Epidemiological and experimental laboratory studies on nutrition and cancer have provided strong evidence that diet, as well as other related variables such as body mass and physical activity, can influence the risk of developing different types of cancer.

Both epidemiological and laboratory research in this field were first developed in the 1970s with the main idea of searching for foods or food constituents which might have carcinogenic effects. This approach was largely inspired by the classical paradigm of chemical carcinogenesis which dominated cancer research for several decades.

In their most simplified form, the chemical or physical carcinogenesis models required the presence of at least one initiator which causes nonreversible DNA mutation and possibly one promoter which provides the mutated cell clones with a selective advantage. Under this traditional, simple and now outdated model, the potential metabolic effects of different diets were considered to be of only marginal secondary importance.

From the late 1960s onwards, the multiplication of population-based cancer registries provided reliable data for the first time on the incidence of specific cancers in entire populations and in specific subgroups defined by age and sex [1, 2]. These data confirmed that there are extremely large variations in cancer inci-
dences around the world which are clearly and strongly related to lifestyle and environmental carcinogens. For instance in lung, liver and bladder cancer these discrepancies could largely be explained by variations in exposure to tobacco, alcohol, hepatitis B virus, bladder parasites (*Schistosoma haematobium*) and some occupational carcinogens.

Cancer registry data, however, also indicated that there were – and to a large extent there still are – large variations in incidence of cancers for which the link with exposure to known chemical, physical or biological carcinogens has never been clearly identified. This is the case for cancers of the breast, prostate, endometrium, colon, rectum and, to some extent, also the stomach, besides the recently established association with chronic infection by *Helicobacter pylori* [3].

The hypothesis that the carcinogenic process in these organs could mainly be determined by diet and related metabolic, anthropometric and hormonal factors was originally supported by a series of early case-control studies, by geographical correlation studies (also called ecological studies) [4] as well as by pioneering work on rodents in experimental laboratory studies carried out in the 1940s [5–8].

During the first half of the 1900s, several researchers noticed that overfed rodents were less resistant to the growth of transplanted tumors [9] and more sensitive to the effect of chemical carcinogens [7]. In particular, Tannenbaum [7], in a large and impressive series of laboratory experiments, found that excess energy intake was more effective than excess fat intake in increasing tumor yield in various chemical carcinogenesis models and that excess fat induced more tumors than excess carbohydrates or excess protein when tested in isocaloric diets.

Several decades later, correlation studies showed that the incidence of (and mortality from) cancer of the breast, colorectum and prostate were positively correlated with the foods most typically consumed in Western societies (*i.e.* meat, total and animal fat, simple sugars) and negatively correlated with the consumption of various vegetable foods (grains, cereals and vegetable fiber) [10]. Geographical correlation studies can, however, only indicate that disease risk and the prevalence of a given factor are correlated across different populations, but they are limited by the methodological and practical possibility of taking into account confounding factors which may create spurious correlations at the population level.

If well-designed, case-control and prospective cohort studies can collect information on a variety of environmental and lifestyle factors at the individual level. Using appropriate statistical methods [11, 12] it is possible to estimate the disease risk for the factor of interest while, at the same time, removing spurious associations due to confounding factors.

During the past 20 years, a considerable number of retrospective case-control and, more recently, prospective cohort studies have been conducted to investigate whether, in each given population, these dietary factors were effectively related to cancer risk at the individual level.
Results of epidemiological and experimental studies on nutrition and cancer have been reviewed in depth in recent years by three independent expert committees:

1. In the UK, the Committee on Medical Aspects of Food and Nutrition Policy (COMA) Working Group on Diet and Cancer [13]
3. At the international level, the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) [15]

The three independent review committees reached broadly similar conclusions. Regarding dietary composition, the three reports agreed that the most clearly identified associations are those between the consumption of vegetables and fruit and reduced risk of various cancers. These protective effects have been most consistently seen in studies on cancers of the digestive and respiratory tracts. Frequent, daily consumption of both vegetables and fruit is associated most strongly with a reduction in risk of cancers of the mouth, pharynx, larynx, esophagus, stomach and lung, while only consumption of vegetables, but not of fruit, is linked to a reduction in risk of colorectal cancer.

Regarding foods which have consistently been identified as being associated with increased cancer risk, the list is much shorter and is limited to red meat and Cantonese-style salted fish. Red meat, mainly beef, but not poultry and fish, are associated with a modest increase in colorectal cancer risk. Consumption of Cantonese-style salted fish has been found to be associated with the risk of nasopharyngeal cancer, which is very frequent in some populations of south-east Asia, particularly southern China, but extremely rare in most other parts of the world, with the exception of the north east of Africa and the Inuit population in North America.

Regarding fruit, vegetables and red meat, the three reports agreed on an important point: while their association with cancer risk can be considered to be reasonably well established on the basis of epidemiological studies, there is no definite clear explanation for the biological mechanisms involved, even though a large number of experimental studies have been carried out, and many different mechanisms have been tested on in vitro and in vivo models.

This conclusion leads to two main recommendations. First of all, it implies that for the time being there is no scientific support for a proven cancer-preventive effect of dietary supplements containing various cocktails of vitamins and minerals also found in vegetables and fruits. The only sound recommendation is therefore to eat lots of fruit and vegetables frequently. Secondly it implies that more research is needed on the biological links between vegetables, fruits and the carcinogenesis process, particularly with randomized trials and observational epidemiological studies.
Another growing and promising area of research concerns the relation between anthropometric characteristics, physical activity and cancer risk. The first evidence that overweight may be linked to increased cancer risk dates back to the 1930s when Tannenbaum [16] conducted a study on mortality in relation to height and weight using the data of various life insurance companies in the USA.

Epidemiological studies conducted during the past 20 years have shown with varying degrees of consistency that excess body mass (usually estimated as kg/m², or body mass index, BMI) is associated with increased risk of cancer of the endometrium, breast and colon. It has been suggested that there may be an association between anthropometric characteristics and risk of cancer at other sites, particularly cancer of the kidney, but the data are more inconsistent. So far, the strongest and most consistent association with body mass has been seen for endometrial cancer, the risk of which is increased 2- to 6-fold in obese compared to lean women, both before and after menopause. A possible biological explanation for this association is that adipose tissue is rich in aromatase which converts androstenedione to estrone, thus increasing estrogenic stimulation of the endometrial mucosa. Several studies have investigated markers of fat distribution, such as weight-to-hip ratio (WHR) or subscapular-to-triceps-skinfold ratio (STR), in relation to endometrial cancer risk with inconsistent results. Some studies found increased risk for markers of abdominal or android obesity (high WHR or STR) after adjustment for BMI, while others did not. The relationship between BMI and breast cancer is even more complex. The majority of case-control and prospective studies found that high BMI increased breast cancer risk in postmenopausal women, while it may slightly reduce risk in premenopausal women [17]. A possible explanation for this apparent paradox is that overweight before menopause could be related to anovulatory cycles, and fewer ovulatory cycles (as determined by pregnancy and lactation) are generally associated with lower breast cancer risk. After menopause, obesity may act as for endometrial cancer by enhancing the peripheral (as opposed to gonadal and surrenal-cortical) production of estrogens.

Recently, several prospective studies in which blood samples were collected and stored at baseline from healthy subjects have shown that high prediagnostic levels of endogenous steroid hormones, mainly estrogens and testosterone, are associated with a 3- to 6-fold increase in breast cancer risk [18–21], while one study showed that high testosterone levels increase prostate cancer risk [22].

These studies also found that low levels of sex hormone-binding globulin (SHBG) are also associated with higher risk of breast cancer. SHBG is synthesized in the liver, and its production is downregulated by insulin. The link between overweight, a sedentary lifestyle and cancer risk may well be mediated in part by these insulin-SHBG-steroid hormone pathways for cancer of the breast [23], colon [24], prostate and possibly other cancers.

An equally important link between diet and cancer risk may exist through the control of the production of insulin-like growth factors (IGFs) and their binding proteins (IGFBPs). Recent studies found that high levels of IGF₁ (adjusted over
the levels of IGFBP3) were significant predictors of the risk of developing cancer of the prostate [25] and colon [26]. More recently, we found in the New York Women’s Health Study that high levels of c-peptide, a serum marker of insulin excretion, were strongly associated with the risk of developing colon cancer [27]. Variations in the levels of estrogens, androgens, IGFs and IGFBPs are probably determined by both environmental and lifestyle factors, as well as by inherited genetic characteristics, as suggested by recent studies on polymorphisms of genes encoding for enzymes regulating steroid hormone metabolism and hormone receptors.

These results on endogenous hormones and anthropometry indicate that the relationship between diet and cancer is much more complex than was previously thought. Research on diet and cancer, based solely on simple dietary questionnaire measurements and mainly retrospective case-control studies, has led to the identification of some major dietary patterns associated with cancer risk (particularly the balance between vegetables, fruits and meat). While these results are sufficient to support some broad and nowadays widely accepted dietary recommendations, cancer prevention would benefit from a better understanding of the biological links between diet and cancer. We believe that traditional case-control studies have little chance of leading to any major breakthrough in our understanding of this matter. Laboratory investigations on human subjects combined with sound epidemiological projects of a prospective nature seem more likely to lead us a step further. This was the strategic choice made by the International Agency for Research on Cancer (IARC) when it decided 10 years ago to give priority in nutrition and cancer studies to the development of prospective cohort studies with repositories of blood samples collected from healthy study subjects. We present here the major developments of the research strategy which led to the realization of the European Prospective Investigation into Cancer and Nutrition (EPIC).

European Prospective Investigation into Cancer and Nutrition

EPIC is a multicenter prospective study aimed at investigating the complex relations between nutrition and various lifestyle factors and the etiology of cancer and other chronic diseases. The study was initiated in 1993 with the collection of data and blood samples in 22 regional centers located in 9 European countries (Denmark, France, Germany, Greece, Italy, Netherlands, Spain, Sweden, and United Kingdom). The list of collaborating centers is given in the Appendix. The field work was completed in 1998 with the inclusion of 484,042 subjects. All provided questionnaire data, and from 387,256 of them blood samples were collected and stored at −196 °C for future analyses on cancer cases and controls.

EPIC was designed with the double aim of improving scientific knowledge on the nutritional factors involved in diet and, as a consequence, of providing the
scientific bases for public health interventions directed to promoting healthier diet and lifestyle.

Diet has changed substantially over the centuries and keeps on changing because of cultural and economic factors. Long-term dietary changes will therefore occur in one way or another. Better knowledge of the foods that can reduce cancer risk may help to determine future changes aimed at a reduction of cancer risk.

**Study Protocol**

In brief, the general protocol for subject recruitment and data collection was as follows. As a rule, eligible study subjects were from the general population residing in a given geographical area, a town or a province. There were, however, a few exceptions: the French cohort was based on members of the health insurance for state school employees (with the aim of facilitating long-term follow-up), a small component of the Italian and Spanish cohorts included members of local blood donor associations, and the Utrecht cohort was based on women attending breast cancer screening.

Eligible subjects were invited by mail to participate in the study. In some cases (e.g. blood donors) the first invitation was by personal contact. Those who accepted signed an informed consent form, and the diet and lifestyle questionnaires were mailed to them to be filled in, generally at home. Study subjects were then invited to a center for blood collection (venipuncture), anthropometric measurements (height, weight, waist, hip and sitting height) and to hand in the completed questionnaire.

Data were collected on a large number of lifestyle and health factors which are of interest in studies on nutrition and cancer, as they may be related to nutritional status or may be known or suspected cancer risk factors. A common core set of questions and possible answers was agreed upon and translated into national questionnaires. This included questions on:

1. Education, socioeconomic status
2. Current job, current and past occupation in industrial settings which might have led to exposure to carcinogens
3. Life status of parents and siblings and their cause of death
4. History of previous illness and disorders or surgical operations
5. Lifetime history of tobacco smoking
6. Lifetime history of consumption of alcoholic beverages
7. Physical activity: occupation, walking, cycling, gardening, housework, physical exercise, climbing stairs
8. Sexual maturation, contraception and reproduction

**EPIC** was initiated in 1990 with methodological studies on dietary assessment and pilot-feasibility studies on subject recruitment and collection of question-
naire data and blood samples. These studies, which took place in 9 different countries between 1990 and 1993, led to a series of publications. In particular, a supplement of the *International Journal of Epidemiology* [28] was devoted to the validity of different dietary assessment methods, of questionnaires on physical activity and on the reproducibility of anthropometric measurements. These studies provided precious information for the finalization of the study protocol. Following the results of these methodological studies conducted in 1990–1992, three dietary methods were adopted:

1. Extensive self-administered dietary questionnaire, which can provide data on up to 300–350 food items per country. This method was used in seven countries.
2. Interview-based dietary questionnaire, very similar in content to the above, but administered by direct computerized interview. This method was used in Spain and in Ragusa (Italy) to increase compliance.
3. Food frequency questionnaire combined with a 7-day record. This method was adopted by the two centers in England.

In addition to the above dietary measurements obtained from all study subjects, it was decided to implement in EPIC a novel methodological approach aimed at calibrating dietary measurements across countries in order to correct for systematic over- or underestimation of dietary intakes. For this purpose, a second dietary measurement was taken from an 8–10% random sample of the cohort using a computerized 24-hour diet recall method developed *ad hoc* [29, 30]. We developed statistical methods to correct for bias in relative risk estimates due to systematic measurement errors in the baseline questionnaire, thereby making the cohort-specific estimates more comparable between study centers [31, 32].

While this general protocol was common to all EPIC centers, the procedure for storage of blood samples differed between two groups of countries due to the fact that the study was originally started in seven European countries (France, Germany, Greece, Italy, Netherlands, Spain, and United Kingdom) where the study followed a common, jointly elaborated protocol as strictly as possible. Blood samples collected from subjects in these seven countries were aliquoted in 28 plastic straws (12 plasma, 8 serum, 4 erythrocytes, 4 buffy coat for DNA) and then split into two sets of 14 aliquots each. One set was stored locally and one shipped to IARC to be stored in liquid nitrogen at –196 °C in the central biological bank. Later on, four additional study centers located in Sweden (Malmö and Umeå) and two in Denmark (Copenhagen and Aarhus) joined EPIC as associated projects. The two Swedish cohorts had started before EPIC and the Danish ones in parallel with EPIC. Their protocols and questionnaires were adapted to become as close as possible to those used in EPIC. Blood samples from these four centers were, however, stored in tubes (not in plastic straws), so for practical reasons they were stored locally, as the EPIC system set up at IARC is not suitable for storing tubes.
### Table 1. Subject recruitment in the EPIC Study (September 1999)

<table>
<thead>
<tr>
<th>Country</th>
<th>Subjects included in the study with questionnaires</th>
<th>Subjects included in the study with blood collection</th>
<th>Completion of subject recruitment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spain</td>
<td>41,446</td>
<td>40,040</td>
<td>1996</td>
</tr>
<tr>
<td>Italy</td>
<td>53,097</td>
<td>53,077</td>
<td>1998</td>
</tr>
<tr>
<td>UK</td>
<td>88,171</td>
<td>43,430</td>
<td>1998</td>
</tr>
<tr>
<td>Netherlands</td>
<td>40,110</td>
<td>36,357</td>
<td>1997</td>
</tr>
<tr>
<td>France</td>
<td>69,321</td>
<td>24,371</td>
<td>1993</td>
</tr>
<tr>
<td>Germany</td>
<td>53,130</td>
<td>50,719</td>
<td>1998</td>
</tr>
<tr>
<td>Greece</td>
<td>27,883</td>
<td>28,632</td>
<td>1999</td>
</tr>
<tr>
<td>Sweden</td>
<td>53,830</td>
<td>53,830</td>
<td>1996</td>
</tr>
<tr>
<td>Denmark</td>
<td>57,054</td>
<td>56,800</td>
<td>1997</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>484,042</td>
<td>387,256</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2. Number of cancer cases expected to have occurred in women in the EPIC study up to the end of 1998 and notified to IARC by the end of 1999

<table>
<thead>
<tr>
<th>Country</th>
<th>Center</th>
<th>Stomach</th>
<th>Colon-rectum</th>
<th>Lung</th>
<th>Breast</th>
<th>Cervix uteri</th>
<th>Corpus uteri</th>
<th>All sites but 173</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>Paris</td>
<td>22</td>
<td>138</td>
<td>45</td>
<td>521</td>
<td>74</td>
<td>99</td>
<td>1,327</td>
</tr>
<tr>
<td>Italy</td>
<td>Florence</td>
<td>7</td>
<td>17</td>
<td>7</td>
<td>49</td>
<td>4</td>
<td>10</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>Varese</td>
<td>6</td>
<td>17</td>
<td>6</td>
<td>60</td>
<td>5</td>
<td>10</td>
<td>154</td>
</tr>
<tr>
<td></td>
<td>Ragusa</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Spain</td>
<td>Turin</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>19</td>
<td>2</td>
<td>3</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Oviedo</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>23</td>
<td>5</td>
<td>6</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Granada</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>20</td>
<td>3</td>
<td>5</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Murcia</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>23</td>
<td>4</td>
<td>6</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Pamplona</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>26</td>
<td>2</td>
<td>5</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>San Sebastian</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>19</td>
<td>3</td>
<td>4</td>
<td>52</td>
</tr>
<tr>
<td>UK</td>
<td>Cambridge</td>
<td>8</td>
<td>40</td>
<td>35</td>
<td>111</td>
<td>9</td>
<td>17</td>
<td>333</td>
</tr>
<tr>
<td></td>
<td>Oxford</td>
<td>6</td>
<td>31</td>
<td>26</td>
<td>131</td>
<td>15</td>
<td>15</td>
<td>332</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Bilthoven</td>
<td>2</td>
<td>11</td>
<td>8</td>
<td>52</td>
<td>5</td>
<td>6</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>Utrecht</td>
<td>9</td>
<td>45</td>
<td>26</td>
<td>127</td>
<td>7</td>
<td>22</td>
<td>346</td>
</tr>
<tr>
<td>Greece</td>
<td>Athens</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>28</td>
<td>5</td>
<td>5</td>
<td>86</td>
</tr>
<tr>
<td>Germany</td>
<td>Heidelberg</td>
<td>2</td>
<td>11</td>
<td>5</td>
<td>30</td>
<td>6</td>
<td>5</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Potsdam</td>
<td>3</td>
<td>13</td>
<td>6</td>
<td>35</td>
<td>7</td>
<td>7</td>
<td>106</td>
</tr>
<tr>
<td>Sweden</td>
<td>Malmö</td>
<td>12</td>
<td>56</td>
<td>30</td>
<td>172</td>
<td>11</td>
<td>39</td>
<td>503</td>
</tr>
<tr>
<td></td>
<td>Umeå</td>
<td>4</td>
<td>17</td>
<td>5</td>
<td>68</td>
<td>7</td>
<td>12</td>
<td>190</td>
</tr>
<tr>
<td>Denmark</td>
<td>Aarhus</td>
<td>2</td>
<td>16</td>
<td>18</td>
<td>43</td>
<td>6</td>
<td>11</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>Copenhagen</td>
<td>6</td>
<td>44</td>
<td>50</td>
<td>117</td>
<td>17</td>
<td>29</td>
<td>389</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>107</td>
<td>499</td>
<td>280</td>
<td>1,683</td>
<td>195</td>
<td>315</td>
<td>4,624</td>
</tr>
</tbody>
</table>
Table 3. Number of cancer cases expected to have occurred in men in the EPIC study up to the end of 1998 and notified to IARC by the end of 1999

<table>
<thead>
<tr>
<th>Country</th>
<th>Center</th>
<th>Stomach</th>
<th>Colon-rectum</th>
<th>Lung</th>
<th>Prostate</th>
<th>All sites but 173</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>Florence</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>2</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Varese</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>1</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Ragusa</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Turin</td>
<td>5</td>
<td>10</td>
<td>21</td>
<td>3</td>
<td>85</td>
</tr>
<tr>
<td>Spain</td>
<td>Oviedo</td>
<td>4</td>
<td>6</td>
<td>16</td>
<td>2</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Granada</td>
<td>2</td>
<td>3</td>
<td>8</td>
<td>1</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Murcia</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Pamplona</td>
<td>8</td>
<td>8</td>
<td>14</td>
<td>5</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>San Sebastian</td>
<td>8</td>
<td>9</td>
<td>19</td>
<td>3</td>
<td>96</td>
</tr>
<tr>
<td>UK</td>
<td>Cambridge</td>
<td>19</td>
<td>47</td>
<td>77</td>
<td>42</td>
<td>311</td>
</tr>
<tr>
<td></td>
<td>Oxford</td>
<td>6</td>
<td>15</td>
<td>23</td>
<td>11</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>Athens</td>
<td>5</td>
<td>6</td>
<td>25</td>
<td>4</td>
<td>78</td>
</tr>
<tr>
<td>Germany</td>
<td>Heidelberg</td>
<td>6</td>
<td>16</td>
<td>30</td>
<td>8</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>Potsdam</td>
<td>6</td>
<td>16</td>
<td>30</td>
<td>9</td>
<td>118</td>
</tr>
<tr>
<td>Sweden</td>
<td>Malmö</td>
<td>17</td>
<td>50</td>
<td>48</td>
<td>90</td>
<td>383</td>
</tr>
<tr>
<td></td>
<td>Umeå</td>
<td>9</td>
<td>20</td>
<td>14</td>
<td>29</td>
<td>151</td>
</tr>
<tr>
<td>Denmark</td>
<td>Aarhus</td>
<td>4</td>
<td>17</td>
<td>28</td>
<td>9</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>Copenhagen</td>
<td>10</td>
<td>43</td>
<td>70</td>
<td>23</td>
<td>304</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>127</td>
<td>297</td>
<td>483</td>
<td>251</td>
<td>2,298</td>
</tr>
</tbody>
</table>

In Sweden, the samples are kept in deep freezers at –70°C, and in Denmark in nitrogen vapor.

Field Work and Subject Recruitment

The field work for the recruitment of study subjects, the collection of questionnaire data and anthropometric measurements, and the collection and storage of blood samples, took place from 1993 to 1998.

Table 1 summarizes the number of subjects for whom questionnaire data and blood samples were collected in each participating country. By September 1999, EPIC included 484,042 subjects who had provided questionnaire data, and from 387,256 of them blood samples had been collected and stored. In addition, as planned in the design of the study for internal calibration of dietary measurements [33, 34], 24-hour diet recalls were collected on a sub-sample of 35,167 subjects, corresponding to about 7.3% of the cohort. The age distribution of the calibration sample was designed to be as close as possible to the age distribution of the expected cancer cases during the first 10 years of follow-up.
These results will endow EPIC with an unusually large power to study the various cancer risk factors of interest. In fact, over 22,000 cancer cases are expected to occur in the EPIC cohorts during the first 10 years of follow-up (by 2005). Tables 2 and 3 provide the expected number of cancer cases, by cancer site and country, expected to have occurred up to the end of 1998.

Follow-Up for Changes in Lifestyle and Health Conditions and for Cancer Incidence and Mortality

In EPIC, cohort members are contacted 3–4 years after recruitment to obtain information on some aspects of lifestyle which are known or strongly suspected to be related to cancer risk: tobacco smoking, alcohol drinking, physical activity, weight, menstruation, pregnancies, menopause, etc. In addition, a series of questions was added on whether the subjects had suffered from any major diseases. The first run of individual follow-up is currently on-going in most EPIC centers, and in a few it has been completed.

Follow-up aimed at the identification of cancer cases occurring among the EPIC cohort is based on population cancer registries in six of the participating countries (Denmark, Italy, Netherlands, Spain, Sweden, United Kingdom) and on a combination of methods including health insurance records, cancer and pathology registries, and active follow-up through study subjects and their next-of-kin in three countries (France, Germany and Greece). Mortality rates are also collected from either the cancer registry or mortality registries at the regional or national level.

A working group, created in 1996 (End-Point Committee), prepared a detailed protocol for the collection and standardization of clinical and pathological data on each cancer site: ‘Guidelines for Collection of End-Point Data in the EPIC Study’ (IARC, 1998). The document is available upon request.

Currently, follow-up is being completed up to December 31, 1997. A delay of at least 18–24 months in obtaining complete follow-up data is unavoidable due to the complex procedures followed by population-based cancer registries for the collection and verification of clinical and pathological diagnoses. On the other hand, these procedures provide complete and reliable follow-up data.

Conclusions

Epidemiological studies on nutrition and cancer have provided strong evidence that dietary patterns, anthropometric characteristics and physical activity play an important role in the etiology of some of the most common cancers. Currently, public health recommendations generally promote the consumption of vegetables and fruits and advise moderation in the consumption of meat and salty foods. While these general recommendations are justified by the current state of the art situations, recent studies on metabolic factors (hormones and biomarkers
of diet) suggest that the relation between nutrition and cancer is probably much more complex and involves various lifestyle factors besides single dietary composition. Most epidemiological studies conducted so far were limited by the fact that they covered single and relatively homogeneous populations with limited variations in dietary habits, and by being based solely on information provided by the study subjects through questionnaires.

The EPIC study was designed with the aim of overcoming these two limitations. Firstly, EPIC includes populations with important variations in dietary intake, particularly of vegetables and fruit. Secondly, with the collection and storage of blood samples, EPIC can provide the material for investigating various nutrition-related metabolic and genetic factors and their possible interactions.

**Acknowledgements**

The work described in this paper is being carried out with the support of the “Europe Against Cancer” Programme of the European Commission.

**Appendix**

**Researchers and Institutes Collaborating in the European Prospective Investigation into Cancer and Nutrition (EPIC)**

| Coordination | International Agency for Research on Cancer (IARC) | Lyon | Dr. Elio Riboli
|             | Dr. Rodolfo Saracci | Dr. Rudolf Kaaks | Ms Nadia Slimani | Ir Ann Linda Van Kappel |
| France      | National Institute for Health & Medical Research (INSERM), Institut Gustave Roussy | Villejuif (Paris) | Dr. Françoise Clavel | Dr. Catherine Guibout |
| Germany     | German Cancer Research Centre | Heidelberg | Dr. Anthony B. Miller | Dr. Jürgen Wahrendorf | Dr. Nikolaus Becker |
|             | German Institute for Human Nutrition | Potsdam-Rehbrücke | Dr. Heiner Boeing | Dr. Anja Kroke | Dr. Manuela Bergmann | Mr. Andro Jeckel |
| Denmark     | Danish Cancer Society | Copenhagen | Dr. Anne Tjønneland | Dr. Hans Storm |
|             | Institute of Epidemiology and Social Medicine, University of Aarhus | Aarhus | Dr. Kim Overvad |
### Appendix (continued)

<table>
<thead>
<tr>
<th>Country</th>
<th>City</th>
<th>Institute/University</th>
<th>Person(s)</th>
</tr>
</thead>
</table>
| Spain   | Barcelona | Catalan Institute of Oncology (ICO) | Dr. Carlos González  
Dr. Antonio Agudo  
Mr. Guilem Pera |
|         | Granada | Andalusian School of Public Health | Dr. Carmen Martinez  
Dr. Maria J. García |
|         | Oviedo | Council for Health and Social Affairs of Asturia | Dr. J. Ramón Quiros  
Dr. Cristina Lasheras |
|         | San Sebastian | Health Administration of Guipuzcoa (Basque Country) | Dr. Miren Dorronsoro  
Dr. Pilar Amiano  
Dr. José M. Begiristain |
|         | Murcia | Council for Health and Social Affairs of Murcia | Dr. Carmen Navarro  
Dr. Maria D. Chirlaque  
Dr. Maria J. Tormo |
|         | Pamplona | Institute of Public Health, Regional Government of Navarra | Dr. Aurelio Barricarte  
Dr. Ana Barcos |
| UK      | Cambridge | Institute of Public Health, University of Cambridge | Dr. Nicholas E. Day  
Ms Suzy Oakes  
Ms Ailsa Welch  
Mr. Robert Luben |
|         |           | Dunn Clinical Nutrition Centre | Dr. Sheila Bingham  
Dr. Kay-Tee Khaw |
|         | Oxford | Imperial Cancer Research Fund, Cancer Epidemiology Unit, University of Oxford | Dr. Timothy Key  
Dr. Gwyneth Davey  
Dr. Valerie Beral |
| Greece  | Athens | University of Athens School of Medicine | Dr. Antonia Trichopoulos  
Dr. Klea Katsouyanni  
Dr. Dimitrios Trichopoulos |
| Italy   | Milan | National Cancer Institute | Dr. Franco Berrino  
Dr. Vittorio Krogh  
Dr. Valeria Pala |
|         | Florence | Centre for Cancer Research and Prevention (CSPO) | Dr. Domenico Palli  
Dr. Eva Butiatti  
Dr. Calogero Saieva |
|         | Turin | Cancer Epidemiology Unit, University of Turin | Dr. Paolo Vineis  
Dr. Benedetto Terracini |
|         | Ragusa | Ragusa Cancer Registry | Dr. Rosario Tamino  
Dr. Lorenzo Gafà  
Dr. Graziella Frasca |
Appendix (continued)

Naples Department of Biomedical Sciences and Human Oncology, University of Naples Dr. Salvatore Panico

Netherlands Utrecht Julius Center for Patient Oriented Research, Utrecht University Medical School Dr. Petra Peeters Ms Margreet A. Pols Dr. Paul A.H. Van Noord Ms Janneke Verloop
Bilthoven National Institute of Public Health & Environmental Protection (RIVM) Dr. Bas Nuens de Mesquita Dr. Marga Ocké Ms Ester Goddijn Dr. Daan Kromhout Dr. Jaap Seidell Dr. Monique Verschuren

Sweden Malmö Departments of Medicine and Community Medicine, Lund University Dr. Göran Berglund Dr. Lars Janzon
Umeå Department of Nutritional Research, University of Umeå Dr. Göran Hallmans

References
Discussion

Participant: Can I ask you how you are funding this project?

Dr. Riboli: Fifty percent from the European Union and the other 50% through national grants. So it is a joint venture between the European Commission and national bodies. We spent 2 years at the beginning in planning the study and trying to get as much information as possible at the smallest cost, and I think we’ve achieved a reasonable result. The total cost per subject for all the biological samples, anthropometric measurements, and sample storage averages 90 dollars.

Dr. Haschke: Can you enlarge on the methodology in relation to selection of the participants? You have a very large number of subjects, but how did you control whether the participants in their respective centers are representative of the background population? Is it possible to draw random samples of the population in this study?

Dr. Riboli: That is an important point. As you can imagine, we had a lot of discussion at the beginning of the study about whether we should give priority to obtaining a fully representative sample of the population, or whether we should give priority to obtaining a sample that would be more suitable for long-term follow-up. It’s a trade-off.

This study was designed to investigate a biological hypothesis – that is, whether given characteristics are related to the occurrence of disease. To make an ambitious comparison, the British Doctors’ study set up by Sir Richard Doll in 1949/1950 [1] was the key study that showed that smoking causes lung cancer. That study included only British doctors, who are certainly not representative of the British population, but it was sufficient to establish that there is a biological relation between smoking and subsequent risk of lung cancer. Then you may say, yes, that’s all very well, but what is the impact of smoking on the British population? So the next step is to obtain data on how many British people smoke, from which you can derive the expected lung cancer incidence; once you know that there is a biological relation, you can estimate the weight of smoking in the British population.

So by analogy, we decided that the most important factor in a study planned to extend over 30 years was to get a population that was relatively ‘easy to follow’. So we invited members of the general population to participate, either through letters or through contact by television, radio, and so on, and we got a response rate that varied between 30 and 70%. We have already done some analyses on key lifestyle factors in the study population and we are very pleased that these do not differ much from the general population. For example, among men, we have a proportion of smokers slightly lower than in the general population, though among women the proportion is slightly higher. Average meat consumption is very close to the national averages. In all, my impression is that we have lost the very poor and the very rich, but we have a good sample of the 90% in the middle.

Dr. Muti: In premenopausal women you mentioned that there is a clearly established association between steroids and breast cancer, but at present the evidence is only for postmenopausal women. For premenopausal women we don’t have this evidence, perhaps because of difficulty in standardizing hormone measurements. But maybe it is also possible that premenopausal women have other kinds of determinants. For instance, there is evidence that in premenopausal women, genetic characteristics are far more important than in postmenopausal women. Also, the insulin resistance hypothesis may be more relevant for premenopausal women, in addition to any effects on ovarian estrogen and androgen production. Perhaps insulin resistance in premenopausal women is an epiphenomenon that is related to other metabolic risk factors.

Dr. Riboli: I do agree with you that the study of the relation between steroid concentrations and cancer risk in premenopausal women is a nightmare, because of the variability related to the menstrual cycle. That is particularly true for estrogen. We were perfectly aware of this. All we could do – and we know this is not ideal – was to record the date of the last menstruation, and then give the premenopausal women a postcard on which they wrote
down the date of the three subsequent menses. This card was sent back to our centers. Thus we have the date of the blood collection, the time in hours and minutes of the blood collection, to take care of circadian variation, and the date of the previous menses and of the three following menses. In this way we hoped to obtain a reasonable estimate of whether the blood was collected during the follicular phase or the luteal phase or around the peri-ovulatory phase, though realizing that this was quite an imprecise classification.

Dr. Muti: Could you tell us something about the timing of blood collection in the different centers? Was it collected in the morning, or in the afternoon? Were the subjects fasting?

Dr. Riboli: Blood was collected at any time of the day, because that was the only way in which we could collect blood from 400,000 subjects and still get a reasonable response rate. So in some subjects we got blood when in the fasting state, in others we got blood in the middle of the morning or in the early afternoon. To try to resolve this problem, we asked an additional question at the time of blood sampling, which was the time at which the subject last took any food or drink. Thus we have some idea of the period over which they had been fasting. We also asked when they smoked their last cigarette, because we know this affects vitamin C levels. Finally, we obtained a list of all drugs and supplements consumed over the previous 3 days. Ideally, we would have liked two blood samples per person, but this was not financially possible.

Dr. Mason: Given the increasing interest in how environmental factors during infancy and childhood affect your risk of developing cancer as an adult, and given the fact that EPIC has at least tentative plans to continue follow-up for several decades, I wonder whether you have any plans to include a cohort of infants or children in your sample.

Dr. Riboli: We would have liked to do that, but when I suggested the idea to several funding bodies, I was not successful, probably because the investment was felt to be too long-term. I think it would have been wonderful to have had a subcohort of teenagers; in fact our dream 10 years ago was to have around 20,000 or 30,000 teenagers in the study. We have not lost all hope of convincing the funding bodies to include a subcohort of the children of the EPIC subjects, but for the time being this is only wishful thinking.

Dr. Bloch: Why did only one third of your French subjects have blood samples taken?

Dr. Riboli: That was purely a logistical problem. In France when we started the study 10 years ago there were a lot of restrictions on doing follow-up studies, and in fact there still are. The way to escape this restriction was to use a health insurance company, which provided semi-private health insurance for school employees. This insurance company agreed to help us in providing the medical data on their members. The price to pay for this was that the 90,000 women who agreed to take part were scattered all over France, from Paris to the smallest village in the Alps, so blood collection was quite complicated. We set up arrangements for blood collection only in urban areas, and that is why we only got samples from a third of the women. Nevertheless, the response rate was very good.

Dr. Pichard: Regarding the follow-up aspect of this study, if I understood you correctly, you now have cell nuclear material in your freezers. I wonder how you have overcome the legal objections to this. As far as I am aware it is not now legal to store genetic material that is not anonymized.

Dr. Riboli: Well, we have complied so far with all the national regulations regarding this issue, and we had clearance from ethics committees in all the countries involved in the study. One of the key factors is that we have emphasized that we are going to use this material mainly to study genetic polymorphisms in cases where the association with disease risk is unknown but is likely to be weak. All the blood samples are totally anonymous. Only the national collaborators have access to the link between the study subject numbers and their names, and this information is kept in a separate file from the other data.

Participant: Considering the long time period during which your samples will be stored, are you sure that the measurements you will make on hormones and other compounds in 2, or 3, or 10 years will be reliable? I’m concerned about stability.
Dr. Riboli: This is a very important point. I can assure you that at the beginning of the study we were very concerned about the issue of stability. Before deciding on liquid nitrogen, we carried out studies where we measured stability over a 1-year period, at different temperatures and with different systems of storage. We considered temperatures of –20°C, –80°C, and liquid nitrogen, which is –196°C, and we finally decided on liquid nitrogen because that was the only temperature at which there was no sign whatsoever of degradation of peptidic hormones, steroid hormones, or very unstable molecules like vitamin C. As an anecdote I can tell you that we are now doing extensive laboratory analyses on blood samples stored at –80°C since the end of the 1980s, and we are finding levels of carotenoids and steroid hormones that correspond with what you would expect in fresh blood.

References