Cancer epidemiology in the last century and the next decade

Julian Peto

Institute of Cancer Research, Sutton, Surrey SM2 5NG, and London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

By the early 1980s, epidemiologists had identified many important causes of cancer. They had also proposed the 'multi-stage' model of cancer, although none of the hypothesized events in human carcinogenesis had then been identified. The remarkable advances in cell and molecular biology over the past two decades have transformed the scope and methods of cancer epidemiology. There have been a few new discoveries based purely on traditional methods, and many long-suspected minor risks have been estimated more precisely. But modern epidemiological studies often depend on genetic, biochemical or viral assays that had not been developed 20 years ago.

Many types of cancer vary in incidence by more than an order of magnitude between different populations, and every type is rare in some part of the world[1]. The convergence towards local cancer rates seen among immigrants (Fig. 1) excludes a genetic explanation of these differences. By the 1960s, cancer epidemiologists had therefore concluded that most cancers are in principle preventable and many could be avoided by a suitable choice of lifestyle and environment[2]. Many specific causes of cancer are now known, the most important being smoking, obesity and a few oncogenic viruses, but a large proportion of global variation for common cancers such as breast, prostate, colon and rectum remains unexplained.

Environmental and lifestyle causes of cancer

Carcinogenic effects of tobacco

The most important discovery in the history of cancer epidemiology is the carcinogenic effect of tobacco. Lung cancer incidence increases rapidly among continuing smokers[3], so the risk is greatest in those who begin to smoke when young and continue throughout life. The large increase in male cigarette smoking in Britain during and after the First World War therefore caused an unprecedented epidemic among men born around 1900, and by 1955 the rate in British men aged under 55 was the highest in the world[4]. Over the five decades since British[5] and American[6] epidemiologists reported that 'cigarette smoking is a factor, and an important factor, in the production of carcinoma of the lung'[5], there has been a marked reduction in tar levels of British cigarettes[7] and in smoking among British men[8]. As a result, their lung cancer rate below age 55 has fallen by more than two-thirds since 1955; it is now among the lowest in the developed world and is still declining[8]. Similar changes occurred 20 years later in America, where cigarette smoking increased rapidly during the Second World War[1]. Women in most Western countries began smoking later than men and fewer have stopped, so their lung cancer rates are either still increasing or falling less rapidly[4]. Male lung cancer rates are still increasing in most developing countries and in Eastern Europe, where consumption of cigarettes remains high and in some areas is still increasing[4].

For many years the carcinogenic effects of tobacco were thought to be restricted largely to the lung, pancreas, bladder and kidney, and (synergistically with alcohol) the larynx, mouth, pharynx (except nasopharynx) and oesophagus[1]. More recent evidence indicates that several other types of cancer, of which the most important worldwide are stomach, liver and (probably) cervix, are also increased by smoking[9–11]. The relative importance of different smoking-related diseases varies widely between populations, as smoking usually multiplies the background rate due to other factors. In China, where liver cancer is common, smoking causes more premature deaths from liver cancer than from heart disease (ref. 11 and Fig. 2). The overall proportion of male cancer deaths caused by smoking in China in 1990 was 22% and rising[11], whereas that in Britain fell from 44% in 1990 to about 36% by 2000[4,12].

The effect of diet and overweight

Dietary epidemiology is notoriously complex owing to the variety of foods and their many constituents and to intercorrelations and temporal changes in their patterns of use. Cancer risks in old age may also depend as much on diet in early life as on current habits[13]. Apart from drinking alcohol, consumption of various foods contaminated with aflatoxin[14], and a few local customs (such as feeding Chinese-style salted fish to infants, which causes nasopharyngeal cancer[14]), no single dietary factor shows a strong and consistent enough effect to establish it unequivocally as an important carcinogen or anti-carcinogen[13]. Extensive research during the past two decades has shown that rates for various cancers correlate fairly consistently with certain aspects of diet, but opinions still differ on the strength of the evidence. Doll and Peto[10] conclude that about a third of British cancer deaths might eventually prove to be avoidable by dietary change but only those due to obesity are definitely avoidable. In contrast, an American expert panel[15] recently concluded that about a third of cancers worldwide would probably be prevented by adoption of their quantitative recommendations on daily consumption of various foods. Another British report made qualitatively similar recommendations but stated that current evidence is insufficient to determine optimal quantities[13].

There is now a consensus that cancer is commoner in those who are overweight[16]. The evidence on weight is strongest for post-menopausal breast cancer and cancers of the endometrium, gall-bladder and kidney, but several other sites contribute to the overall cancer risk[16,17]. About 5% (3% in men, 6% in women) of all incident cancers in the European Union might be prevented if no-one's body-mass index (BMI; weight divided by the square of height) exceeded

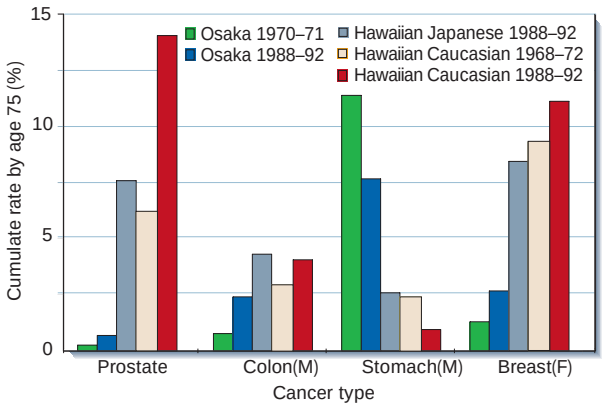

Figure 1 Cancer rates in migrants become similar to those in the local population. Cancer rates in 1990 among Japanese migrants to Hawaii, and around 1970 and 1990 in Japan (Osaka) and in Hawaiian Caucasians. Local rates for prostate, colon and breast cancer increased over time (due partly to increased completeness of diagnosis and registration, particularly for prostate cancer in Hawaiian Caucasians) and stomach cancer decreased; but the effects of migration were larger.

25 kg m^{-2} (ref. 17). Exclusion of smoking-related cancers would increase this estimate to about 7%. A large prospective cohort of non-smokers in America, where obesity is more prevalent, provides the strongest evidence on BMI and cancer mortality[18]. The authors did not calculate an attributable fraction, but their data suggest that about 10% of all cancer deaths among American non-smokers (7% in men and 12% in women) are caused by overweight[18]. It is, however, not clear how much the risk can be reduced by weight reduction in those who are already overweight. Mortality from non-malignant diseases is increased in those who are either too thin or too fat[18].

Radical changes in national dietary habits would not be easy to achieve even if there were a consensus on which foods are relevant. Dietary supplements such as vitamins or other micronutrients seem an attractive alternative, but they may not have the same effects as the foods that contain them, and some may even be harmful[13]. The only reliable way to assess their effectiveness is in large randomized trials that continue for many years. For example, there is substantial epidemiological evidence that consumption of foods containing beta-carotene correlates with reduced risk of lung cancer[13,15], but 12 years' treatment in a large randomized trial showed no benefit[19], and in two shorter trials the lung cancer risk was higher in those who received beta-carotene supplements[20,21]. Aspirin and folate supplements probably reduce colorectal cancer incidence but may take a decade or more to do so[22]. The American panel[15] concluded that various cancers were likely to be reduced by foods containing adequate amounts of carotenoids, vitamins C and E and selenium, but neither they nor the British panel[13] recommended that these micronutrients should be taken as supplements.

Reproductive and hormonal factors

The effects of reproductive factors on breast and ovarian cancer have long been assumed to reflect underlying hormonal processes[1], and this is confirmed by the effects of both endogenous[23,24] and exogenous[25] hormones. Breast cancer incidence is transiently increased by pregnancy and while oestrogens are administered as oral contraceptives or hormone replacement therapy, and is permanently lowered by late menarche, early menopause, early first childbirth and high parity[25]. Endometrial cancer incidence is also increased by hormone replacement therapy[25]. Ovarian cancer incidence declines with increasing parity[26], and both endometrial and ovarian cancers are less common in oral contraceptive users[25].

The Western diet is associated both with earlier age at menarche and with post-menopausal obesity, which increases endogenous oestrogen production and hence breast cancer risk[15]. Breast cancer incidence is much higher in most Western countries than in many developing countries, and this is partly (and perhaps largely) accounted for by these dietary effects combined with later first childbirth, lower parity and shorter breastfeeding[27,28]. The development of cancers of the testis and prostate may also depend on hormonal effects, but apart from the increased risk in an undescended testis, no behavioural or reproductive correlate is strongly predictive of these diseases[29].

Viruses, bacteria and parasites

The most important discoveries of the past two decades in cancer epidemiology relate to the carcinogenic effects of infectious pathogens that had not been characterized 20 years ago. *Helicobacter pylori*, a chronic gastric bacterial infection that can cause gastric ulcers, is a major factor in the development of stomach cancer[30]. More than 100 human papillomaviruses (HPVs) have been sequenced, and DNA from a phylogenetic subgroup of sexually transmitted HPVs that includes HPV16, HPV18 and HPV45 is detectable in virtually all cervical cancers worldwide[31]. These and other HPVs are also found in other anogenital cancers and may also cause cancers of other sites (head and neck, oesophagus and skin)[32]. The contribution of hepatitis-B virus (HBV) to liver cancer in high-incidence regions has long been recognized[33], although the synergistic effect of smoking is a more recent discovery[11]. The hepatitis-C virus (HCV) is similarly carcinogenic[33]. About one-fifth of all human cancers worldwide arise in the stomach (9%), liver (6%) or cervix (5%), and most of these would be prevented if these infections could be eradicated[34].

Other pathogens that cause a substantial cancer risk in certain populations include Epstein–Barr virus (EBV; associated with various B-cell malignancies and nasopharyngeal cancer), malaria (the major cofactor with EBV for Burkitt's lymphoma in Africa), human T-cell lymphotropic virus type 1 (some T-cell leukaemias and lymphomas), HIV (non-Hodgkin's lymphoma), human herpesvirus 8 (Kaposi's sarcoma, with HIV), schistosomiasis (bladder and colon cancer) and liver flukes (cholangiosarcoma)[30,35,36]. There is also strong epidemiological evidence for an infective aetiology in childhood leukaemia[37], but no specific pathogen has been implicated. The incidence of several virally induced cancers is further increased by specific cofactors such as dietary aflatoxin (liver), salted fish (nasopharynx) and smoking (liver

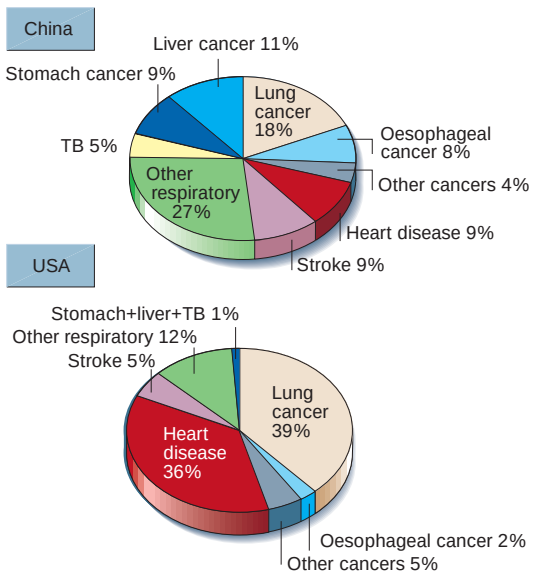

Figure 2 Smoking kills different populations in different ways. Deaths below age 70 in 1990 caused by smoking in China[11] and the United States[4]. 'Other cancers' are mouth, pharynx, larynx, bladder and pancreas.

and cervix), and if SV-40 is an important cause of mesothelioma, asbestos must also be classed as a cofactor[38].

Therapeutic immunosuppression causes a marked increase in the incidence of non-melanoma skin cancer and some virally induced cancers[39]. The discovery that many other epithelial cancers, notably lung, colon, rectum, bladder and prostate (but not breast), are also increased by immunosuppression (Table 1 and ref. 40) suggests that unidentified viruses may be important in these cancers as well. The alternative is the long-standing but equally speculative theory that many non-viral cancers are normally kept in check by immunosurveillance (see article in this issue by Rosenberg, pages 380-384).

Occupational and environmental carcinogens

About a dozen specific occupational exposures and several complex mixtures, particularly the combustion products of coal, have caused high risks of certain cancers (predominantly lung cancer) in heavily exposed workers[10]. Exposure levels for many industrial hazards have been progressively reduced in many Western countries since the 1930s, and by the late 1970s it was assumed, probably correctly, that the occupational exposure levels then current would contribute a very small proportion of future cancer incidence[1]. But uncontrolled asbestos use had been widespread in the European construction industry from the 1950s to the mid-1970s, when public concern led to a rapid reduction. The resulting epidemic of mesothelioma in building and other workers born after 1940 did not become apparent until the 1990s owing to the long latency of the disease. Incidence rates are still rising, and asbestos exposure prior to 1980 may eventually cause 250,000 mesotheliomas and a similar number of lung cancers in Western Europe[41].

This tragic episode was largely avoidable, as the carcinogenic effects of asbestos were known by 1960[42,43]; but it illustrates the major weakness of epidemiology as an early warning system. The increase in cancer incidence caused by increased exposure to a carcinogen might not be detectable for several decades, and laboratory testing must remain the first line of defence against potentially dangerous new agents, particularly those affecting endocrine or paracrine signalling that could be biologically active at very low levels. The unexplained increase in testicular cancer in many Western countries could be due to such compounds, although a dietary or viral explanation would also be plausible. The possibility of germ-cell damage is of particular concern, as environmental mutagens must cause some heritable changes, but the effect is so small that no known or suspected mutagen, including ionizing radiation, has measurably increased the frequency of germ-line mutation in humans[44]. Epidemiological studies of markers such as DNA adducts in the lung or chromosomal aberrations in lymphocytes might also provide early warning of a potential hazard. But such direct or indirect measures of mutagenic or transforming potency have never detected an important carcinogen and even today cannot provide quantitative estimates of risk.

Epidemiological data on human cancer rates still provide the only reliable evidence that the cancer risks caused by long-established activities such as working in an oil refinery or living near a high-voltage power line are not large. Apart from skin cancers due to sunlight, the only substantial and widespread cancer risk known to be caused by an avoidable environmental factor in developed countries is the further increase in lung cancer among smokers caused by indoor radon escaping from the ground or from building materials, although both indoor and outdoor air pollution from fossil fuels may also contribute to the risk in smokers[45]. The risk to non-smokers is relatively trivial in developed countries, but burning fossil fuels indoors without adequate ventilation certainly contributes to the high lung cancer rates even in non-smokers seen in parts of China[11].

Genetic epidemiology of cancer

Polymorphisms in candidate genes

There have been many studies comparing the prevalence in cancer patients and unaffected controls of common polymorphisms in genes involved in the metabolism of external or endogenous mutagens or in the production or processing of sex hormones or their analogues. A few polymorphisms in such genes seem to alter the risk substantially, such as the *N*-acetyltransferase 2 (NAT2) slow acetylator phenotype, which increases the risk of bladder cancer[46], particularly in workers heavily exposed to certain aromatic amines[47]. (Fast acetylators are not immune, however, as in one factory all 19 men who distilled β-naphthylamine developed bladder cancer[48].) But systematic meta-analysis reveals little or no effect for most such polymorphisms, and the pooled data for the minority that are statistically significant usually suggest odds ratios of less than two, and often much less[46,49–51]. Thus, for example, early reports suggested a more than doubled lung cancer risk associated with glutathione *S*-transferase μ1 (GSTM1) deficiency, but the pooled results of subsequent genotyping studies give an odds ratio of only 1.14 (95% confidence interval, 1.03–1.25)[50].

Polymorphisms in oncogenes or tumour-suppressor genes may also confer a moderately increased cancer risk. An example is the I1307K single nucleotide polymorphism (SNP) in the *APC* gene, which is carried by about 1 in 20 Ashkenazi Jews and almost doubles their colon cancer risk[52]. To estimate the individual effects of rare polymorphisms will require very large studies, but their average effect can be observed. The increased cancer risk associated with rare alleles of the HRAS1-associated minisatellite was among the first such associations reported. Such alleles, which are carried by about 5% of the population, increase the risk of several common cancers by a factor of 1.5 to 2 (ref. 53).

There have been various reports of statistically significant gene–environment interactions, such as a much larger lung cancer risk due to passive smoking in women who were GSTM1-deficient[54], or an increased breast cancer risk due to smoking in postmenopausal women that was confined to NAT2 slow acetylators[55]. In these examples, however, the estimates of the risk in susceptibles (although not their lower confidence limits) were inconsistent with the much lower overall effect of passive smoking on lung cancer[56] or of smoking on breast cancer (which is nil) in larger studies[57]. Many apparently significant gene–gene or gene–exposure interactions will arise by chance, but some will be real. The interaction between methylenetetrahydrofolate reductase and plasma folate in colorectal cancer is a plausible candidate[58]. The effects of such polymorphisms in combination with each other and with environmental risk factors

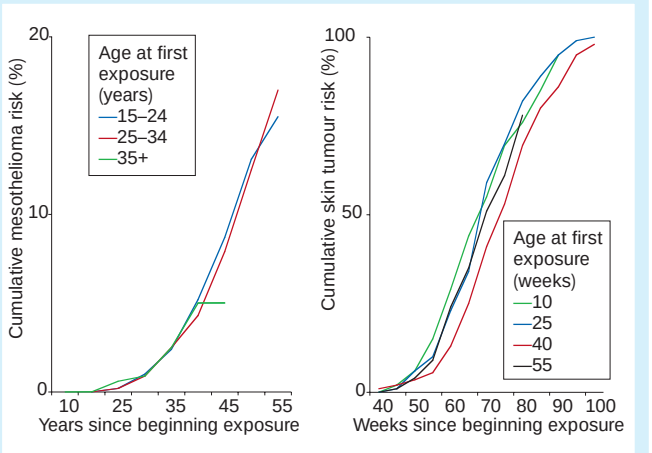

Figure 3 Age has no effect on susceptibility to some carcinogens. Left panel, cumulative mesothelioma risk in US insulation workers. Right panel, cumulative skin tumour risk in mice treated weekly with benzo(a)pyrene. Mesothelioma rates in humans[65] and skin tumour rates in mice[64] depend on time since first carcinogenic exposure but not on age, suggesting an initiating effect of these carcinogens. Lung cancer incidence in smokers depends on duration of smoking but not on age, and stops increasing when smoking stops[67], indicating both early- and late-stage effects. Radiation-induced cancer incidence increases with age at exposure above age 20, suggesting predominantly late-stage effects[3], although the large effect of childhood irradiation also indicates an early-stage effect.

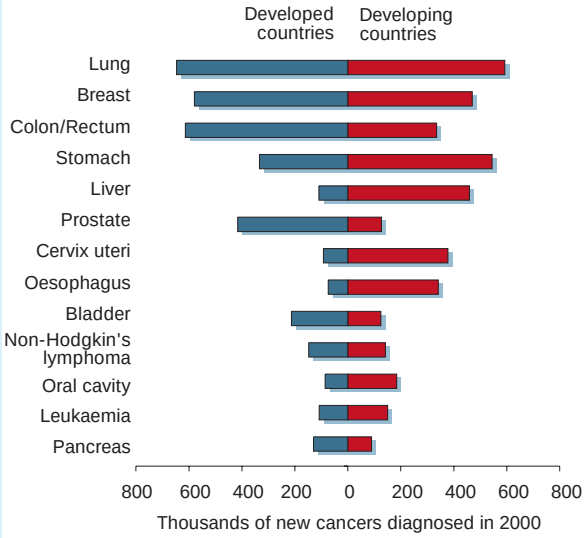

Figure 4 Global cancer incidence in developed and developing countries. (From ref. 77. Sites contributing over 2% of the 10 million cancers worldwide.)

could be substantial, but their total contribution to cancer incidence will not be known until data on risk factors and extensive genotyping are available for very large numbers of patients and controls.

Familial risks for common cancers

Highly penetrant hereditary conditions such as polyposis coli, Li-Fraumeni syndrome and familial retinoblastoma cause at most a few per cent of the majority of cancers, and from an epidemiological perspective the genetic basis of the roughly twofold increase in incidence of the same type of cancer in first-degree relatives of patients with most common cancers (ref. 59 and article in this issue by Ponder, pages 336-341) is a more important question. A mendelian gene must confer a risk an order of magnitude greater than that in non-carriers for the risk in patients' relatives to be twice that in the general population[60]. The individual effects of the common polymorphisms described above are thus far too small to account for much of this familial risk, although synergistic combinations could do so.

An important first step is to estimate the proportion of cancers of each type that arise in susceptible individuals and the contribution to this overall familial effect that can be accounted for by known genes. Breast cancer is so common that twin and sibling risks can be estimated fairly precisely. The high risk in patients' identical twins indicates that susceptible women contribute a high proportion, and perhaps even the majority, of overall breast cancer incidence[61]. This must be due mainly to 'low-penetrance' genes. Most multiple-case families[62], but only about 2% of all cases[63], are due to mutations in *BRCA1* or *BRCA2*. The genetic epidemiology of colon cancer is quantitatively similar, although there has been less extensive sequencing of known genes in unselected cases. The quantitative contribution of penetrant genes to overall cancer incidence for common cancers such as prostate, melanoma and stomach has not yet been determined. 'Low-penetrance' susceptibility conferring a site-specific lifetime risk of the order of 30–50% may underlie many cancers, but it would almost never cause large numbers of cases in a family. If many genes contribute to the large genetic effects that seem to underlie many common cancers, they may be discoverable only through advances in our understanding of carcinogenic mechanisms.

Mechanisms of carcinogenesis

Age-incidence patterns for non-hormone-dependent carcinomas, and the effects of timing and dose-level of various agents alone and in combination (particularly smoking, alcohol, ionizing radiation and some occupational carcinogens), are parsimoniously explained by the 'multi-stage' model of carcinogenesis. The evidence underlying this early work, which preceded the identification of any of the hypothesized sequence of heritable events in human carcinogenesis[3], seems sometimes to be rather neglected. For example, the epidemiological and experimental evidence that somatic ageing processes per se play little or no role in carcinogenesis (Fig. 3 and refs 3, 64, 65) was not discussed in a recent review on cancer and ageing that argued exactly the opposite[66].

The incidence rate of cancer is presumably proportional both to the rate of the final rate-limiting step in carcinogenesis and to the number of premalignant cells that have undergone all but this final step. The rapid increase in the lung cancer incidence rate among continuing smokers ceases when they stop smoking, the rate remaining roughly constant for many years in ex-smokers[67]. The fact that the rate does not fall abruptly when smoking stops indicates that the mysterious final event that triggers the clonal expansion of a fully malignant bronchial cell is unaffected by smoking, suggesting a mechanism involving signalling rather than mutagenesis. Such data are still generating new mechanistic hypotheses[61,68].

The future of cancer epidemiology

Over the next decade, cancer epidemiologists will be increasingly preoccupied with genetically susceptible subgroups. Comparison of the DNA in cancerous and normal cells from the same patient may lead directly to the identification of most of the genes that are commonly mutated in carcinogenesis. Candidate genes are also being identified on the basis of structural homologies from the human genome sequence. Extensive sequence or SNP comparisons between affected relatives and between cancer patients and controls may define combinations of polymorphisms or inherited defects in such genes that identify a few percent of the population whose average lifetime risk may be as high as 50% for a particular cancer. An alternative possibility is that susceptibility genes underlying phenotypic characteristics such as mammographic density[69] or chromosomal instability[70] that correlate with cancer risk and exhibit mendelian segregation will be found by linkage. Genes involved in DNA repair are likely to prove particularly important. Assays for defective DNA repair correlate consistently with substantially increased susceptibility[71], and chromosomal aberrations predict increased cancer risk irrespective of carcinogenic exposure[72].

Once they are identified, susceptible people might benefit disproportionately from screening or prophylaxis, while those at low risk would be reassured. But there will also be penalties. A different susceptible minority will be identified for each disease, and a high proportion of the population may eventually suffer the consequences of being classed as genetically susceptible to some major risk. The hazards of screening for cancer susceptibility are illustrated by the

Table 1 Cancer incidence in immunosuppressed transplant patients*

Cancer site	No. of cases	Expected no.	Ratio
Non-melanoma skin	127	5.1	24.7
Thyroid and other endocrine	30	2.1	14.3
Mouth, tongue and lip	22	1.6	13.8
Cervix, vulva and vagina	39	3.6	10.8
Non-Hodgkin's lymphoma	25	2.4	10.3
Kidney and ureter	32	3.5	9.1
Bladder	26	4.7	5.5
Colorectal	38	10.5	3.6
Lung	30	12.5	2.4
Brain	10	4.1	2.4
Prostate	11	5.2	2.1
Melanoma	7	4.1	1.7
Breast	15	13.6	1.1

*Source: ref. 40.

widespread introduction of testing for prostate-specific antigen in the United States, which has reduced prostate cancer mortality only marginally but has led to a sharp increase in recorded incidence and considerable post-operative psychosexual and physical morbidity. Striking gene–environment interactions may be discovered, but most causes of cancer are likely to increase the risk by a smaller amount but a similar factor in those who are less susceptible. If smokers are less likely to stop smoking on discovering that their lifetime lung cancer risk is 'only' 10%, the population death rate might even be increased by such knowledge.

Advances in genetic and molecular understanding will increasingly enable epidemiologists to quantify the relationships between risk factors and specific events in carcinogenesis. Direct monitoring of changes in the genes that underlie carcinogenesis or their products is likely to provide sensitive and specific measures that can be correlated both with cancer incidence and with exposure to carcinogenic agents or activities. Characteristic mutations in DNA from subclinical cancers[73] or their precursor lesions[74,75] can already be quantified, and serum levels of hormones such as oestrogen[23] and prolactin[24], or growth factors such as insulin-like growth factor-I, as well as chromosomal damage itself[72], are predictive of increased risk for certain cancers.

The most significant developments in cancer epidemiology may result from discoveries in virology and tumour immunology. The speculation that unidentified viruses (perhaps including some animal viruses[76]) are associated with many human cancers is consistent with the large increase in overall cancer rates seen in immunosuppressed patients[40]. The difficulty is that an unknown virus might mimic the epidemiological effects of dietary or genetic mechanisms. Thus, for example, the migrant patterns for prostate cancer (Fig. 1) might be due partly to an endemic infection, as they are for stomach cancer[30]. Viruses usually act synergistically with other carcinogens and therefore provide alternative approaches to risk reduction. Perhaps the best way to prevent mesothelioma following heavy asbestos exposure will be by targeting SV-40 (ref. 38). The crucial issue is which of the increased risks in Table 1 reflect an unknown viral aetiology and which reflect immunosurveillance targeted at non-viral tumour markers. Some cancers may well be preventable by vaccination with tumour-specific antigens or by some less specific immunostimulation (see articles in this issue by Rosenberg, pages 380-384, and Appelbaum, pages 385-389).

Current priorities in cancer prevention

The large differences in the pattern of cancer incidence between developed and developing countries (Fig. 4)[77] imply different priorities for prevention, but at an individual level the most important difference is between smokers and non-smokers, particularly in developed countries. Table 2 shows approximate percentages of future cancer deaths in the United States that would be avoided by successively removing the effects of smoking, known infections, alcohol, sunlight, current occupational and environmental pollution, inactivity and obesity. The additional effect of specific dietary recommendations such as those of the American panel[15] is much more speculative. Avoidance of overweight and prevention or treatment of oncogenic infections are the most important aims for non-smokers; but it is absurd for smokers in the West to worry about anything except stopping smoking.

Tobacco causes one-third of all cancer deaths in developed countries. About 15% of cancers worldwide are caused by known infectious agents[34]. HBV alone causes almost as many cancers as smoking in China, and can be prevented by vaccination. HPV vaccines that are already being tested may be able to prevent almost all cervical cancers[78], and if the prevalence of *Helicobacter pylori* can be reduced, many stomach cancers would be avoided. The belated elimination of asbestos by many Western countries will eventually prevent the great majority of mesotheliomas and many lung cancers. (Whether almost all mesotheliomas are caused by crocidolite, amosite or tremolite is still contentious[79,80], but all forms of asbestos cause lung cancer.) Various cancer screening tests are partially effective, and cervical screening is very effective.

Table 2 US Cancer deaths that would be avoided by eliminating known risks

Cause	Deaths (%) avoided after removing preceding causes	
	Current smokers*	Non-smokers
Smoking	60	—
Known infections†	2	5
Alcohol‡	0.4	1
Sunlight	0.4	1
Air pollution§	0.4	1
Occupation‖	0.4	1
Lack of exercise¶	0.4	1
Diet#		
Overweight (BMI>25 kg m^{-2})	4	10
Other dietary factors	4–12?	10–30?
Presently unavoidable☆	About a quarter	At least half

Approximate percentages of future cancer deaths among smokers and non-smokers in the United States that would be prevented by successive elimination of smoking, known infections, alcohol, sun exposure, current levels of workplace and air pollution, lack of exercise, and obesity. The estimates for the residual effect of specific dietary factors in people of normal weight are much less certain.

*For causes other than smoking, the table shows the additional reductions that could be achieved after smoking has been eliminated, so synergism with smoking is ignored. About 60% of cancers among smokers are due to smoking[4], so all other percentages for smokers are 40% of those for non-smokers. The percentages for smokers relate to cancer risks during a smoker's lifetime and ignore risks during the 15 years of extra life gained on average when a smoking-related death is prevented by stopping smoking.

†All cervical cancers (HPV) plus half of all stomach cancers (*H. pylori*) gives 4% for non-smokers. A further 1% is added for all other infections. These two cancers and those caused by EBV, HBV and HCV are much more common in many developing countries. HBV alone, which can be prevented by vaccination, causes almost as many cancers as smoking in China. About 15% of cancers worldwide would be prevented if all known carcinogenic infections were eliminated[34].

‡Alcohol contributes about 0.5% due to upper aerodigestive cancers in non-smokers and a similar number due to breast cancer, although the reduction in heart disease among moderate drinkers outweighs their increased cancer risk.

§The lung cancer risks caused by radon and smoke from fossil fuels are largely confined to smokers. But lung cancer accounts for 2% of all cancer deaths in non-smokers, and a substantial proportion may be due to indoor and outdoor air pollution, including indoor radon. Mesotheliomas caused by non-occupational asbestos exposure contribute less than 0.1% of cancer mortality.

‖Construction and maintenance workers are still sometimes exposed to asbestos and some workers are quite heavily exposed to dusts containing crystalline silica. Occupational exposure to chemicals causes some bladder and other cancers. Current exposures to asbestos and other known occupational carcinogens are much less than they were 40 years ago in most developed countries, and occupational lung cancer risks are much lower in non-smokers. Heavy exposure to occupational carcinogens still occurs in many developing countries.

¶As well as preventing heart disease, regular exercise reduces the risk of colon cancer and probably also of breast cancer, although the quantitative effect is uncertain[10,15].

#Mortality from various cancers and other causes increases progressively with increasing body-mass index (BMI) above 25 kg m^{-2}, and overall mortality also increases as BMI falls below about 22 kg m^{-2} (refs 17, 18). The additional benefits of dietary recommendations for those with BMI between 22 and 25 kg m^{-2} are very uncertain.

☆A few per cent of cancers are avoidable by the combined effects of measures that most people would not consider, such as having many children at an early age, followed by oophorectomy (surgical removal of one or both ovaries). Modern chemotherapy, radiotherapy and radiography usually do more good than harm, and cosmic radiation is unavoidable. The presumably environmental causes of the increased incidence of testicular cancer and non-Hodgkin's lymphoma are unavoidable because they are not yet known.

The threat to epidemiology of the new ethics

Government action is essential to protect epidemiological research from the increasing burden of 'ethical' laws or conventions that bear no relation to patients' physical or psychological well-being. Examples in Britain include the recent directive by the General Medical Council that doctors who notify their patients to cancer registries without obtaining fully informed consent may face disciplinary action, and the earlier delay in introducing anonymous HIV testing of discarded blood samples on the grounds that 'screening must confer some benefit on the patient'[81]. Under the Data Protection Act it may even be illegal to use historical personnel records to study the

mortality of factory workers, as it is impractical to obtain informed consent from all former workers. Legislation is urgently needed to restore the long-established principle that consent is not mandatory for access to medical or civil records for bona fide medical research that has no effect on the individuals concerned and has been approved by a competent ethics committee. □

1. Doll, R. & Peto, R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J. Natl Cancer Inst.* **66,** 1191–1308 (1981).
2. World Health Organization. *Prevention of Cancer* (*WHO Tech. Rep.* 276) (WHO, Geneva, 1964).
3. Doll, R. An epidemiological perspective of the biology of cancer. *Cancer Res.* **38,** 3573–3583 (1978).
4. Peto, R., Lopez, A. D., Boreham, J., Heath, C. & Thun, M. *Mortality from Tobacco in Developed Countries, 1950-2000* (Oxford Univ. Press, Oxford, 1994).
5. Doll, R. & Hill, A. B. Smoking and carcinoma of the lung. Preliminary report. *Br. Med. J.* **ii,** 739–748 (1950).
6. Wynder, E. L. & Graham, E. A. Tobacco smoking as a possible etiologic factor in bronchogenic carcinoma. *J. Am. Med. Assoc.* **143,** 329–336 (1950).
7. Wald, N., Doll, R. & Copeland, G. Trends in tar, nicotine, and carbon monoxide yields of UK cigarettes manufactured since 1934. *Br. Med. J. (Clin. Res. Ed.)* **282,** 763–765 (1981).
8. Peto, R. *et al.* Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. *Br. Med. J.* **321,** 323–329 (2000).
9. Doll, R. Cancers weakly related to smoking. *Br. Med. Bull.* **52,** 35–49 (1996).
10. Doll, R. & Peto, R. in *Oxford Textbook of Medicine* 4th edn (eds Warrell, D. A., Cox, T. M. & Firth, J. D.) (Oxford Medical Publications, Oxford, in the press).
11. Liu, B. Q. *et al.* Emerging tobacco hazards in China: 1. Retrospective proportional mortality study of one million deaths. *Br. Med. J.* **317,** 1411–1422 (1998).
12. Peto, R., Lopez, A. D., Boreham, J., Thun, M. & Heath, C. Mortality from tobacco in developed countries: indirect estimation from national vital statistics. *Lancet* **339,** 1268–1278 (1992).
13. Working Group on Diet and Cancer of the Committee on Medical Aspects of Food and Nutrition Policy. *Nutritional Aspects of the Development of Cancer* (Dept Health Rep. Health Social Subjects 48) (The Stationery Office, London, 1998).
14. IARC. *Some Naturally Occurring Substances: Food Items and Constituents, Heterocyclic Aromatic Amines and Mycotoxins* (*IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* 56) (IARC, Lyon, 1993).
15. American Institute for Cancer Research/World Cancer Research Fund. *Food, Nutrition and the Prevention of Cancer: A Global Perspective* (American Institute for Cancer Research, Washington, 1997).
16. Josefson, D. Obesity and inactivity fuel global cancer epidemic. *Br. Med. J.* **322,** 945 (2001).
17. Bergstrom, A., Pisani, P., Tenet, V., Wolk, A. & Adami, H. O. Overweight as an avoidable cause of cancer in Europe. *Int. J. Cancer* **91,** 421–430 (2001).
18. Calle, E. E., Thun, M. J., Petrelli, J. M., Rodriguez, C. & Heath, C. W. Body-mass index and mortality in a prospective cohort of U.S. adults. *N. Engl. J. Med.* **341,** 1097–1105 (1999).
19. Hennekens, C. H. *et al.* Lack of effect of long-term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease. *N. Engl. J. Med.* **334,** 1145–1149 (1996).
20. Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group. The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. *N. Engl. J. Med.* **330,** 1029–1035 (1994).
21. Omenn, G. S. *et al.* Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N. Engl. J. Med.* **334,** 1150–1155 (1996).
22. Janne, P. A. & Mayer, R. J. Chemoprevention of colorectal cancer. *N. Engl. J. Med.* **342,** 1960–1968 (2000).
23. Hankinson, S. E. *et al.* Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. *J. Natl Cancer Inst.* **90,** 1292–1299 (1998).
24. Hankinson, S. E. *et al.* Plasma prolactin levels and subsequent risk of breast cancer in postmenopausal women. *J. Natl Cancer Inst.* **91,** 629–634 (1999).
25. IARC. *Hormonal Contraception and Post-Menopausal Hormonal Therapy* (*IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* 72) (IARC, Lyon, 1999).
26. Casagrande, J. T. *et al.* 'Incessant ovulation' and ovarian cancer. *Lancet* **2,** 170–173 (1979).
27. Pike, M. C., Krailo, M. D., Henderson, B. E., Casagrande, J. T. & Hoel, D. G. 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. *Nature* **303,** 767–770 (1983).
28. Newcomb, P. A. *et al.* Lactation and a reduced risk of premenopausal breast cancer. *N. Engl. J. Med.* **330,** 81–87 (1994).
29. Swerdlow, A. J. in *Principles and Practice of Genitourinary Oncology* (eds Raghavan, D., Scher, H. I., Leibel, S. A. & Lange, P.) 643–652 (Lippincott-Raven, New York,1997).
30. IARC. *Schistosomes, Liver Flukes and Heliobacter pylori* (*IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* 61) (IARC, Lyon, 1994).
31. Walboomers, J. M. *et al.* Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.* **189,** 12–19 (1999).
32. IARC. *Human Papillomaviruses* (*IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* 64) (IARC, Lyon, 1995).
33. IARC. *Hepatitis Viruses* (*IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* 59) (IARC, Lyon, 1994).
34. Parkin, D. M., Pisani, P., Munoz, N. & Ferlay, J. The global health burden of infection associated cancers. *Cancer Surv.* **33,** 5–33 (1999).
35. IARC. *Epstein-Barr Virus and Kaposi's Sarcoma Herpesvirus/Human Herpesvirus 8* (*IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* 70) (IARC, Lyon, 1997).
36. IARC. *Human Immunodeficiency Viruses and Human T-cell Lymphotropic Viruses* (*IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* 67) (IARC, Lyon, 1996).
37. Kinlen, L. J. Infection and childhood leukemia. *Cancer Causes Cont.* **9,** 237–239 (1998).
38. Carbone, M., Rizzo, P. & Pass, H. Simian virus 40: the link with human malignant mesothelioma is well established. *Anticancer Res.* **20,** 875–877 (2000).
39. Kinlen, L. J., Sheil, A. G., Peto, J. & Doll, R. Collaborative United Kingdom-Australasian study of cancer in patients treated with immunosuppressive drugs. *Br. Med. J.* **2,** 1461–1466 (1979).
40. Birkeland, S. A. *et al.* Cancer risk after renal transplantation in the Nordic countries, 1964-1986. *Int. J. Cancer* **60,** 183–189 (1995).
41. Peto, J., Decarli, A., La Vecchia, C., Levi, F. & Negri, E. The European mesothelioma epidemic. *Br. J. Cancer* **79,** 666–672 (1999).
42. Wagner, J. C., Sleggs, C. A. & Marchand,P. Diffuse pleural mesothelioma and asbestos exposure in the North Western Cape Province. *Br. J. Ind. Med.* **17,** 260–271 (1960).
43. Doll, R. Mortality from lung cancer in asbestos workers. *Br. J. Ind. Med.* **12,** 81–86 (1955).
44. Venitt, S. & Phillips, D. H. in *Environmental Mutagenesis* (eds Phillips, D. H. & Venitt, S.) 1–20 (Bios Scientific, Oxford, 1995).
45. Dockery, D. W. *et al.* An association between air pollution and mortality in six U.S. cities. *N. Engl. J. Med.* **329,** 1753–1759 (1993).
46. Hein, D. W. *et al.* Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. *Cancer Epidemiol. Biomarkers Prev.* **9,** 29–42 (2000).
47. Cartwright, R. A. *et al.* Role of N-acetyltransferase phenotypes in bladder carcinogenesis: a pharmacogenetic epidemiological approach to bladder cancer. *Lancet* **2,** 842–845 (1982).
48. Doll, R. *Possibilities for the Prevention of Cancer* (Royal Society Lecture for the Public, 13 November 1986) (Royal Society, London).
49. Johns, L. E. & Houlston, R. S. Glutathione S-transferase μ1 (GSTM1) status and bladder cancer risk: a meta-analysis. *Mutagenesis* **15,** 399–404 (2000).
50. Houlston, R. S. Glutathione S-transferase M1 status and lung cancer risk: a meta-analysis. *Cancer Epidemiol. Biomarkers Prev.* **8,** 675–682 (1999).
51. Dunning, A. M. *et al.* A systematic review of genetic polymorphisms and breast cancer risk. *Cancer Epidemiol. Biomarkers Prev.* **8,** 843–854 (1999).
52. Woodage, T. *et al.* The APCI1307K allele and cancer risk in a community-based study of Ashkenazi Jews. *Nature Genet.* **20,** 62–65 (1998).
53. Krontiris, T. G., Devlin,B., Karp, D. D., Robert, N. J. & Risch, N. An association between the risk of cancer and mutations in the HRAS1 minisatellite locus. *N. Engl. J. Med.* **329,** 517–523 (1993).
54. Bennett, W. P. *et al.* Environmental tobacco smoke, genetic susceptibility, and risk of lung cancer in never-smoking women. *J. Natl Cancer Inst.* **91,** 2009–2014 (1999).
55. Ambrosone, C. B. *et al.* Cigarette smoking, N-acetyltransferase 2 genetic polymorphisms, and breast cancer risk. *J. Am. Med. Assoc.* **276,** 1494–1501 (1996).
56. Weinberg, C. R. & Sandler, D. P. Gene-by-environment interaction for passive smoking and glutathione S-transferase M1? *J. Natl Cancer Inst.* **91,** 1985–1986 (1999).
57. IARC. *Tobacco Smoking* (*IARC Monographs on the Evaluation of Carcinogenic Risks to Humans,* 38) (IARC, Lyon, 1986).
58. Ma, J. *et al.* Methylenetetrahydrofolate reductase polymorphism, dietary interactions, and risk of colorectal cancer. *Cancer Res.* **57,** 1098–1102 (1997).
59. Houlston, R. S. & Peto, J. in *Genetic Predisposition to Cancer* (eds Eeles, R. A., Ponder, B. A. J., Easton, D. F. & Horwich, A.) 208–226 (Chapman & Hall, London, 1996).
60. Peto, J. in *Cancer Incidence in Defined Populations. Banbury Report 4* (eds Cairns, J., Lyon, J. L. & Skolnick, M.) 203–213 (Cold Spring Harbor Laboratory Press, 1980).
61. Peto, J. & Mack, T. M. High constant incidence in twins and other relatives of women with breast cancer. *Nature Genet.* **26,** 411–414 (2000).
62. Ford, D. *et al.* Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am. J. Hum. Genet.* **62,** 676–689 (1998).
63. Peto, J. *et al.* Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *J. Natl Cancer Inst.* **91,** 943–949 (1999).
64. Peto, R., Parish, S. E. & Gray, R. G. in *Age-Related Factors in Carcinogenesis* (*IARC Sci. Publ.* 58) (eds Likhachev, A., Anisimov, V. & Montesano, R.) 43–53 (IARC, Lyon, 1986).
65. Peto,J. in *Models, Mechanisms and Aetiology of Tumour Promotion* (*IARC Sci. Publ.* 56) (eds Borzsonyi, M., Day, N. E., Lapis, K. & Yamasaki, H.) 359–371 (IARC, Lyon, 1984).
66. DePinho, R. A. The age of cancer. *Nature* **408,** 248–254 (2000).
67. Halpern, M. T., Gillespie, B. W. & Warner, K. E. Patterns of absolute risk of lung cancer mortality in former smokers. *J. Natl Cancer Inst.* **85,** 457–464 (1993).
68. Janin, N. A simple model for carcinogenesis of colorectal cancers with microsatellite instability. *Adv. Cancer Res.* **77,** 189–221 (2000).
69. Pankow, J. S. *et al.* Genetic analysis of mammographic breast density in adult women: evidence of a gene effect. *J. Natl Cancer Inst.* **89,** 549–556 (1997).
70. Roberts, S. A. *et al.* Heritability of cellular radiosensitivity: a marker of low-penetrance predisposition genes in breast cancer? *Am. J. Hum. Genet.* **65,** 784–794 (1999).
71. Berwick, M. & Vineis, P. Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. *J. Natl Cancer Inst.* **92,** 874–897 (2000).
72. Bonassi, S. *et al.* Chromosomal aberrations in lymphocytes predict human cancer independently of exposure to carcinogens. European Study Group on Cytogenetic Biomarkers and Health. *Cancer Res.* **60,** 1619–1625 (2000).
73. Anker, P., Mulcahy, H., Chen, X. Q. & Stroun, M. Detection of circulating tumour DNA in the blood (plasma/serum) of cancer patients. *Cancer Metastasis Rev.* **18,** 65–73 (1999).
74. Ponten, F. *et al.* Molecular pathology in basal cell cancer with p53 as a genetic marker. *Oncogene* **15,** 1059–1067 (1997).
75. Khrapko, K. *et al.* Constant denaturant capillary electrophoresis (CDCE): a high resolution approach to mutational analysis. *Nucleic Acids Res.* **22,** 364–369 (1994).
76. zur Hausen, H. Proliferation-inducing viruses in non-permissive systems as possible causes of human cancers. *Lancet* **357,** 381–384 (2001).
77. Parkin, D. M., Bray, F. I. & Devesa, S. S. Cancer burden in the year 2000. The global picture. *Eur.J.Cancer* (special supplement; in the press).
78. Schiller, J. T. & Lowy, D. R. Papillomavirus-Like Particle Vaccines. *J. Natl Cancer Inst. Monogr.* **28,** 50–54 (2000).
79. Cullen, M. R. Chrysotile asbestos: enough is enough. *Lancet* **351,** 1377–1378 (1998).
80. McDonald, A. D. *et al.* Mesothelioma in Quebec chrysotile miners and millers: epidemiology and aetiology. *Ann. Occup. Hyg.* **41,** 707–719 (1997).
81. Black, D. *et al.* HIV testing on all pregnant women. *Lancet* **2,** 1277 (1987).

Acknowledgements

Parts of the section on environmental and lifestyle factors are based on the review on avoidable causes of cancer by Richard Doll and my brother Richard Peto[1] and their updated summary[10]. I am also grateful to them and to many other colleagues for commenting on earlier drafts, and to the Cancer Research Campaign for support.