METHODOLOGY

The evaluation of the diet/disease relation in the EPIC study: considerations for the calibration and the disease models

Pietro Ferrari, 1* Nicholas E Day, 2 Hendriek C Bosshuizen, 3 Andrew Roddam, 4 Kurt Hoffmann, 5 Anne Thiebaut, 6 Guillem Pera, 7 Kim Overvad, 8 Ellyv Lund, 9 Antonia Trichopoulou, 10 Rosario Tumino, 11 Bo Gullberg, 12 Teresa Norat, 13 Nadia Slimani, 14 Rudolf Kaaks 14 and Elio Riboli 13

Accepted 30 October 2007

Background International multicentre studies on diet and cancer are relatively new in epidemiological research. They offer a series of challenging methodological issues for the evaluation of the association between dietary exposure and disease outcomes, which can both be quite heterogeneous across different geographical regions. This requires considerable work to standardize dietary measurements at the food and the nutrient levels.

Methods Within the European Prospective Investigation into Cancer and Nutrition (EPIC), a calibration study was set up to express individual dietary intakes according to the same reference scale. A linear regression calibration model was used to correct the association between diet and disease for measurement errors in dietary exposures. In the present work, we describe an approach for analysing the EPIC data, using as an example the evaluation of the association between fish intake and colorectal cancer incidence.

Results Sex- and country-specific attenuation factors ranged from 0.083 to 0.784, with values overall higher for men compared with women. Hazard ratio estimates of colorectal cancer for a 10 g/day increase in fish intake were 0.97 [95% confidence interval (CI): 0.95–0.99] and 0.93 (0.88–0.98), before and after calibration, respectively.

Conclusions In a multicentre study, the diet/disease association can be evaluated by exploiting the whole variability of intake over the entire study. Calibration may reduce between-centre heterogeneity in the diet–disease relationship caused by differential impact of measurement errors across cohorts.

Keywords Calibration, multicentre study, measurement error, diet, EPIC

* Corresponding author. IARC-WHO, 150 cours Albert-Thomas, 69372 Lyon cedex 08, France. E-mail: Ferrari@iarc.fr

1 International Agency for Research on Cancer (IARC-WHO), Lyon, France.
2 Strangeways Research Laboratory, University of Cambridge, UK.
3 National Institute of Public Health and the Environment (RIVM), Bilthoven, Netherlands.
4 Cancer Epidemiology Unit, University of Oxford, Headington, Oxford, UK.
5 German Institute of Human Nutrition, Postdam-Rehbrücke, Germany.
6 INSERM (Institut National de la Santé et de la Recherche Médicale), ERI-20, Institut Gustave Roussy, Villejuif, France.
7 Catalan Institute of Oncology, Barcelona, Spain.
8 Department of Clinical Epidemiology, Aalborg Hospital, Aarhus University Hospital, Aalborg, Denmark.
9 Corresponding author. IARC-WHO, 150 cours Albert-Thomas, 69372 Lyon cedex 08, France. E-mail: Ferrari@iarc.fr

368
Introduction

Over the past 30 years, much epidemiological research has focused on the relationship between diet-related factors and chronic diseases, particularly cancer and cardiovascular diseases. For this purpose, several study designs have been employed using different types of information, such as ecological correlations, migrant and time-trend studies and case-control and cohort studies.

Analytical epidemiology studies aimed at investigating the relationship between diet and chronic diseases face important challenges, very often related to diet ascertainment and confounding bias. Moreover, study populations often show limited variations in nutrient intake. In addition, diet is a complex mixture of foods and nutrients with many correlated elements, very difficult to measure with sufficient accuracy and precision, in order to distinguish the role of different foods and nutrients in epidemiological investigations.

Random and systematic measurement errors in individual dietary assessment may bias estimates of diet–disease associations. These effects can be reduced by increasing the overall heterogeneity of the dietary exposures studied, thus limiting the proportion of variability attributable to random error. This was the rationale behind the European Prospective Investigation into Cancer and Nutrition (EPIC), a multicentre cohort study on diet and cancer consisting of 23 regional centres in 10 Western European countries with widely varying dietary practices and cancer incidence rates.

In the EPIC study, information on habitual individual dietary intake were assessed using different validated dietary questionnaires (Q) across participating countries to capture geographical specificity of diet. For proper comparison of dietary measurements between cohorts, a calibration sub-study was established for the first time in a large, multicentre European study. This approach involved an additional dietary assessment common across study populations to re-express individual dietary intakes according to the same reference scale. For this, a single 24 h dietary recall (24-HDR), (R) was collected, as the EPIC reference calibration method, from a stratified random sample of 36 900 subjects from the entire EPIC cohort, using a software program (EPIC-SOFT) specifically designed to standardize the dietary measurements across study populations. A linear regression calibration model was used to adjust for possible systematic over- or under-estimation in dietary intake measurements and correct for attenuation bias in relative risk estimates.

In this work, some issues related to the statistical analyses of these complex data are presented and discussed. In particular, we address how the between- and within-cohort components of the calibration model were used for measurement error correction at the individual and at the ecological level. Furthermore, we address aspects on the evaluation of the diet/disease relationship in the risk model, to exploit fully the overall variability of exposure in the EPIC study, and to evaluate heterogeneity of associations across cohorts. The evaluation of the relationship between diet and cancer incidence in the EPIC study is illustrated with an example of the assessment of the association between fish intake and colorectal cancer risk. The variability of fish intake within- and between-EPIC cohorts, before and after calibration, is evaluated.

Material and Methods

EPIC included 366 521 women and 153 457 men, mostly aged 35–70 years, recruited in 23 centres located in 10 European countries. After exclusion of 22 432 prevalent cases of any cancer other than non-melanoma skin cancer at the date of enrolment, 10 208 subjects in the top and bottom 1% of the distribution of the ratio of reported total energy intake to energy requirement were excluded from the analysis to reduce the impact of outlier values. In addition, subjects with missing questionnaire data or missing dates of diagnosis or follow up, representing 2% of the participants were excluded. A total of 478 040 subjects were included in the analysis. Details on the follow-up of study subjects included in the analyses are described by Riboli et al.

The calibration model

A calibration model was used to correct for errors in dietary measurements, the relation between food groups and/or nutrients and cancer incidence. Given the multicentre nature of EPIC, the calibration model should reflect two levels of evidence, by making dietary exposures comparable across centres (between-group calibration) and correcting the diet/disease relationship for measurement errors (within-group calibration).

For study subject \(i = 1, \ldots, n_j\) in group \(j = 1, \ldots, J\), be it centre or country, let \(T_{ij} = (T_{ij,1}, \ldots, T_{ij,K_j})\) denote...
For the $k_i$-th error-prone variables, $k_i = 1, \ldots, K_i$, it is assumed that the reference measurements $\mathbf{R}_{ij,k_i} = T_{ij,k_i} + \varepsilon_{R_{ij,k_i}}$, which requires that the errors $\varepsilon_{R_{ij,k_i}}$ are strictly random and that variation around individuals’ unknown true intake is totally due to within-person random variability or to random measurement errors in reporting individuals’ diet, i.e. $\text{Cov}(\varepsilon_{R_{ij,k_i}}, T_{ij,k_i} | \text{group} = j) = 0$. On the other hand, the corresponding questionnaire measurements $Q_{ij,k_i}$ are assumed to contain random and systematic within-person measurement errors.

In line with the calibration methodology introduced by Rosner et al., the reference measurements were linearly regressed in a multivariate calibration model on the dietary questionnaire measurements to estimate the factors that quantify the impact of bias on observed (naive) risk estimates. A very general form of the EPIC multivariate calibration model can be defined, for each group $j$, as

$$\mathbf{R}_{ij} = \alpha_j + Q_{ij} \boldsymbol{A}_j + \mathbf{Z}_{ij} \boldsymbol{\Gamma}_j + \varepsilon_{ij}$$

(1)

The terms $\varepsilon_{ij}$ capture the residual random variability within each group and are assumed to be multivariate and $\varepsilon_{ij}$ i.i.d. $(0, \mathbf{\Sigma}_{ij})$. The group-specific $K_2 \times K_1$ coefficient matrices $\boldsymbol{\Gamma}_j$ model the effect of confounder variables $\mathbf{Z}_{ij}$, consistently with covariates included in the disease model. The columns of the group-specific matrices $\mathbf{A}_j = (k_1, \ldots, k_{K_i})_j$ are $K_i \times 1$ vectors of coefficients that relate the $k_1$-th reference measurements $\mathbf{R}_{ij,k_1}$ to the vector of dietary questionnaire variables $Q_{ij}$ given the error-free confounding variables $\mathbf{Z}_{ij}$.

It is worth noting that in situations where more than one variable is measured with errors ($K_i > 1$), the level of bias of the $k_1$-th parameter that quantifies the diet/disease association, has a multiplicative factor, referred to as attenuation factor (elements in the main diagonal of the matrix $\mathbf{A}_j$), and an additive component referred to as contamination factors (CFs) (off-diagonal elements, in the $k_1$-th column of $\mathbf{A}_j$), due to residual confounding by the other ($K_i - 1$) variables measured with error in the vector $Q_{ij}$ included in model (1). The way relative risk estimates are affected by attenuation and CFs is outlined in Appendix 1.

The regression calibration allows, for each variable measured with errors, the imputation of predicted values, $E(\mathbf{T}_{ij} | Q, Z)$, for the entire cohort, indicating average true intake levels for given questionnaire measurements which are used as predictors in a disease model. In model (1), the terms $\varepsilon_{ij}$ are $1 \times K_i$ group-specific intercept terms. In this way, the between-group component of the calibration model is taken into account, as $E(\mathbf{T}_{ij} | Q, Z)$ values are centred on group-specific $R$ means, thus correcting for group-specific biases, while the terms in $\mathbf{A}_j$ capture the within-group component in the calibration model.

Empirical evidence suggested the use of gender-specific calibration models to take into account gender heterogeneity in the accuracy of $Q$ measurements. In the current study, a multivariate calibration model was used for fish, energy from fat sources and energy from non-fat sources. Calibration was not performed for physical activity values because no reference measurements are available in EPIC for this variable. The sampling distributions of days (Monday–Thursday vs Friday–Sunday) and seasons of 24-HDR administration was corrected using a set of weights in (1) to reproduce an even distribution of recalls across weekday and season.

A general form of the calibration model is described in (1). In this study, country-specific calibration models, rather than centre-specific, were implemented. This allowed greater stability for parameter estimates in (1) to be reached, particularly in those centres with relatively small sample size. Country-specific models were adjusted by centre, to take into account the geographical specificity of dietary variables.

Analyses performed at the country-level showed that residuals were rather homogeneous over the range of intake of fish, energy from fat and energy from non-fat sources, respectively. In some countries, however, the distribution of residuals of fish intake in model (1) was skewed, with a notable frequency of zero values in $R$ measurements. An approach to take into account a heteroskedastic structure of data in the calibration model (1) was employed, but results were very similar to the ordinary approach.

### Variability of exposure

In order to evaluate the variability of dietary variables, a mixed model was used to estimate the between- and the within-centre variance components of dietary exposures, which reflect variability at the individual and at the aggregate level, respectively. Variance components of dietary exposures were estimated using the following mixed model

$$Y_{ij} = \alpha_E + \varepsilon_{E_{ij}}$$

$$\alpha_E = \alpha_E + u_{E_{ij}}, \quad u_{E_{ij}} \sim N(0, \sigma^2_u), \quad \varepsilon_{E_{ij}} \sim N(0, \sigma^2)$$

(2)

with, $Y_{ij}$ being each element of vectors $Q_{ij}$, $R_{ij}$ as well as the predicted values estimated in (1). The mixed model enables the estimation of the proportion of total variation in the exposure attributable to the between-centre component, expressed by intra-class correlation coefficient (ICC) $= \sigma^2_u / (\sigma^2 + \sigma^2)$. Parameter estimates were obtained by restricted maximum likelihood estimation.


While the two components of energy intake approximated a normal distribution, the distribution of fish intake was rather skewed, with a varying amount of zero-values across centres, particularly for \( R \) measurements. Variance components are fairly robust to departures from normality in mixed models.\(^\text{14} \) However, although analyses performed after log-transformation of dietary variables or using positive \( Q \) measurements only showed that variance components and ICC estimates were similar to estimates obtained using untransformed variables, these quantities should be interpreted cautiously.

All analyses were run using the SAS Statistical Software.\(^\text{15} \) For all analyses, \( P \)-values <0.05 were considered statistically significant.

The disease model
The association between fish intake and colorectal cancer risk was evaluated using Cox proportional hazards regression\(^\text{16} \) with age as the primary time variable.\(^\text{17,18} \) The analyses were stratified by centre to control for differences in follow-up procedures, questionnaire design and other centre effects. These analyses included 2,279,075 person-years in 4.8 years of follow-up on average since 1992, and 1,329 colorectal cancer cases with complete and satisfactory data from a total of 15 centres and 10 countries, as detailed by Norat et al.\(^\text{19} \)

If the overall exposure distribution allows it, the use of pre-defined cut-offs is advisable, in order to better model the expected dose–response relationship over the entire range of exposure in EPIC. In this study, pre-defined cut-points for fish intake (0–20/20–40/40–80/80–120 g/day) were used. In order to perform isocaloric comparisons and to partly control for the error in the estimated intake of fish, energy intake was included to reduce the effect of error correlation between the different dietary components.\(^\text{20,21} \) Estimated energy intake was divided into energy from fat and energy from non-fat sources, since it is mostly the fat components of diet that contribute to fish intake.\(^\text{19,22} \) In addition, to improve the performance of these isocaloric comparisons, weight was consistently included in the model.\(^\text{23} \)

Height, smoking (never, former and current smoker) and job-related physical activity (not employed, sedentary occupation, standing occupation, manual and heavy manual work activities) were also incorporated in the model. Furthermore, to explore potential differences by sex, gender-specific as well as overall disease models were used.

To test hazard ratios (HRs) for overall significance, the differences in log-likelihood between models with and without the set of indicator variables used to model the association between categories of exposure and disease risk was compared with a \( \chi^2 \)-distribution, with degrees of freedom equal to the number of indicators minus one. Similarly, tests of trend were computed using continuous variables, according to a \( \chi^2 \)-distribution with one degree of freedom. Tests of heterogeneity across countries of associations between each dietary factor and colorectal cancer risk were performed using indicator variables to model the effect of centres and interaction terms between dietary factors and country-specific indicator variables. \( P \)-values were derived after comparison of the log-likelihood difference between models with and without interaction terms with a \( \chi^2 \)-distribution with seven and nine degrees of freedom in men and women, respectively. Similarly, tests for heterogeneity were performed using predicted values of intake after linear calibration.

After correction for measurement errors, the variability of parameter estimates in the disease model should take into account the uncertainty stemming from the regression calibration model. Rosner et al.\(^\text{3} \) developed an analytical formula to obtain corrected standard errors (SE) for the parameter estimates using a Taylor series expansion. In a multicentre design, the use of group-specific calibration models produces a set of group-specific matrices \( A \). In the EPIC study, instead of using a meta-analytical approach, thus pooling centre-specific measurement error corrected \( \log(HR) \) estimates,\(^\text{24} \) and applying a multivariate Taylor series expansion to correct the SE of pooled estimates,\(^\text{11} \) a bootstrap sampling procedure was used\(^\text{25,26} \) to compute corrected SE for the overall calibrated parameter estimates.

Sampling in the calibration model is performed to estimate the coefficients in the calibration model and obtain predicted values to estimate the corrected association for dietary exposure, \( \hat{\beta} \). This procedure yields a set of parameter estimates, \( \hat{\beta}_m, m = 1, \ldots, M \), where \( M \) is the number of bootstrap samplings. The variance of the corrected parameter can be approximated by

\[
\hat{V}(\hat{\beta}) \approx \frac{1}{M} \sum_{m=1}^{M} \hat{V}(\hat{\beta}_m) + \frac{1}{M} \sum_{m=1}^{M} \left( \hat{\beta}_m - \frac{1}{M} \sum_{m=1}^{M} \hat{\beta}_m \right)^2
\]

where \( \hat{V}(\hat{\beta}_m) \) is the variance of \( \hat{\beta}_m \) using the \( m \)-th draw. A total of 300 repetitions were used, and results appeared to be virtually unchanged using larger values of \( M \), thus providing a sufficiently stable estimate of the SE of the corrected regression coefficient. The test for trend for the de-attenuated estimate was calculated referring to a standard normal distribution for the de-attenuated estimate divided by the bootstrap-derived SE.

Although measurement error correction is expected to inflate the variability of the parameter that quantifies the diet/disease relationship, in two cases in this study the statistical significance of the calibrated parameter was higher than the uncorrected HR estimate (e.g. fish intake in men). This might be attributable to the fact that in a multicentre study, if the regression parameters in the calibration model are estimated with great precision, calibration can lead to a reduction of the between-centre heterogeneity in
the overall diet–disease relationship caused by differential impact of measurement errors, thus leading to a relatively more precise estimate of the overall relative risk, as illustrated in Appendix 2.

To assess heterogeneity of de-attenuated associations across countries, the uncertainty from the calibration model was not taken into account as such uncertainty [reflected in second part of expression (3)] was minimal.

Results

In Table 1, attenuation factors estimated in gender- and country-specific calibration models are presented for fish intake. Attenuation factors were low in women in the UK (0.083) and Norway (0.212), as well as in Sweden (0.236 in men and 0.191 in women). Similar values were observed in men and women, with the exception of the UK (0.518 vs 0.083 for men and women, respectively) and Italy (0.531 vs 0.297). In the same table, CFs, which are the coefficients in model (1) relating reference measurements of energy from fat and energy from non-fat sources to Q measurements of fish intake were generally negative, and ranged substantially across countries. To interpret accurately the CFs, it should be noted that their absolute values reflect the difference in scale between energy and fish intakes.

The square root of variance components of the dietary variables, as well as ICC values, are shown in Table 2. For fish intake, the amount of variability explained by between-centre differences is very similar in the two genders, and lower in R compared with Q measurements. This is mostly due to the large within-person day-to-day variability of fish intake in the single R measurement available per subject.

More pronounced increases in ICC values are

<table>
<thead>
<tr>
<th>Country</th>
<th>Fish AF</th>
<th>Fish SE</th>
<th>Energy from fat CF</th>
<th>Energy from fat SE</th>
<th>Energy from non-fat sources CF</th>
<th>Energy from non-fat sources SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>0.531</td>
<td>0.086</td>
<td>1.288</td>
<td>0.431</td>
<td>0.878</td>
<td>0.666</td>
</tr>
<tr>
<td>Spain</td>
<td>0.571</td>
<td>0.064</td>
<td>0.051</td>
<td>0.226</td>
<td>0.041</td>
<td>0.269</td>
</tr>
<tr>
<td>UK</td>
<td>0.518</td>
<td>0.122</td>
<td>1.525</td>
<td>0.731</td>
<td>0.826</td>
<td>0.867</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.675</td>
<td>0.138</td>
<td>2.884</td>
<td>1.151</td>
<td>1.066</td>
<td>1.408</td>
</tr>
<tr>
<td>Greece</td>
<td>0.784</td>
<td>0.188</td>
<td>1.342</td>
<td>0.631</td>
<td>0.843</td>
<td>0.850</td>
</tr>
<tr>
<td>Germany</td>
<td>0.321</td>
<td>0.044</td>
<td>0.131</td>
<td>0.331</td>
<td>0.509</td>
<td>0.367</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.236</td>
<td>0.051</td>
<td>0.086</td>
<td>0.297</td>
<td>0.028</td>
<td>0.306</td>
</tr>
<tr>
<td>Denmark</td>
<td>0.521</td>
<td>0.061</td>
<td>2.149</td>
<td>0.340</td>
<td>0.574</td>
<td>0.426</td>
</tr>
<tr>
<td>Norway</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>0.381</td>
<td>0.039</td>
<td>0.506</td>
<td>0.178</td>
<td>0.567</td>
<td>0.199</td>
</tr>
<tr>
<td>Italy</td>
<td>0.297</td>
<td>0.046</td>
<td>0.115</td>
<td>0.231</td>
<td>0.790</td>
<td>0.341</td>
</tr>
<tr>
<td>Spain</td>
<td>0.449</td>
<td>0.068</td>
<td>0.520</td>
<td>0.282</td>
<td>0.249</td>
<td>0.289</td>
</tr>
<tr>
<td>UK</td>
<td>0.083</td>
<td>0.049</td>
<td>0.412</td>
<td>0.266</td>
<td>0.634</td>
<td>0.322</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.542</td>
<td>0.067</td>
<td>1.789</td>
<td>0.462</td>
<td>1.664</td>
<td>0.539</td>
</tr>
<tr>
<td>Greece</td>
<td>0.628</td>
<td>0.151</td>
<td>1.228</td>
<td>0.649</td>
<td>0.262</td>
<td>0.715</td>
</tr>
<tr>
<td>Germany</td>
<td>0.491</td>
<td>0.056</td>
<td>1.354</td>
<td>0.327</td>
<td>1.061</td>
<td>0.395</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.191</td>
<td>0.050</td>
<td>0.866</td>
<td>0.277</td>
<td>0.321</td>
<td>0.294</td>
</tr>
<tr>
<td>Denmark</td>
<td>0.423</td>
<td>0.063</td>
<td>1.240</td>
<td>0.304</td>
<td>0.317</td>
<td>0.426</td>
</tr>
<tr>
<td>Norway</td>
<td>0.212</td>
<td>0.042</td>
<td>0.572</td>
<td>0.152</td>
<td>0.714</td>
<td>0.190</td>
</tr>
</tbody>
</table>

*The multivariate calibration model was gender- and country-specific. Adjustment for age (years), height and weight, smoking status, physical activity at work and centre was performed. A set of weights was employed to model the effect of ‘day of recall’ and ‘season of recall’ in R measurements.
observed in $E(T|Q,Z)$ compared with $Q$ measurements for energy from non-fat sources compared with energy from fat.

In Table 3, the number of colorectal cancer cases, of person-years (PY) and multivariable hazard ratios (HR) (95% confidence intervals) of the association between fish intake and risk of colorectal cancer in EPIC, by categories of fish intake.

The corrected SEs of calibrated parameters estimated with a bootstrap sampling approach did not show appreciable variation compared with uncorrected SEs. Linear calibration reduced the heterogeneity of associations across countries in men. Overall, after linear calibration, for men and women combined, the relation between fish and colorectal cancer was statistically not heterogeneous in the EPIC countries.

Discussion

Multicentre epidemiological studies on diet and chronic diseases are relatively new and they offer a series of challenging methodological issues on the standardization of dietary measurements at the food and nutrient levels. In the EPIC study, much effort has been invested to make dietary exposures comparable across centres. This issue was addressed with a calibration study and the preparation of a common

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measurement</th>
<th>Men Between</th>
<th>Within</th>
<th>ICC</th>
<th>Men Between</th>
<th>Within</th>
<th>ICC</th>
<th>Women Between</th>
<th>Within</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish</td>
<td>Q</td>
<td>19</td>
<td>30</td>
<td>0.285</td>
<td>20</td>
<td>30</td>
<td>0.308</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>23</td>
<td>82</td>
<td>0.070</td>
<td>14</td>
<td>63</td>
<td>0.046</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$E(T</td>
<td>Q,Z)$</td>
<td>22</td>
<td>15</td>
<td>0.670</td>
<td>14</td>
<td>9</td>
<td>0.694</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy from fat</td>
<td>Q</td>
<td>110</td>
<td>290</td>
<td>0.127</td>
<td>90</td>
<td>230</td>
<td>0.132</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>100</td>
<td>435</td>
<td>0.050</td>
<td>54</td>
<td>321</td>
<td>0.028</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$E(T</td>
<td>Q,Z)$</td>
<td>105</td>
<td>156</td>
<td>0.310</td>
<td>58</td>
<td>95</td>
<td>0.272</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy from non-fat sources</td>
<td>Q</td>
<td>171</td>
<td>428</td>
<td>0.138</td>
<td>131</td>
<td>337</td>
<td>0.131</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>177</td>
<td>533</td>
<td>0.099</td>
<td>110</td>
<td>389</td>
<td>0.075</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$E(T</td>
<td>Q,Z)$</td>
<td>163</td>
<td>193</td>
<td>0.419</td>
<td>111</td>
<td>134</td>
<td>0.410</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 Frequency of colorectal cancer cases, person-years (PY) and multivariable hazard ratios (HR) (95% confidence intervals) of the association between fish intake and risk of colorectal cancer in EPIC, by categories of fish intake.

Table 2 Square root of between- and within-centre variance components, and ICC values of fish intake (g/day), energy from fat and energy from non-fat sources (Kcal/day) from questionnaire (Q), 24-HDR (R) measurements and predicted values ($E(T|Q,Z)$), estimated in model (2).

<table>
<thead>
<tr>
<th>Exposure Categories (g/day)</th>
<th>Men Cases/PY</th>
<th>Men Observed HR</th>
<th>Women Cases/PY</th>
<th>Women Observed HR</th>
<th>ALL Cases/PY</th>
<th>ALL Observed HR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–20</td>
<td>98/130612</td>
<td>1</td>
<td>149/314850</td>
<td>1</td>
<td>247/445463</td>
<td>1</td>
</tr>
<tr>
<td>20–40</td>
<td>75/116758</td>
<td>0.70 (0.52–0.95)</td>
<td>158/301953</td>
<td>1.00 (0.80–1.26)</td>
<td>233/418711</td>
<td>0.89 (0.74–1.06)</td>
</tr>
<tr>
<td>40–80</td>
<td>184/193676</td>
<td>0.83 (0.64–1.08)</td>
<td>244/486012</td>
<td>0.86 (0.68–1.08)</td>
<td>428/679688</td>
<td>0.86 (0.72–1.02)</td>
</tr>
<tr>
<td>80–120</td>
<td>125/156518</td>
<td>0.61 (0.45–0.82)</td>
<td>178/379961</td>
<td>0.71 (0.55–0.91)</td>
<td>303/536479</td>
<td>0.67 (0.55–0.81)</td>
</tr>
<tr>
<td>&gt;120</td>
<td>60/75485</td>
<td>0.64 (0.44–0.92)</td>
<td>58/127332</td>
<td>0.70 (0.50–0.99)</td>
<td>118/202818</td>
<td>0.67 (0.53–0.87)</td>
</tr>
</tbody>
</table>

$P$-value$^b$ – 0.007 0.021 0.001

$^a$In the Cox models, age was used as primary time variable, and adjustment for energy from fat, energy from non-fat sources, height, weight, physical activity at work (not employed, sedentary occupation, standing occupation, manual and heavy manual work activities), smoking status (never, former, current smoker) was performed. Models were stratified by centre and gender (in combined analyses only).

$^b$P-value for overall significance, according to chi-squared distribution with four degrees of freedom (dof).
Table 4 Multivariable-standardized observed and measurement error corrected HR (95% confidence intervals) of the association between fish intake, energy from fat and energy from non-fat sources, and colorectal cancer risk in EPIC, using exposure on a continuous scale.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Men</th>
<th>Men</th>
<th>Women</th>
<th>Women</th>
<th>ALL</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Linear calibrated</td>
<td>Observed</td>
<td>Linear calibrated</td>
<td>Observed</td>
<td>Linear calibrated</td>
</tr>
<tr>
<td>Fish Intake</td>
<td>Estimate</td>
<td>-0.0030</td>
<td>-0.0065</td>
<td>-0.0035</td>
<td>-0.0096</td>
<td>-0.0033</td>
</tr>
<tr>
<td>HR (10 g/day)</td>
<td>0.970</td>
<td>0.937</td>
<td>0.966</td>
<td>0.908</td>
<td>0.968</td>
<td>0.960</td>
</tr>
<tr>
<td>SE\textsuperscript{b}</td>
<td>0.0015</td>
<td>0.0031</td>
<td>0.0015</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.0011</td>
</tr>
<tr>
<td>(P)-value\textsuperscript{c}</td>
<td>0.0455</td>
<td>0.0360</td>
<td>0.0196</td>
<td>0.0329</td>
<td>0.0018</td>
<td>0.0050</td>
</tr>
<tr>
<td>Corrected SE\textsuperscript{d}</td>
<td>-</td>
<td>0.0032</td>
<td>-</td>
<td>0.0045</td>
<td>-</td>
<td>0.0026</td>
</tr>
<tr>
<td>(P)-value\textsuperscript{e}</td>
<td>-</td>
<td>0.0422</td>
<td>-</td>
<td>0.0329</td>
<td>-</td>
<td>0.0050</td>
</tr>
<tr>
<td>(P)-value\textsuperscript{f}</td>
<td>0.0696</td>
<td>0.2510</td>
<td>0.5229</td>
<td>0.5606</td>
<td>0.0957</td>
<td>0.2449</td>
</tr>
<tr>
<td>Energy from fat</td>
<td>Estimate\textsuperscript{g}</td>
<td>-0.0049</td>
<td>-0.0066</td>
<td>0.0025</td>
<td>0.0022</td>
<td>-0.0012</td>
</tr>
<tr>
<td>HR (100 kcal/day)</td>
<td>0.952</td>
<td>0.936</td>
<td>1.025</td>
<td>1.022</td>
<td>0.988</td>
<td>0.965</td>
</tr>
<tr>
<td>SE\textsuperscript{b,g}</td>
<td>0.0020</td>
<td>0.0035</td>
<td>0.0020</td>
<td>0.0044</td>
<td>0.0014</td>
<td>0.0028</td>
</tr>
<tr>
<td>(P)-value\textsuperscript{c}</td>
<td>0.0172</td>
<td>0.0610</td>
<td>0.2208</td>
<td>0.6221</td>
<td>0.3975</td>
<td>0.2024</td>
</tr>
<tr>
<td>Corrected SE\textsuperscript{d,g}</td>
<td>-</td>
<td>0.0036</td>
<td>-</td>
<td>0.0045</td>
<td>-</td>
<td>0.0029</td>
</tr>
<tr>
<td>(P)-value\textsuperscript{e}</td>
<td>-</td>
<td>0.0652</td>
<td>-</td>
<td>0.6260</td>
<td>-</td>
<td>0.2260</td>
</tr>
<tr>
<td>(P)-value\textsuperscript{f}</td>
<td>0.5446</td>
<td>0.5771</td>
<td>0.7080</td>
<td>0.3448</td>
<td>0.3600</td>
<td>0.4010</td>
</tr>
<tr>
<td>Energy from non-fat sources</td>
<td>Estimate\textsuperscript{g}</td>
<td>0.0015</td>
<td>0.0010</td>
<td>0.0007</td>
<td>0.0021</td>
<td>0.0011</td>
</tr>
<tr>
<td>HR (100 kcal/day)</td>
<td>1.015</td>
<td>1.011</td>
<td>1.007</td>
<td>1.022</td>
<td>1.011</td>
<td>1.022</td>
</tr>
<tr>
<td>SE\textsuperscript{b,g}</td>
<td>0.0013</td>
<td>0.0030</td>
<td>0.0014</td>
<td>0.0034</td>
<td>0.0014</td>
<td>0.0022</td>
</tr>
<tr>
<td>(P)-value\textsuperscript{c}</td>
<td>0.2639</td>
<td>0.7252</td>
<td>0.6102</td>
<td>0.5327</td>
<td>0.2488</td>
<td>0.3370</td>
</tr>
<tr>
<td>Corrected SE\textsuperscript{d,g}</td>
<td>-</td>
<td>0.0032</td>
<td>-</td>
<td>0.0035</td>
<td>-</td>
<td>0.0023</td>
</tr>
<tr>
<td>(P)-value\textsuperscript{e}</td>
<td>-</td>
<td>0.7435</td>
<td>-</td>
<td>0.5426</td>
<td>-</td>
<td>0.3521</td>
</tr>
<tr>
<td>(P)-value\textsuperscript{f}</td>
<td>0.7228</td>
<td>0.3520</td>
<td>0.6076</td>
<td>0.6855</td>
<td>0.5212</td>
<td>0.5301</td>
</tr>
</tbody>
</table>

\textsuperscript{a}In the Cox models, age was used as primary time variable, and adjustment for energy from fat, energy from non-fat sources, height, weight, physical activity at work (not employed, sedentary occupation, standing occupation, manual and heavy manual work activities), smoking status (never, former, current smoker) was performed. Models were stratified by centre and gender (in combined analyses only).

\textsuperscript{b}Calculated SE in the disease model.

\textsuperscript{c}\(P\)-value for significance of log(HR), according to a chi-squared distribution with 1 dof.

\textsuperscript{d}Corrected SE after bootstrap sampling.

\textsuperscript{e}\(P\)-value for significance of log(HR) using corrected SE, according to a chi-squared distribution with 1 dof.

\textsuperscript{f}\(P\)-value for heterogeneity across countries, according to a chi-squared distribution with 7 and 9 dof, for men and women, respectively.

\textsuperscript{g}(x10).
departures from linearity between dietary exposure and disease should be consistently evaluated. For observed $Q$ measurements, this might be achieved with the inclusion of an indicator variable ($0 = \text{non-consumer}/1 = \text{consumer}$) in the model, which allows the diet/disease relationship to be evaluated among consumers only. Similarly, the same indicator can be used in the linear regression calibration to evaluate possible departure from linearity if the indicator is statistically significant, and possibly use calibration models using non-zero $Q$ values only. In the current study, although the percentage of non-consumers of fish intake according to $Q$ measurements ranged from 0.9 (Denmark) to 24.5% (UK), no evidence of departure from linearity has been observed in the calibration and in the disease model.

The regression calibration approach has a theoretical justification based on a linear approximation, provided that the errors in $R$ and $Q$ measurements are uncorrelated. Correlation between errors in self-reported dietary measurements is likely to exist as a consequence of study subjects’ tendency to consistently under- or over-estimate dietary intake.9,29–31 Attenuation factors are overestimated in the case of correlation between errors in $R$ and $Q$ measurements,22,32 at least when the univariate case is considered, i.e. the case of calibrating the relationship of one dietary exposure and disease risk.31 As a consequence, the linear calibration will result in a conservative correction.

In the case of multiple exposure measured with errors, high levels of error correlations between questionnaire measurements can lead to bias of regression estimates in multivariable calibration models. Although it has been suggested that energy adjustment, either by energy density or by direct adjustment, can reduce the effect of correlation between random measurement errors,21,33,34 it should be noted that these considerations apply in situations where the correlation between errors in $Q$ and $R$ measurements is not too sizeable, and when truly unbiased reference measurements are available. Both conditions are likely to challenge a correct evaluation of many diet–disease associations in nutritional epidemiology, including the one investigated in this work. In the present analysis, energy adjustment was performed in the multivariate calibration model. Although the observed correlation coefficients between fish and energy from fat and energy from non-fat sources do not reflect the true association between these variables, because of measurement errors in dietary exposure, as well as of correlation between errors, our findings suggest that error correlations between fish and energy intake are not likely to be very high.

In general, energy adjustment seems to be particularly effective in the case of macronutrient intakes,8,20,21 for which higher correlations between errors are expected. Even after energy adjustment, error correlation in $Q$ and $R$ measurements cannot be ruled out, and findings on the aetiological role of dietary exposure factors on disease outcome need an extremely cautious interpretation.

Linear calibration assumes that $R$ measurements provide an unbiased estimate of true exposure. The paper by Ferrari et al.31 showed under reporting in EPIC $R$ measurements, thus suggesting the presence of some systematic errors. Under the assumption that under reporting is either randomly distributed across subjects within groups or is of the same magnitude across groups, $R$ measurements could still provide a reference scale without absolute validity but common across subjects within group (at the individual level), and for between-group comparison (at the aggregate level).35 However, results from the OPEN validation study31 showed that 24-HDR measurements contain, at least for total energy and protein intakes, intake-related bias, which occurs when the tendency to under or overestimate is different for subjects with a high intake than for subjects with a low intake. Intake-related bias is not randomly distributed and it is likely to be related to study subjects’ characteristics, thus providing evidence on the limitations of the use of 24-HDRs as reference measurements.

The EPIC calibration model was adjusted for weight, age, and a set of weights was used to model the effect of week day of 24-HDR, as well as the season of the year in which the 24-HDR was administered. Although this adjustment was performed consistently in order to partially control the systematic component of measurement errors,36 it cannot be excluded that systematic error is still present and bias individual $R$ measurements.

In the EPIC study, fish intake showed a significant protective effect for colorectal cancer.19 The effect of dietary exposure on cancer incidence was assessed by modelling the effect of these factors as continuous variables, or by using categories of exposure, to capture dose–response effects other than linear. In a multicentre study, the use of EPIC-wide categories of exposure ensures that the comparisons are made over the variability of intake of the entire EPIC cohort, although adjusting for centre would limit the evaluation of the diet/disease relationship within each group. In some cases, the distribution of intakes could vary widely by centre, so the within-centre distribution of cases by category of intake can be very heterogeneous across centres. This approach is therefore valid on condition that each category of consumption is represented by cases from each centre so that, within each centre, comparisons are made across the full range of intakes. In addition, it is assumed that the diet/disease association is the same across the entire exposure range. The power of the comparison will be weaker than if each centre had the same distribution, but the comparisons and reported HRs refer to the full range of variation in intake.
Regression calibration concentrates the values of ‘calibrated’ exposure within each centre around the mean for that centre as a consequence of the shrinkage in the distribution of predicted values. In this study, the analysis of variance components of predicted intake values (i.e. after calibration) showed strong shrinkage towards centre-specific mean values, where estimates of ICCs of \( E(T|Q,Z) \) values were 2-fold higher than for \( Q \) measurements.

In the current study, no measurement error correction was performed for physical activity and smoking status, for which no reference measurements are available. Although the observed correlation between these factors and the dietary variables investigated in this study was not sizeable, the evaluation of their confounding role in the disease model is likely to be attenuated by measurement error. Therefore, the presence of bias in corrected HR estimates arising from the inclusion of variables measured with error in the disease model cannot be entirely ruled out.

Measurement errors bias the relative risk estimate in the disease model, and cause power loss. Within single study centres, calibration may correct for bias but does not recuperate power losses due to random errors. In addition, in multivariate analyses with more than one variable measured with error, incorrect tests and confidence intervals (CIs) can be produced in the disease model. In this study, a relatively more precise estimate of the overall HR was observed, compared with a naïve approach that does not take into account measurement errors in dietary variables. However, the possibility cannot be excluded that, in the context of measurement error correction with multiple variables measured with errors, the nominal values of statistical tests in the disease model are not fully controlled for.

A bootstrap sampling procedure was set up to take into account the uncertainty of estimating the attenuation factors in the calibration model compared with using a meta-analytical approach using centre-specific corrected SEs, but in contrast to the latter it does not use any specific assumptions to derive asymptotic estimates corrected SE. This could be an advantage, particularly in the case of food groups, with a sizeable frequency of zero consumers in the 24-HDR measurements. In this study, bootstrapping did not practically inflate the variability of the corrected risk parameters, possibly because of the large sample size of the EPIC calibration study, thus suggesting that the observed SEs provide sufficiently accurate estimates.

Here, a strategy for statistical analyses within the EPIC study has been presented and discussed. This strategy entails group-specific calibration models, to take into account the regional specificity and accuracy of the dietary questionnaires, and an overall disease model stratified (in the Cox model sense) by centre to evaluate the diet/disease association before and after measurement error correction. A very valuable alternative would be to use group-specific disease models, and to pool the different estimates following a meta-analytical approach in close similarity with what is done within the Pooling Project. Such a strategy was considered in this study and showed very similar results to those presented here.

Acknowledgements

The EPIC study was funded by Europe Against Cancer Programme of the European Commission (SANCO); Ligue contre le Cancer (France); Societ 3M (France); Mutuelle Generale de l’Education Nationale; Institut National de la Sante et de la Recherche Medecale (INSERM); German Cancer Aid; German Cancer Research Center; German Federal Ministry of Education and Research; Danish Cancer Society; Health Research Fund (FIS) of the Spanish Ministry of Health; the participating regional governments and institutions of Spain; Cancer Research UK; Medical Research Council, UK; the Stroke Association, UK; British Heart Foundation; Department of Health, UK; Food Standards Agency, UK; the Wellcome Trust, UK; Greek Ministry of Health; Greek Ministry of Education; Italian Association for Research on Cancer; Italian National Research Council; Dutch Ministry of Public Health, Welfare and Sports; World Cancer Research Fund (WCRF); Swedish Cancer Society; Swedish Scientific Council; Regional Government of Skane, Sweden; Norwegian Cancer Society. We are very thankful to the reviewers for their enlightening comments on the developments of this work.

Conflict of interest: None declared.

References

6 Slimani N, Deharveng G, Charrondiere RU et al. Structure of the standardized computerized 24-hour diet recall
EVALUATION OF THE DIET/DISEASE RELATION IN THE EPIC STUDY

Appendix 1

For each group \( j \), the uncorrected log-relative risk estimate of the first variable measured with errors, \( \hat{\beta}_{1,\text{obs}} \), can be expressed as

\[
\hat{\beta}_{1,\text{obs}} = \sum_{k=1}^{K} \lambda_{1,k}, \hat{\beta}_{k,\text{true}}
\]

where the terms \( \lambda_{1,k} \) indicate the elements of the vector \( \lambda_{1} \) in matrix \( A_{g} \), and \( \hat{\beta}_{k,\text{true}} \) express the true risk estimates for the variables measured with errors. The parameter \( \beta_{1,\text{obs}} \) is the result of a linear combination, which reflects a multiplicative bias, captured by the attenuation factor, \( \lambda_{1} \), as well as an additive component, associated with the CFs \( \lambda_{1,k} \) (for \( k \neq 1 \)), and the true risk estimates for the other \( K-1 \) variables \( \hat{\beta}_{k,\text{true}}(k \neq 1) \).

Appendix 2

In the case of a multicentre study \( (j=1, \ldots, J \text{ centres}) \), it is assumed, for simplicity, that the dietary variable under investigation has the same variability \( (\text{var}(Q)) \), and the same number of cancer cases \( (D) \) in all centres [e.g. \( \text{var}(Q) = \text{var}(Q) \) and \( D = D_{j} \) for each \( j \)]. Let's consider the case of one variable measured with error, where the centre-specific estimates that quantify the diet–disease associations are the parameter \( \beta_{j} (= \log(\text{RR})) \), whose variability can be approximated by

\[
\text{var}(\hat{\beta}_{j}) = \frac{1}{D_{j} \cdot \text{var}(Q)} = \frac{1}{D \cdot \text{var}(Q)}
\]

Following a meta-analytical approach, the pooled estimate can be derived as

\[
\hat{\beta}_{\text{pooled}} = \frac{\sum_{j} (\hat{\beta}_{j}/\text{var}(\hat{\beta}_{j}))}{\sum_{j} (1/\text{var}(\hat{\beta}_{j}))}
\]

and

\[
\text{var}(\hat{\beta}_{\text{pooled}}) = \frac{1}{\sum_{j} (1/\text{var}(\hat{\beta}_{j}))}
\]

Therefore, the ratio of the pooled parameter over its SE is equal to

\[
\frac{\hat{\beta}_{\text{pooled}}}{\text{SE}(\hat{\beta}_{\text{pooled}})} = \sqrt{\frac{\sum_{j} (\hat{\beta}_{j} \cdot \text{var}(Q)D)}{J \cdot \text{var}(Q)D}} = \sqrt{\frac{\sum_{j} \hat{\beta}_{j}}{J}}
\]

(A1)

The \( P \)-value associated to this ratio is computed according to a standard normal distribution.

It can be assumed that the centre-specific attenuation factors \( (\lambda_{j}) \) are inversely proportional to the magnitude of the observed associations. This is equivalent to assume that the heterogeneity of \( \beta_{j} \) estimates is entirely attributable to measurement errors with high degree of specificity in the different centres, e.g. \( \beta_{j}/\lambda_{j} = c_{j} = c \).

The de-attenuated pooled estimate can be obtained as

\[
\hat{\beta}_{\text{de-att, pooled}} = \frac{\sum_{j} (\hat{\beta}_{j}/\lambda_{j})/\text{var}(\hat{\beta}_{j})}{\sum_{j} (1/\text{var}(\hat{\beta}_{j}))} = \frac{1}{\sum_{j} \lambda_{j} \cdot \text{var}(Q)}
\]

This last result stems from the assumption that estimates of \( \lambda_{j} \) coefficients are measured with great precision, so that \( \text{var}(\lambda_{j}) \rightarrow 0 \). In this way, \( \text{var}(\hat{\beta}_{j}/\lambda_{j}) = (1/D \cdot \lambda_{j} \cdot \text{var}(Q)) \). The variance of the de-attenuated pooled estimate can be obtained as

\[
\text{var}(\hat{\beta}_{\text{de-att, pooled}}) = \frac{1}{\sum_{j} (1/\text{var}(\hat{\beta}_{j}))} = \frac{1}{\sum_{j} (1/\text{var}(\lambda_{j}))}
\]

and for the de-attenuated ratio

\[
\hat{\beta}_{\text{de-att, pooled}}/\text{SE}(\hat{\beta}_{\text{de-att, pooled}}) = c \sqrt{D \cdot \text{var}(Q) \sum_{j} \lambda_{j}^{2}}
\]

(A2)

It is shown that the de-attenuated ratio in (A2), under the hypotheses made, is greater than the ratio observed in (A1), thus indicating a gain in the relative precision of the overall estimate after regression calibration.

\[
c \sqrt{D \cdot \text{var}(Q) \sum_{j} \lambda_{j}^{2}} > \sqrt{D \cdot \text{var}(Q) J} \frac{\sum_{j} \hat{\beta}_{j}}{\sqrt{J}}
\]

Recalling that \( \hat{\beta}_{j}/\lambda_{j} = c \), and \( \hat{\beta}_{j} = \hat{\lambda}_{j} c \), it follows that

\[
c \sqrt{\sum_{j} \lambda_{j}^{2}} > \sum_{j} \lambda_{j} \cdot c \sqrt{J}
\]

(A3)

After definition of the two quantities \( v(\lambda) = \frac{1}{J} \sum_{j} (\hat{\lambda}_{j} - \hat{\lambda}_{j})^{2} \) and \( \hat{\lambda}_{j} = J (\sum_{j} \hat{\lambda}_{j}) / J \),

\[
\sum_{j} \lambda_{j}^{2} = (v(\lambda) + J \hat{\lambda}^{2} / J)
\]

\[
\sum_{j} \hat{\lambda}_{j} = \hat{\lambda} \sqrt{J}
\]

Expression (A3) is equivalent to

\[
v(\lambda) + (J - 1) \hat{\lambda}^{2} > 0
\]

(A4)

Expression (A4) always applies, as a sum of squares plus the product of the number of centres times a square is a positive quantity. For this reason, standing the assumptions made, the associated \( P \)-value will be lower after measurement error correction, thus indicating that in cases where the between-centres heterogeneity in the diet–disease associations is largely attributable to differential measurement errors across centres, measurement error correction can lead to a relative gain in statistical power.