnaires, food diaries, biomarkers, EPIC, ischaemic heart disease

Introduction

It is accepted that very large cohort studies are needed for investigating the complex interaction between genetic and environmental factors in common chronic diseases. Nevertheless, although there has been rapid progress in identifying genetic variations that contribute to these diseases, precise and quantitative techniques for the measurement of environmental exposures have not progressed so rapidly. The majority of data relating diet and genetic factors to individual risk of disease in large cohorts relies on a simple method, the food frequency questionnaire (FFQ).
However, misclassification from FFQ leads to attenuation of associations between diet and disease that cannot be corrected even in very large studies.6–11 This concern has lately been described as a ‘crisis’ with hundreds of millions of dollars invested in epidemiological studies using the FFQ.10 The importance of measuring exposure accurately is emphasized by the fact that the subtle effects of gene variants may only become detectable if the presence of certain environmental exposures is also measured.12,13

We have previously used biomarkers to assess the accuracy of different dietary measurement techniques used in our cohort study, the Norfolk arm of EPIC (European Prospective Investigation of Cancer).14 As a result of this validation study with biomarkers we used three different methods to assess diet, a 24h recall, a FFQ and a detailed food diary kept over 7 days.15 Increased breast cancer risk was strongly associated with saturated fat intake assessed using the detailed 7 day food diary but not a FFQ.16 In a larger US-based study, relative risks for breast cancer for women with fat intake assessed by 4 day food records were significant (P for trend 0.02) but not for fat intake when assessed by a FFQ (P for trend 0.24).17 Measurement error in the assessment of diet is an acknowledged contributor to the controversy surrounding the role of fat in breast cancer risk.18

Here, we present data on 12,474 individuals from the EPIC Norfolk cohort who completed both a FFQ and food diary when recruited in 1992–94. We compare the relationship between intake from diet assessed by these two methods with nutritional biomarkers collected in the same individuals. We also show the relationship between dietary vitamin C and fruit and vegetable intake from these two methods compared with plasma vitamin C in risk of ischaemic heart disease (IHD). The results have implications for large investments of public funds in cohort studies of interactions between genetic and nutritional factors within a number of populations such as the UK Biobank, the USA and in Asia.1–4

Methods

EPIC Norfolk is a cohort of men and women recruited at age 45–75 years. In 1993, a total of 35 Norfolk medical practices agreed to participate and invited 77,630 individuals to take part in the study. A Health and Lifestyle questionnaire was sent to respondents that was returned with completed informed consent forms by 30,445 participants. The questionnaire included questions on smoking, alcohol consumption, socio-economic status, social class, occupational history, use of medication, dietary changes, history of disease, short family history of main disease end points and reproductive history for women.14 Exercise was measured by the means of a simple physical activity index, previously validated.19 Of the participants, 25,639 then agreed to a health examination at which a blood sample, spot urine sample and data on height, weight, respiratory function, anthropometry and blood pressure were collected by trained nurses. All 30,445 participants have been followed up for their health status by flagging with the National Health Service Central Register for death and cancer incidence, the local East Anglian Cancer Registry and the district data base of all in-patient hospital activity. Incident cases of IHD were identified from hospital discharge data or were those who had an underlying cause of IHD (ICD 120-125) on their death certificate. Those who had been told by their doctor at baseline that they had had a heart attack or stroke were excluded. As part of follow up all participants were invited back for a second health check at the beginning of 1998 at which the protocol was repeated. A total 15,783 individuals attended, of whom 15,025 had attended the first health check. Permission for the study was obtained from The Norfolk and Norwich Hospital Ethics Committee. Consent was provided by the participants for the use of their medical records, to attend a health check and for blood samples to be used at a later date.

Dietary data were obtained using FFQ and 7 day food diaries.15 The 130-item FFQ was sent with the appointment for the medical examination and were brought to the clinic (usually within about 2 weeks) where they were checked for completeness by nurses at the medical examination. The completed questionnaires were designed to estimate habitual intake over the previous year. Nutrients were computed using an in-house programme, the CAFE (Compositional Analyses from Frequency Estimates) programme.20 At the medical examination, nurses asked the participants to complete a 7 day food diary comprising an A5, 45 page colour booklet containing food portion photographs and detailed instructions in which the description, preparation and amounts of foods eaten at main meals, snacks and between meals over a week could be recorded.15,21 Nutrients were calculated using a custom-designed dietary assessment software program, DINER (Data Into Nutrients for Epidemiological Research) program.21

Non-fasting serum total cholesterol, high density lipoprotein cholesterol and triglyceride levels were measured with an RA 1000 Technicon analyzer (Bayer Diagnostics, Basingstoke, UK). Urine specimens were frozen without preservative at −20°C. Between 1998 and 2002, the urine samples were thawed and assayed for creatinine on a Roche Cobas Mira Plus analyzer, and sodium and potassium concentrations by flame photometry (IL 943; Instrumentation Lab, Warrington, UK). For vitamin C, 0.25 ml plasma was collected in citrate tubes and stabilized with 0.5 ml 10% metaphosphoric acid prepared fresh weekly. Samples were stored at −70°C until analysis (within 1 week of collection at Addenbrookes Hospital, Cambridge, UK). The plasma ascorbic acid concentration was estimated using a flurometric assay, as
Results

Tables 1 and 2 show the linear relationships between the intakes of nutrients assessed by both methods and the related biomarker. Trends across quintiles and differences between the top and bottom quintiles of dietary intake data were all \( P < 0.0001 \) for the data derived from the diary. However, the coefficients were markedly attenuated for data obtained from the FFQ when compared with all biomarkers, especially so for plasma vitamin C and dietary vitamin C, for urinary potassium and dietary potassium when assessed by FFQ and for urinary and serum diadzein and genistein (\( Z P < 0.05 \), Tables 1 and 2). Dietary and plasma vitamin C associations were similar without adjustment for carbohydrate and fat (coefficient trend diary 5.62 (CI 5.4–5.9) \( P < 0.001 \), trend FFQ 4.2 (CI 3.9–4.4) \( P < 0.001 \), \( Z P < 0.05 \). There was no association between HDL and total fat [for example coefficient Q5 vs Q1 diary 0.014 (CI −0.02 to 0.05) \( P = 0.435 \), nor between triglycerides and total fat [for example coefficient diary Q5 vs Q1 −0.03 (CI −0.11 to 0.05) \( P = 0.458 \). The positive trends between fibre and plasma HDL were weaker than the inverse trends for carbohydrate, for example coefficient Q5 vs Q1 diary 0.027 (CI 0.01–0.05) \( P = 0.019 \). Alcohol was included throughout the analyses and had significant associations with blood lipids, especially HDL [for example Q1 vs Q5 diary 0.178 (CI 0.16–0.20) \( P = <0.001 \). However, protein had no significant effect on blood lipids (for example for HDL Q1 vs Q5 diary −0.007 (CI −0.03 to 0.02) \( P = 0.577 \) and inclusion of protein did not appreciably affect the magnitude of the regression coefficients [for example for carbohydrate and HDL Q1 vs Q5 diary −0.175 (CI −0.21 to −0.14) \( P < 0.001 \). Results shown in the Tables were, therefore, not adjusted for protein. Results for sex specific analyses were consistent with the combined analyses (data not shown).

Table 3 shows comparisons between diet at the first health check and plasma lipids and vitamin C from the second health check 4 years later in the 7370 individuals for whom data was available. Associations between blood lipids and saturated fat, carbohydrate and dietary fibre measured by the diary remained strong, but associations measured by the FFQ were attenuated, as before, and especially so for plasma HDL and carbohydrate (\( Z P < 0.05 \)). Mean plasma vitamin C had increased from 52 (SEM 0.19) at the first health check to 62 (SEM 0.31) mmol/l at the second health check and the relationship between diet and plasma vitamin C assessed by the first health check diary was lessened and the difference between methods was no longer apparent (\( Z P > 0.05 \).

Table 4 shows quintile means of plasma vitamin C and fruits and vegetables, together with HR in individuals who were free of IHD at the time of recruitment but who later developed IHD. As previously reported, there was an inverse relationship (\( P < 0.001 \)) between first health check plasma vitamin C as a marker for fruit and vegetable intake and risk of IHD.23

When vitamin C intake and fruit and vegetable intake were assessed using the food diary, an inverse relationship was observed (\( P < 0.001 \)). However, the relationships between vitamin C and fruit and vegetables and IHD risk were not apparent when assessed by the FFQ. This is despite the almost 2-fold greater apparent intake of fruits and vegetables when assessed by the FFQ compared with the food diary (Table 4). The Figure 1 shows quintile IHD HR estimates for fruits and vegetables and plasma vitamin C.

Discussion

Validation of dietary methods should ideally be carried out using quantitative recovery biomarkers.
Table 1  Linear regression data on 12 474 men and women who completed both a 7 day diary and a FFQ at baseline in 1993–1997 in EPIC-Norfolk

<table>
<thead>
<tr>
<th>Item</th>
<th>Diary n = 12 474</th>
<th>FFQ n = 12 474</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>CI trend</td>
</tr>
<tr>
<td></td>
<td>Coefficient</td>
<td>CI trend</td>
</tr>
<tr>
<td>Trends</td>
<td>Q1 vs Q5</td>
<td>Q1 vs Q5</td>
</tr>
<tr>
<td>Plasma LDL and saturated fat percentage energy</td>
<td>0.028</td>
<td>0.01 to 0.04</td>
</tr>
<tr>
<td>Plasma HDL and dietary carbohydrate</td>
<td>−0.044</td>
<td>−0.05 to −0.04</td>
</tr>
<tr>
<td>Plasma triglycerides and dietary fibre</td>
<td>−0.031</td>
<td>−0.04 to −0.02</td>
</tr>
<tr>
<td>Urinary sodium and dietary sodium (mg)</td>
<td>3.996</td>
<td>3.40 to 4.58</td>
</tr>
<tr>
<td>Urinary potassium and dietary potassium (mg)</td>
<td>2.967*</td>
<td>2.61 to 3.32</td>
</tr>
<tr>
<td>Plasma vitamin C and dietary vitamin C</td>
<td>5.564*</td>
<td>5.33 to 5.79</td>
</tr>
</tbody>
</table>

Individuals were categorized into quintiles of the distribution in sex specific nutrient intake by either method. This table shows linear regression results for the trend across quintiles and top vs lower quintile and plasma LDL, HDL, triglycerides and vitamin C and urinary sodium and potassium. All results were adjusted for sex, age, smoking, exercise, weight, systolic blood pressure, use of antihypertensive and cholesterol lowering drugs and intake of energy and the energy yielding nutrients alcohol, fat and carbohydrate, except for the plasma LDL and percentage saturated fat energy regressions, which were adjusted for carbohydrate only and the plasma HDL and dietary carbohydrate which were adjusted for total fat only. Blood samples were collected in a non-fasting state and urines were casual spot samples corrected for creatinine.

*Z P < 0.05  β coefficients FFQ vs diary.
Table 2  Linear regression data on 596 men and women who completed both a 7 day diary and a FFQ at baseline in 1993–1997 in EPIC-Norfolk

<table>
<thead>
<tr>
<th></th>
<th>Diary n = 596</th>
<th>FFQ n = 596</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urinary daidzein and dietary daidzein</td>
<td>0.071* 0.043–0.099 &lt;0.001 0.284 0.158–0.410 &lt;0.001</td>
<td>0.024* −0.01 to 0.05 0.109 0.112 −0.018 to 0.241 0.090</td>
</tr>
<tr>
<td>Urinary genistein and dietary genistein</td>
<td>0.060* 0.035–0.085 &lt;0.001 0.225 0.113–0.336 &lt;0.001</td>
<td>0.016* −0.01 to 0.04 0.220 0.086 −0.029 to 0.201 0.144</td>
</tr>
<tr>
<td>Serum daidzein and dietary daidzein</td>
<td>0.070* 0.039–0.100 &lt;0.001 0.300 0.163–0.438 &lt;0.001</td>
<td>0.021* −0.01 to 0.05 0.202 0.115 −0.027 to 0.258 0.113</td>
</tr>
<tr>
<td>Serum genistein and dietary genistein</td>
<td>0.053* 0.030–0.076 &lt;0.001 0.190 0.086–0.294 &lt;0.001</td>
<td>0.015* −0.01 to 0.04 0.207 0.080 −0.028 to 0.188 0.146</td>
</tr>
</tbody>
</table>

Individuals were categorized into quintiles of the distribution in sex-specific phytoestrogen intake by either method. This table shows linear regression results for the trend across quintiles and top vs lower quintile and plasma or urinary phytoestrogens. Blood samples were collected in a non-fasting state and urines were casual spot samples corrected for creatinine.

*Z P < 0.05 \( \beta \) coefficients FFQ vs diary.

Table 3  Linear regression data on 7401 men and women who completed both a 7 day diary and a FFQ at baseline in 1993–1997 in EPIC-Norfolk and attended a second health check 1997–2000

<table>
<thead>
<tr>
<th>Item</th>
<th>Diary n = 7401</th>
<th>FFQ n = 7401</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trends</td>
<td>Q1 vs Q5</td>
</tr>
<tr>
<td></td>
<td>Coefficient</td>
<td>CI</td>
</tr>
<tr>
<td>Plasma LDL and saturated fat percentage energy</td>
<td>0.032</td>
<td>0.01 to 0.05</td>
</tr>
<tr>
<td>Plasma HDL and dietary carbohydrate</td>
<td>−0.043*</td>
<td>−0.06 to −0.03</td>
</tr>
<tr>
<td>Plasma triglycerides and dietary fibre</td>
<td>−0.027</td>
<td>−0.05 to −0.01</td>
</tr>
<tr>
<td>Plasma vitamin C and dietary vitamin C</td>
<td>3.39</td>
<td>2.97 to 3.81</td>
</tr>
</tbody>
</table>

Individuals were categorized into quintiles of the distribution in sex specific nutrient intake by either method. This table shows linear regression results for the trend across quintiles and top vs lower quintile and plasma LDL, HDL, triglycerides and vitamin C from bloods taken at the second health check. All results were adjusted as in Table 1. Blood samples were collected in a non-fasting state.

*Z P < 0.05 \( \beta \) coefficients FFQ vs diary.
such as doubly labelled water or 24 h urine collections. We used this approach to inform our early decision to obtain information using three different methods in the EPIC Norfolk cohort. The availability and expense of quantitative recovery biomarkers made their use not possible on this, the largest population comparison sample of dietary methods incorporating biomarkers ever studied. However, plasma vitamin C, plasma and spot urinary phytosterols, and spot urinary sodium and potassium are well established comparison biomarkers of dietary exposure, because they are correlated with dietary intake on an individual basis even though quantitative recoveries are not obtained.

LDL has not previously been suggested as a biomarker of fat intake since it is affected by a number of other factors besides diet. Nonetheless, carefully controlled metabolic and cross sectional studies have established that percentage energy from saturated fatty acids is one of the major determinants of serum LDL and total cholesterol. We have shown elsewhere in EPIC Norfolk that associations with plasma LDL and total saturated fat and percentage energy from total saturated fat were similar and we have used the convention of percentage energy from saturated fat here. Investigations of the effect of fibre have previously concentrated on its effect on plasma cholesterol rather than on serum triglyceride levels and there are conflicting results from intervention studies, with some studies showing no effects and others inverse effects of fibre on serum triglycerides. Carbohydrate has not previously been suggested as a marker of triglycerides, but carbohydrate typically increases triglyceride concentrations more than does fat, which generally does not raise serum triglyceride levels. Effects of carbohydrate in this study were, however, weaker than those of fibre. Cross sectional inverse associations between fibre intake and serum triglycerides have been shown in another prospective study, the Framingham study of women. We measured serum lipids in the non-fasting state. However, blood sampling would have had no major effect on our analysis as indicated in a meta-analysis that showed no differences in triacylglycerol-associated IHD risk between non-fasting and fasting participants.

Table 4 Quintile means and HR per quintile increase in variable using Cox regression for vitamin C in 678 cases of IHD (mortality and incident hospital admissions) in 11 134 individuals who completed both a 7 day diary and a FFQ.

<table>
<thead>
<tr>
<th>Item</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma vitamin C mmol/l</td>
<td>28.9</td>
<td>46.3</td>
<td>55.1</td>
<td>63.1</td>
<td>79.9</td>
</tr>
<tr>
<td>Food diary</td>
<td>79</td>
<td>160</td>
<td>226</td>
<td>306</td>
<td>486</td>
</tr>
<tr>
<td>FFQ</td>
<td>200</td>
<td>324</td>
<td>422</td>
<td>536</td>
<td>824</td>
</tr>
</tbody>
</table>

All HR were adjusted for sex, age, smoking, exercise, weight, systolic blood pressure and intake of energy and the energy yielding nutrients alcohol, fat and carbohydrate, and for plasma cholesterol. Subjects reporting use of antihypertensive and cholesterol lowering drugs and self reported change in diet over the past year and reporting diabetic, low salt and low fat diet at baseline were excluded.

Figure 1 Plasma vitamin C and fruit and vegetable intake from different methods and risk of developing IHD in 678 cases in EPIC Norfolk Cohort of 11 134 free of CHD at baseline. HR per quintile are shown.

EPIC NORFOLK COHORT STUDY 983
present study, we found no association between HDL and total fat, but we did find a strong inverse association with carbohydrate. Fat is inversely associated with carbohydrate intake and it is well established in the literature that isoenergetic substitution of saturated, monounsaturated and polyunsaturated fatty acids (but not trans fatty acids) with carbohydrate lowers HDL levels.\textsuperscript{36,40–42} Lower HDL concentrations are also found in both individuals and populations that habitually consume low fat, high carbohydrate diets.\textsuperscript{36,40,42}

Using these comparison markers for exposure to fat, carbohydrate, fibre, vitamin C, potassium, sodium and phytoestrogens, we show that there were strong and consistent associations between biomarkers and dietary intake when assessed by the food diary. However, associations with biomarkers were attenuated when intake was assessed by the FFQ (Tables 1 and 2). Table 3 shows that the associations with fat, fibre and carbohydrate biomarkers remained strong for dietary intake measured by the food diary 4 years earlier but were attenuated when dietary intake was measured 4 years earlier by the FFQ, especially so for carbohydrate and plasma HDL. There were no differences in coefficients of associations between the biomarkers and intakes of vitamin C from different methods using second health check data (Table 3). However, plasma vitamin C had increased between the two health checks, likely due to an increase in dietary intake, as national data show that there has been a substantial (30–35%) increase in dietary intake of vitamin C between 1987–2000.\textsuperscript{43} A change in intake, rather than loss of power from a reduction in sample size by 40%, thus explains why the relationship between intake from the first health check diary and plasma vitamin C at the second health check was lessened. It is noteworthy that the magnitudes of the $b$ coefficients between intake of vitamin C from the FFQ and plasma vitamin C were similar on both occasions, as was the attenuated relation between the intake from the food diary and the changed plasma vitamin C up to 4 years later. This suggests that the association between plasma vitamin C and intake of vitamin C from the FFQ was less specific than the relation between the actual intake measured from the diary and plasma vitamin C at the first health check. Unfortunately no coded second health check diaries are available as yet to confirm this point. Plasma vitamin C is most highly correlated with dietary vitamin C for up to 30 days preceding blood collection ($r = 0.5–0.6$), within the time the FFQ was generally completed and blood was taken at the first health check.\textsuperscript{44} This, taken together with the data in Table 3 suggests there is little evidence that the results shown in the Tables 1 and 2 can be attributed to the fact that the collection of data from the biomarkers was closer in time to the food diary (within 1 week) than that from the FFQ (completed before attending the clinic where biomarkers were collected).

We have previously shown that plasma vitamin C is associated with reduced risk of IHD. Plasma vitamin C is a biomarker of fruits and vegetable consumption and unlikely to be the active protective factor, since supplements of pure vitamin C have had no effect on IHD mortality.\textsuperscript{45} It has been suggested that, as socioeconomic factors are also related to plasma vitamin C, these could account for associations between plasma vitamin C and fruits and vegetables with IHD risk.\textsuperscript{43} However, as data on all exposures were available on the same individuals, it is unlikely that the apparent lack of effect of fruits and vegetables on IHD risk when assessed by FFQ is accounted for by socioeconomic factors, especially as the FFQ is recommended because it is less demanding for participants to complete.\textsuperscript{10} For consistency, we used the same adjusting factors for energy and risk factors when comparing methods with biomarkers, and assessing HR for IHD, throughout.

Table 4 shows that the association between risk of IHD and consumption of fruits and vegetables would not have been evident had a FFQ been used. However, using the food diary, there was a relative risk reduction of 37%, 0.63 (0.48–0.82) in the top quintile of fruit and vegetable intake (equivalent to six portions per day) compared with the bottom quintile. This is of a much greater magnitude than that of the relative risk reduction of only 12% for an increment of five servings daily of 0.88 (95% CI 0.81–0.95) for cardiovascular disease in a large US study of 3634 cardiovascular cases in a cohort of 109 635 individuals in which intake was determined using a FFQ.\textsuperscript{46} In the present study, and in the US study, fruits and vegetables assessed by the FFQ were remarkably high. It is difficult to compare results between populations quantitatively but findings from the FFQs might suggest that the average population health target of five portions or fruit and vegetables (400 g/day) had been surpassed. However, FFQ are well known to systematically overestimate fruits and vegetable intakes\textsuperscript{15} and Table 4 shows that less than three portions per day were consumed on average in the middle quintile when assessed by the food diary, an amount similar to that derived from UK national food survey average data.\textsuperscript{43}

We have been able to assess associations between biomarkers and dietary intake using different methods using a very limited range of available biomarkers, and our findings may differ for other nutrients and foods. For example in a smaller comparison of 4949 individuals in EPIC Norfolk, the correlations between n-3 plasma fatty acids and fatty fish intake assessed by food diary and FFQ were the same,\textsuperscript{47} whereas in our earlier study using recovery biomarkers, intake of protein and potassium as estimated by the food diary was more closely related to 24 h urine biomarkers for protein and potassium ($r = 0.60–0.70$) than when estimated by FFQ ($r = 0.27–0.50$).\textsuperscript{48} In the OPEN study, correlations between the biomarkers
doubly labelled water and 24 h urine nitrogen with energy intake and protein intake assessed by two 24 h recalls were higher \((r=0.24-0.41)\) than when assessed by FFQ \((r=0.10-0.33)\).\(^4\) These and other validation studies investigating the relationship between biomarkers and different methods are, however, generally small, involving <500 subjects,\(^6\) so that investigation of the predictive value of different methods in relation to disease end points is not possible. The present set of data are, however, sufficiently large to compare plasma vitamin C, a biomarker of intake of fruits and vegetables, and intake from different methods of dietary assessment in relation to risk of IHD risk.

Energy adjustment has been recommended to attempt to correct for measurement error, but effects can vary widely depending on the nutrient concerned, with negligible effects in the case of vitamin C, but with greater and unpredictable effects for items with a high error correlation such as fat.\(^32\) In addition, total energy intake is imperfectly measured and we, therefore, adjusted for weight and activity throughout, as these are more closely related to energy expenditure and, in energy balance, energy intake.\(^24\),\(^53\)

Data shown here comparing different dietary methods suggest that those methods most closely related to biomarkers are associated with an improved association between diet and chronic disease, in this case fruit and vegetable intake and IHD, probably due to reduced measurement error and thus improved accuracy. Improved accuracy is a cost effective way of avoiding substantial infrastructure requirements for recruitment and follow up of massive population cohorts in which interaction between gene variants and dietary or other environmental variable exposure is to be assessed.\(^53\) Records, recalls or other ways of documenting real time intake are more expensive to analyse, but written records can be stored for later analysis in nested case control studies, as can biological markers. Self administered CD and web-based methods are under development. The advantage of heterogeneity and calibration from multi-cohort studies in food habits can also be used to improve accuracy, at least for some items, as in the Europe wide EPIC study.\(^54\)

**Acknowledgements**

EPIC Norfolk is funded by the Medical Research Council, Cancer Research UK, Food Standards Agency, World Cancer Research Fund, British Heart Foundation, Yen Ling Low by the Agency for Science, Technology and Research (Centros, Singapore).

**Conflict of interest:** None declared.

---

**KEY MESSAGES**

- In assessing diet in cohort investigations of gene nutrient interactions in chronic disease risk, a simple method, a FFQ, is usually used.
- In this study of 12474 men and women, we used a second method, a food diary, to assess dietary intake.
- Cross sectional associations between dietary intake derived from the food diary and nutritional biomarkers in blood and urine were stronger than those derived from the FFQ.
- Risk of IHD was associated with plasma vitamin C \((P<0.001)\) and intake of vitamin C and fruit and vegetables assessed by food diary \((P\text{ quintile trends }<0.001, 0.001)\) but not by the FFQ \((P\text{ quintile trends }0.923, 0.186)\).

---

**References**

In this issue, Bingham et al. present analyses comparing nutrient intakes estimated by a 1-week diet record (DR) or food frequency questionnaire (FFQ) to a series of biomarkers and incidence of coronary heart disease (CHD). Such triangulation analyses can provide valuable evidence on the relative validity of different dietary assessment methods and are thus of considerable interest. Unfortunately, the results provided by Bingham et al. are difficult to interpret due to the design of the study and their choice of statistical models.

In the design of a study in which multiple dietary assessment methods are compared with biomarkers, the temporal relationships are critical. Underlying this is the concept that, for epidemiologic applications, the dietary methods are meant to represent intake over an extended period, usually many months or years, not just the few days or weeks around the time the data are collected. Thus, if the biomarker is sensitive to recent intake (true for most urinary and plasma measurements, but less true for biomarkers using adipose or nail tissue) and is collected close in time to a measure of short-term intake, such as a DR or several 24-h recalls, the correlations will tend to be exaggerated. This is a form of correlated error that can result from fluctuations in diet due to seasonality or random variation that will be reflected in both measurements (see Chapter 6). These issues are different for a FFQ, which typically asks about diet over the past year, and should therefore reflect dietary intake over a much longer period, not just recent diet. In the cross-sectional analyses by Bingham et al., the FFQ, DR and biomarkers were all collected in close temporal proximity, which would tend to bias the correlations in favour of the DR. Their argument in the discussion that the FFQ was collected within 30 days of the biomarkers misses the critical point. The solution to this bias is to collect the biomarker at some time remote from the collection of the dietary intakes assessments, either before or later. Fortunately, Bingham et al. have done this. In an earlier, although smaller, validation study, plasma vitamin C...