Carcinogenesis vol.23 no.3 pp.381–387, 2002

COMMENTARY

Gene–environment interaction and aetiology of cancer: what does it mean and how can we measure it?

Paul Brennan

Unit of Environmental Cancer Epidemiology, International Agency for Research on Cancer, 150 cours Albert Thomas, F-69008 Lyon, France

Email: brennan@iarc.fr

One form of defence against cancer development involves a series of genes whose role is to metabolize and excrete potentially toxic compounds and to repair subtle mistakes in DNA. Much laboratory and epidemiological research over the past decade has concentrated on the identification of these genes and an assessment of their role in cancer aetiology. Of particular interest has been whether the risk of cancer associated with a particular environmental exposure differs with respect to functionally different polymorphisms of these genes, i.e. gene–environment interaction. A large number of studies have been conducted for numerous genes and also for all common cancer sites, although results have been very inconsistent and therefore inconclusive. This is partially due to the inadequate sample size of most studies to detect modest effects and the over-reporting of positive associations identified in subgroups of the dataset. There is also much confusion about the meaning of 'gene–environment interaction', what type of studies should be conducted to study it and also how it should be measured. Furthermore, the very purpose of these studies is not clear; are they attempting to identify high-risk individuals, or are they simply trying to further understand the cancer process? This review will explore these questions and provide some recommendations to help ensure that the next phase of gene–environment interaction studies, which are likely to be much larger and based on many more genes, also provide some clearer answers.

Introduction

A malignant tumour is the result of a series of DNA alterations in a single cell, or clones of that cell, which lead to loss of normal function, aberrant or uncontrolled cell growth and often metastases. Several of the genes, which are frequently lost or mutated, have been identified including those whose function is to induce cell proliferation under specific circumstances (e.g. ras and myc proto-oncogenes) and genes which are programmed to halt proliferation in damaged cells (e.g. p53 and APC tumour suppressor genes). Other mutations are also necessary, including in genes involved in DNA repair, cell-cycle control, angiogenesis and telemorase production. The pattern of losses and mutations is complex, although mutation or loss of at least one proto-oncogene and one or more tumour suppressor genes in a single cell, resulting in uncontrolled and unchecked cellular proliferation, is likely to occur in nearly all tumours (1).

With the exception of rare familial cancers which are primarily caused by a germline inheritance of a specific mutation, a sporadic cancer may acquire mutations as a result of genotoxic exposure to external or internal agents (such as tobacco carcinogens, dietary factors, infectious agents and sex hormones) and consequent DNA adduct formation. The likelihood of a mutation occurring and persisting in subsequent clones may be heavily dependant on the efficiency with which potentially toxic exposures are metabolized and excreted, and also the efficiency with which small mistakes in DNA replication are rectified. It is this 'caretaker' role of carcinogenesis that is likely to vary strongly between individuals because of the population variability in polymorphic genes that regulate these processes. Much effort has been focused on the identification of these low risk but highly prevalent genes, although success has been very limited. This review will explore some of the main problem areas in the search for low-risk cancer genes, and propose some solutions for future studies.

Abbreviations: OR, odds ratio; PAHs, polycyclic aromatic hydrocarbons; UC, ulcerative colitis.

Which common polymorphisms may modify the multistage cancer process?

Some examples of genes that are currently attracting interest are included in Table I, and include genes that are responsible for conversion of exogenous exposures into intermediate metabolites (phase 1 enzymes), notably genes of the cytochrome P450 family (2). Parodoxically, most of these intermediate metabolites are highly reactive with DNA and are responsible for adduct development and subsequent DNA mutations, and their metabolization and excretion is necessary via a further group of phase II enzymes including those from the glutothione and *N*-acetyltransferase families (3,4). Another potentially important source of inter-individual variability in relation to the development of cancer is DNA repair capacity (5). There is substantial variation in DNA repair capacity between individuals although it appears to vary little between identical twins (6), indicating that it is largely under genetic control. A fourth group of cell-cycle control genes includes tumour suppresser genes such as p53 codon 72 polymorphism, and genes thought to regulate apoptosis such as Cyclin D1 (CCND1). Some DNA repair genes such as XPD may also be involved in regulating apoptosis (7) although the extent of functional differences in most cell-cycle control and DNA repair polymorphisms is currently unclear. Similarly, genes involved in regulation and development of the immune system may be important for viral and haematopoetic neoplasms, although little data are available. Finally, genes that influence behaviour, such as smoking, alcohol consumption and excess calorie intake have the potential to substantially affect cancer risk. For example, several genes have been suggested to be associated with the ability to quit smoking, including CYP2A6 and the dopamine D2 receptor gene (8,9).

Why are we interested in genes with a low risk?

Functional variation in the families of genes listed in Table I is likely to have a subtle effect on cancer risk for an individual,

© Oxford University Press

Table I. Preliminary list of candidate genes that may influence the risk of developing various cancers

Type of gene	Gene
Phase I polymorphisms	CYP1A1, CYP1A2, CYP2A6, CYP2D6, CYP2E1, ADH2, ADH3, MPO mEH
Phase II polymorphisms	GSTM1, GSTT1, GSTP1, NAT1, NAT2, ALDH2, NQO1, SULT1A1, SOD2
DNA repair genes	XRCC1, XRCC3, XPD, XPF, ERCC1
Immune function genes	IL1A, IL1B, IL2, IL6, TNF, HLA Class I/II
Cell-cycle control genes	TP53, HRAS
Nicotine addiction and other receptor genes	CYP2A6, DAT1, DRD2, DRD4, RARA

Table II. Population attributable risk (PAR) and sample size requirements for various combinations of OR and prevalence of polymorphism (P)

Prevalence	OR				
	1.25	1.5	2.0	5.0	25
50%	**11**	**20**	**33**	**66**	**92**
	1267	*387*	*136*	*30*	*13*
20%	**4.8**	**9.1**	**16.7**	**44.0**	**82.8**
	1850	*535*	*172*	*28*	*8*
5%	**1.2**	**2.4**	**4.8**	**16.6**	**54.5**
	6020	*1689*	*516*	*69*	*11*

Population attributable risk (in bold) and sample size (in italic) required to detect each OR assuming a power of 80% and $P < 0.05$.

but may have a large population impact because the relevant polymorphisms may be highly prevalent. For example, a polymorphism which increases risk by only 50% but is also present in half of the population would account for 20% of all cases, similar to a high-risk gene with an increased risk of 5-fold which is present in only 5% of the population (Table II).

Identifying low-risk cancer genes will also be important for increasing our knowledge of carcinogenicity. Most exogenous and endogenous exposures (e.g. diet, tobacco, alcohol, air pollution and hormones) are complex mixtures and we are still unclear about the process by which they exert their carcinogenic effect. Identifying metabolizing and DNA repair genes involved with these exposures will help to clarify the process by which a cancer develops, and may thus indirectly lead to prevention. For example, exposure to polycyclic aromatic hydrocarbons (PAHs) may be important causes of colon cancer, with possible exposure mainly from tobacco consumption and dietary consumption of charred meat (10). The glutathione *S*-transferase M1 and T1 genes code for enzymes which may be involved in PAH metabolism and a role for these genes in colon cancer would enhance the credibility of a causal association with PAHs, especially if this association was restricted to smokers, or those who were likely to consume well cooked meat. A similar example involves the alcohol and aldehyde dehydrogenase genes (mainly ADH2, ADH3 and ALDH2) and their possible role with oro-pharyngeal and laryngeal cancer. ADH2 and ADH3 are primarily responsible for the metabolization of ethanol to acetaldehyde and the speed at which the metabolization occurs is dependant upon various ADH alleles, with ADH2*2 and ADH3*1 alleles representing fast metabolizers (11). If an association between head and neck cancers and these ADH alleles is confirmed this will provide strong evidence that the metabolite acetaldehyde, is the causative agent in alcohol-related carcinogenesis, and provide evidence against the main alternative hypotheses (e.g. that alcohol acts merely as a solvent for tobacco-related carcinogens).

Will identification of many low-risk genes allow identification of high-risk individuals?

Although of importance on the population level, another argument to justify the identification of susceptibility genes is that they may also be useful in identifying individuals who are at a very high cancer risk because of their increased genetic susceptibility in response to a genotoxic exposure, or because they have inherited several low-risk types whose combined effect results in a high risk. The implications of such findings are unclear although one assumption is that they could lead to chemoprevention programmes directed to high-risk individuals, perhaps to enable them to quit smoking, reduce alcohol consumption, adopt dietary changes or avoid particular occupational exposures. This strategy is problematic on both pragmatic and theoretical grounds. On a pragmatic level, as was pointed out by Rose (12), modest behavioural changes which occur on a population level are likely to lead to a greater public health impact than larger behavioural changes which are restricted to a small high-risk group. On a theoretical level, it is far from certain that even complete knowledge of all genes potentially involved in cancer development would allow identification of population subgroups at a particularly high risk if exposed to a particular carcinogen. An alternative viewpoint is that sporadic cancers may have a strong stochastic element, and a deterministic approach that attempts to accurately predict risk based on complete knowledge of genetic susceptibility may not be appropriate.

The inherent stochastic nature of cancer is supported by several observations in animal and twin studies. Genetically identical animals kept in as similar an environment as possible will not behave the same upon exposure to environmental carcinogens. While it is possible to estimate the proportion of animals that will develop a malignancy at a particular exposure level, there appears to be a random element determining which particular animals will develop tumours. The closest possible controlled situation among humans is the twin study. A recent Scandinavian study analysed 44 788 pairs of twins and reported that the correlation of sporadic cancers within monozygotic twins is not particularly striking, especially when compared with non-identical twins (13). For example, in pairs where one male twin developed a lung cancer, the other twin developed a lung cancer by the age of 75 years in 11% of monozygotic pairs and 10% of dizygotic pairs. Given that the lifetime risk of current smokers is ~15% (14), and assuming that these twins were over-represented by smokers, which is also known to correlate strongly within twins (15), the results provide little

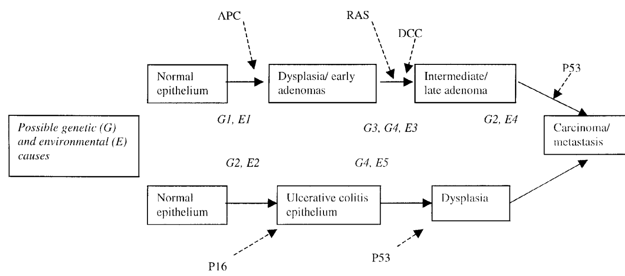

Fig. 1. Two possible causal pathways for colorectal cancer, including likely mutations and potential genetic and environmental exposures.

evidence of a genetically susceptible population at a high risk of developing lung cancer.

It is possible that the current model incorporating genetic susceptibility and environmental exposures may have to be extended to include a stochastic element, similar to the three dimensional biologic model proposed by Lewontin (16). Given previous observations from twin studies, it is also possible that the stochastic element may be just as important, if not more important, than the environmental and genetic components.

How can we measure gene–environment interaction?

A possible definition of gene–environment interaction is that it occurs when a genetic and environmental exposure work together to cause a disease outcome in some or all cases. If either exposure is absent then the disease will not occur in these cases. The assumption, which is commonly used to test for interaction, is that if the joint effect for two exposures is greater than the product of the individual effects then interaction is present. However, an alternative model of interaction, which is often ignored, is the additive model. In order to demonstrate how both of these models of interaction may be relevant in carcinogenesis, the colorectal model of tumour growth is presented in Figure 1.

Based on analysis of intestinal biopsy samples from different stages of malignant growths, Vogelstein *et al.* found that both copies of the APC tumour suppressor gene on chromosome 5 were typically mutated in dysplastic lesions and early adenomas, and the mutated RAS oncogene was often detected in more advanced adenomas. Still more advanced polyps tended to have a mutated tumour suppressor gene termed DCC, with p53 tumour suppressor genes being most commonly mutated only at the end of the carcinogenic process (17). Potential genetic and environmental exposures, which may be related to specific mutations, have been introduced into this model (e.g. E_1, E_3, E_4, G_1, G_2, G_3, G_4). A second alternative pathway for colorectal cancer also exists and is included in Figure 1, along with hypothetical environmental and genetic causes for each stage of colorectal cancer development. Ulcerative colitis (UC) is an autoimmune disease leading to chronic inflammation of the colon mucosa and an increased risk of colorectal cancer (18). The mutation pattern of UC-linked colon cancer appears to be distinct to adenoma-linked colon cancer with very few APC mutations being detected and an excess of p16 mutations, as well as a high proportion of p53 mutations being detected in UC dysplastic lesions (19). Similarly, it is probable that there are distinct causes of UC-related colorectal cancer, which are completely independent of adenoma-related colorectal cancer (an example from Figure 1 includes G1 and E1 for adenoma-related colon cancer and G2 and E5 for UC-related colon cancer). In this scenario risk factors, which appear on distinct causal pathways, are independent risk factors for the cancer and do not interact. Importantly, the combined risk for having two independent risk factors (e.g. G1 and E5 or G2 and E1) can be shown to be additive (20):

$$\text{i.e. OR(G1,E5)} = \text{OR(G1)} + \text{OR(E5)} - 1$$

A further conclusion is that joint effects, which depart from an additive model, involve exposures on the same causal pathway. For example, the exposure E1 will increase the mutation rate of the APC gene, resulting in a larger number of modified clones with APC mutations, and increasing the probability that one of these will develop further and pick up a sufficient number of mutations to become malignant. Exposure E1 will interact with all other genetic and environmental exposures which operate later in this pathway (G2, G3, G4, E3, E4) and the joint risk of E1 with any of these may be predicted by the multistage cancer model to be the product of their individual risks (21). The results provide an initial framework for testing for interaction between two causes in that a departure from an additive effect for two exposures implies that they are both involved in the same causal pathway, although possibly at different mutation sites. A measure of departure from the additive model is the relative excess risk due to interaction (RERI), which may be calculated as:

$$\text{RERI} = \text{OR(GE)} - \text{OR(E)} - \text{OR(G)} + 1$$

A third situation exists when two exposures are acting on the same mutation (e.g. E2 and G2 in the development of UC and subsequent p16 mutations). In this case the resulting joint effect is not possible to predict and may be multiplicative, greater than multiplicative or even less than multiplicative. This relationship will be dependant upon how the risk of a particular mutation changes with changing carcinogenic dose. If, as may be the case for many metabolic polymorphisms, the mutation rate is proportional to the dose, then a multiplicative effect is likely to be observed. Alternatively, a greater then multiplicative effect will be observed when the mutation rate is disproportionately increased with dose. Another typical situation where a greater than multiplicative effect may be observed is when one risk factor is completely dependant on the other for its carcinogenic potential, i.e. there is no increased risk for one risk factor in the absence of the other. Other possible scenarios are also feasible including thresholds with a less than additive effect being observed. A more extreme situation may result from exposures which are protective at one dose level and risk factors at another dose level, a possible example being beta-carotene and lung cancer. In summary, there is no reason to expect that a greater than multiplicative effect will be observed for interaction between exposures at a single mutation site and several other possibilities may be equally plausible. Also, it may not be possible to distinguish this form of interaction for a specific mutation from an interaction between causes at different sites on a pathway when a multiplicative effect between exposures is observed. However, a greater than multiplicative effect will strongly imply simultaneous interaction acting on a specific mutation (Table III).

Finally, the existence of separate causal pathways for any particular cancer, as depicted for colorectal cancer in Figure 1, is likely to be the exception rather than the rule indicating that additive interaction between established risk factors will be rare. Another possible example of separate causal pathways

Table III. Interpretation of joint OR regarding type of statistical interaction, which is present

(E)	(G)	OR	Possible interpretations of OR_{eg}
–	–	1	
+	–	$OR_{eg'}$	$OR_{eg} = OR_{eg'} + OR_{e'g} - 1$. Consistent with additive model. E and G likely to be involved in different causal pathways
–	+	$OR_{e'g}$	$OR_{eg} = OR_{eg'} * OR_{e'g}$. Consistent with multiplicative model. E and G likely to be involved in same causal pathway, possibly at different mutation site
+	+	OR_{eg}	$OR_{eg} > OR_{eg'} * OR_{e'g}$. Greater than multiplicative model. E and G likely to act simultaneously at same mutation site

(E), Environmental exposure; (G), genetic exposure; $OR_{eg'}$, OR for E in absence of G; $OR_{e'g}$, OR for G in absence of E; OR_{eg}, OR in presence of both E and G.

may include pharyngeal cancers, a high proportion of which are HPV positive and which do not appear to be related to the traditional risk factors for HPV negative pharyngeal cancers including alcohol and tobacco (22).

Why have previous studies of genetic polymorphisms provided such inconsistent results?

The literature on genetic polymorphisms and cancer is vast with hundreds of studies being published each year on possible gene–cancer relationships. The total sum of knowledge obtained from this effort is however disappointing, with very little consensus on which positive associations are likely to be true. This lack of consistency of study findings, and confusion in their interpretation, is probably due to several reasons including small sample size of studies conducted up to now, publication bias by both authors and journal editors and also over-interpretation of subgroup findings.

The average size of current case-control studies for genetic polymorphisms is in the region of 150–300 cases and a similar number of controls. As can be seen from Table II, these sample sizes would be appropriate for detecting common polymorphisms which increase risk by ~2-fold, although they are not sufficient for detecting such increased risks for rarer polymorphisms (e.g. 5%) or for common polymorphisms which increase risk by <100%. This is problematic as recent large meta-analysis of published studies indicates that when an increased risk does exist for a particular polymorphism it is likely to be very moderate. For example, two recent meta-analyses of NAT2 slow acetylation and bladder cancer risk have been conducted on 22 published studies comprising over 2000 cases and 3000 controls (23,24). Both meta-analyses found an increased risk of ~40% for slow acetylators, with a 95% confidence interval (CI) of between 20 and 60%. Given that ~40% of Caucasian populations are slow acetylators, this increased risk is important. However, none of the studies that investigated this relationship had a sufficiently large sample size to detect this. Similarly, a recent meta-analysis of the GSTM1 null polymorphism and lung cancer comprising 43 studies and over 8000 cases and 11 000 controls reported an increased risk of 17% for the GSTM1 null polymorphism, with a 95% range between 7 and 27% (S.Benhamou, W.J.Lee, A.-K.Alexandrie *et al*., submitted for publication). Again, even this small increased risk may be important on the population level because of the high proportion of the population who carry the GSTM1 null genotype. Given their biological function, NAT2 and GSTM1 were two of the strongest candidate genes for bladder and lung cancer, respectively, and future studies on common polymorphisms should anticipate that if an increased risk does exist it is likely to be in the range 20–60%, and not 100–200% as previous studies have assumed, and the sample size of future studies should be sufficiently large to detect such small risks.

The conduct of small studies may not be such a major problem if the results of all studies were published in a format that allows subsequent pooling with similar studies. Instead, what appears in the literature is known to be a biased representation of studies that have been conducted, with positive findings being more likely to be reported than negative findings (25). This is due to both selection by the investigator choosing which findings to report or simply putting off submission of negative findings because of other priorities. The problem is also partly due to journal editors who more eagerly accept for publication positive rather than negative findings. There are several possible approaches to lessening the potential for publication bias, including statistical and graphical tests when conducting meta-analysis (26), contacting investigators and asking for access to unpublished data and pooling this with published data (27) and also special sections in journals for the reporting of non-positive findings. While all of these provide useful information they cannot exclude the possibility of publication bias. Instead it is probable that the most reliable evidence will come from large single studies, or coordinated multicentre studies which test hypotheses of genetic susceptibility, and which undertake to report all genotyping results.

Another complicating factor in interpreting results from genetic studies is that many studies will report no overall positive findings with the exception of one particular subgroup, e.g. among women, among lightly exposed, among squamous cell cancer cases only, etc. Subsequent reporting of the findings may concentrate strongly on the subgroup analysis and an appropriate biological model for why any increased risk should be restricted to this group. What such reports often ignore is that if no overall effect is observed (OR = 1.0) then an increase in one group implies a similar decrease in risk in the remaining group (e.g. among men, among heavily exposured, among adenocarcinomas). Also, such findings are rarely replicated in subsequent studies. As a general rule, and especially in the absence of a strong *a priori* biological model, it is therefore prudent to establish a main effect for a particular polymorphism before attempting to identify whether the risk is increased in particular subgroups. If this rule were adopted then the frequency of reporting false positive findings would decrease substantially.

Another potential reason for discrepant results is that genotyping laboratories may vary with respect to the sensitivity and specificity of the technique used resulting in different levels of misclassification in genotype data. The effect of even moderate levels of misclassification (e.g. a sensitivity and specificity of 90%) can sharply attenuate any true increased

Table IV. Sample size required to provide an 80% power to detect (i) departure from additive interaction and (ii) 2-fold gene–environment interaction under assumption of multiplicative model between a dichotomous environmental (E) and genetic (G) risk factor

Prevalence		OR and sample size under additive model						OR and sample size under multiplicative model		
E	G	$OR_{e'g}$	$OR_{eg'}$	OR_{eg}[a]	OR_{eg}[b]	n (P=0.05)	n (P=0.001)	OR_{eg}[c]	n (P=0.05)	n (P=0.001)
0.4	0.4	2.0	2.0	3.0	4.0	1344	2648	8.0	609	1200
0.4	0.1	2.0	2.0	3.0	4.0	2971	5849	8.0	1144	2260
0.1	0.1	2.0	2.0	3.0	4.0	4713	9195	8.0	2336	4684
0.4	0.4	1.5	1.5	2.0	2.25	10 224	20 134	4.5	566	1116
0.4	0.1	1.5	1.5	2.0	2.25	23 751	46 752	4.5	1185	2340
0.1	0.1	1.5	1.5	2.0	2.25	50 267	98 618	4.5	2555	5099
0.4	0.4	1.0	10.0	10.0	10.0	N/A	N/A	20.0	1032	2037
0.4	0.1	1.0	10.0	10.0	10.0			20.0	1027	2037
0.1	0.1	1.0	10.0	10.0	10.0			20.0	2442	4899

$OR_{eg'}$, OR for e in absence of g; $OR_{e'g}$, OR for g in absence of e.
[a]OR_{eg}, OR in presence of both e and g. Expected under additive model.
[b]OR under assumption of multiplicative model.
[c]OR expected under 2-fold interaction. Sample size formulas taken from: Garcia-Closas M and Lubin J, 1999 (38). Assumes equal number of controls.

risk resulting in a significantly lower study power, especially for detecting any gene–environment interaction (28). Controlling for genotype misclassification is, however, feasible through good study design. The extent of random misclassification may be controlled through blindly genotyping a proportion of samples twice, whereas systematic errors may require inclusion of 'gold standard' previously sequenced DNA cell lines, or at least the inclusion of a separate laboratory with a different genotyping methodology. As the use of high-throughput genotyping based on various technologies become routine, prior validation and the inclusion of a sample of doubles throughout the genotyping process should become a standard part of the conduct and reporting of genotyping studies.

How big should gene–environment interaction studies be?

The sample size estimates in Table II indicate that future studies will have to be significantly larger than the current average of 150–300 case-control pairs if they are to accurately detect moderate effects of genetic susceptibility, and that sample sizes in the range of 500–2000 case-control pairs are required as an initial step in order to detect the possible main effects of genes. A second stage of analysis may concentrate on detecting departures from both additive and multiplicative interactions for the effects of genes with particular exposures, as well as for assessment of groups of genes together and also the effect of genes in particular subgroups. As can be seen from Table IV, such studies will require significantly larger sample sizes. When relatively large independent effects are observed for both genetic and environmental exposure then a sample size in the range of 1500–5000 case-control pairs will be sufficient to detect departures from additive interaction, with an alternative hypothesis of multiplicative interaction. However, when the risk for either environmental or genetic factor is close to 1.0 then the expected OR under both additive and multiplicative model are very similar (equivalent when $OR_{eg'}$ or $OR_{e'g}$ = 1.0), and the sample size required to differentiate between additive and multiplicative models quickly becomes unrealistic. Sample sizes to detect for moderate departures from multiplicative interaction (e.g. $>$2-fold excess) do, however, remain fairly stable under a number of assumptions and indicate that studies in the range of 1000–5000 case-control pairs will be required. Studies of this magnitude for most common cancers are currently being planned, although conducting such large studies for rarer cancer sites present difficulties which may be best overcome through the coordination of multicentre studies, or through establishing consortia of independent investigators who start individual studies (D.Hunter, M.Bray, D.Burns *et al*., submitted for publication).

Multiple testing and the penalty for peeking

While recent studies have restricted analysis to one or a few genes of primary interest, the various high-throughput genotyping techniques allowing for the analysis of large numbers of SNPs will result in an increase of between 10- and 100-fold of the number of genes being tested. For example, a DNA microarray including SNPs for known relevant polymorphisms of all the genes listed in Table I may be constructed and genotyping conducted for a cost in the region of $100 per subject (assuming an average of five SNPs per gene and $0.5 per genotype). A potential concern in such studies is multiple testing and the worry that it will not be possible to distinguish the few true positive results from the abundance of false positive results. Several solutions to this have been proposed in the past including correcting for the overall significance level by incorporating the Bonferroni correction, keeping a part of the study sample to retest positive findings, and Bayesian techniques which incorporate previous results. While all three possibilities are problematic, the proposition to keep separate a retest sample is distinctly inefficient and cannot be recommended (29). Regarding the Bonferroni coefficient, often described as 'the penalty for peeking', this can have the practical advantage that if planned in advance a study sample size can be increased to allow for multiple testing. As can be seen from Table IV, the sample size required if a 0.1% significance level is adopted instead of 5% is to roughly double the sample size. Therefore, in order to test 50 hypotheses with an overall null P-value of 0.05, the sample size only has to be increased 2-fold. This ratio will depend on various parameters although it is never more than a 3-fold increase (30). While the Bonferroni coefficient is crude and not intuitively appealing, there is no substitute for a large study when testing multiple hypotheses.

A third potential methodology for analysing large numbers of hypotheses, are Bayesian techniques, which incorporate a prior belief of strength of evidence for particular genetic

associations. The rationale is that genes with a positive association but little prior support are more likely to represent false positive associations than genes with substantial prior support. The Bayesian analysis incorporates this prior knowledge resulting in stronger associations for genes with substantial prior support. Bayesian techniques are, however, little used in the context of multiple testing. One drawback is that the prior is subjective and investigators may differ strongly in the perceptions of the validity of previous results. In practice, the level of prior support is informally included in the discussion of results, and should play a major role in the interpretation of positive findings. In general, if more emphasis were placed on interpretation of results in the light of number of hypotheses tested and the strength of previous findings then a Bayesian calculation of posterior evidence would not be necessary.

An alternative pharmacokinetic model has also been proposed for the joint modelling of large numbers of metabolizing genes and exposures, allowing for the inclusion of phenotypic and functional genetic information (D.C.Thomas and V.Cortessis, submitted for publication). This approach aims to model metabolic pathways including multiple genetic and environmental factors instead of individual genes or simple two-way gene–gene/gene–environment interactions. While intuitively appealing, it does suppose a good understanding of the underlying pathway and its successful application has yet to be demonstrated.

Do we need other study designs?

While nearly all studies of genetic susceptibility and cancer are from case-control studies, several other study designs have been proposed including family-based designs and case-only designs (31,32). Family-based designs involve typing not only the affected individuals but their parents as well. The genotype which an individual does not inherit provides a virtual control that is perfectly matched for ethnic group and removes the possibility of bias from this source (31). A comparison of actual and observed genotype frequencies may be used to calculate genotype relative risks. This method may also be extended for cases with information on only one parent. A specific case of family-based studies is the transmission disequilibrium test (TDT), which identifies parents who are heterozygous for a risk allele. This may again be interpreted as a matched case-control study and the OR associated with the risk allele is simply the number of times the affected offspring inherits the risk allele from a heterozygous parent divided by the number of times the affected offspring inherits the alternative allele. Parents who are not heterozygous for the risk allele offer no information for this comparison. A variation on this design includes unaffected sibs as controls (31)

Family-based designs are usually proposed to counteract the possibility of confounding by ethnicity between cases and controls. They do however involve considerable extra effort in obtaining parental genetic information and whether this is a worthwhile use of resources is questionable. Similar information can be obtained from a well-conducted matched case-control with a minimal possibility of ethnic differences between cases and controls biasing the results. Also, recent simulation exercises have shown that even under quite extreme circumstances confounding by ethnicity is likely to have a very minor effect (<10%) (33). For most planned studies of cancer the question is academic, as both parents of most cases will have died. However, in family-based studies where the information is available the TDT test does provide an attractive study design for testing association.

The 'case-only' and 'control-only' design

A major problem when conducting studies which attempt to obtain an accurate measure of interaction is that the sample size required increases sharply (Table IV). A recently proposed study design, which aims to be more efficient for assessing gene–environment interaction is the case-only study (32). A group of cases are stratified into a 2×2 table on the basis of two exposure variables and, provided that these exposures are unrelated in the population, the 'case-only' OR (OR_{ca}) is related to the unmeasured individual OR for each exposure separately by the following formula:

$$OR_{ca} = OR_{ge} / (OR_{ge'} * OR_{g'e})$$

For example, a recent case-only study of GSTM1 null genotype and environmental tobacco smoke exposure among 106 never smoking women identified a case-only OR of 2.6, 95% CI 1.1–6.1 (34). This indicates that OR_{ge} is about two times greater than the product of the individual effects ($OR_{ge'}$ $*OR_{g'e}$) implying substantial simultaneous interaction.

The number of subjects required in a case-only study is substantially smaller than the number that would be required for a traditional case-control study for the same power to detect greater than multiplicative interaction (35). An extra strength is that genotyping may be conducted on archived tumour tissue allowing large groups of rare cases to be collected. However, the study design also has several limitations. First of all, it is not possible to assess the individual effects of exposure or genotype, and case-only studies where the main effect of the gene or exposure is not known provide only limited information. Secondly, it is based on a statistical model measuring departures from multiplicative interaction and provides no information on departures from an additive model. Thirdly, the assumption of independence between the exposure and the genotype cannot be evaluated. Modifications to this design incorporating some control information have been proposed (36), although in general the case-only design is only of limited use for identifying gene–environment interaction. An exception to this is when one is investigating risk factors for subgroups of a particular cancer, e.g. exposure to PAHs and the development of lung tumours with a p53 mutation versus lung tumours without such a mutation (37). In this situation the case-only study has reverted back to a case-control study with case-control status defined according to some tumour characteristic.

Another alternative study design, which is probably underutilized is the cross-sectional study of genetic determinants of behaviour in non-diseased population ('control-only' design). Genes that influence lifestyle factors such as smoking or alcohol consumption, or anthropometric factors such as weight control, or endogenous factors such as sex hormone levels, are also prime candidate genes for cancer development. The same approach may be used to investigate genetic determinants of markers for increased cancer risk such as markers of chromosome instability or poor DNA repair (7). Such studies have the potential to provide important functional information on candidate genes and are of much interest.

Future perspectives and recommendations for assessing gene–environment interaction

Previous gene–environment interaction studies have been based on a relatively limited sample size, have analysed only one or

several genes and have been of only limited success. It is likely that the defining feature of future studies will be much larger samples of cases and controls, and the simultaneous analysis of large numbers of candidate genes. In the absence of a strong *a priori* biological hypothesis, it may be prudent to establish a main effect for a particular polymorphism using stringent criteria before attempting to identify whether the risk is increased in particular subgroups. Also with an increasing awareness of the strong detrimental effect of misclassification it is likely that more genotyping studies will include a validation component and will publish the results of this along with the main results. Any random misclassification may be assessed by including a sample of blind doubles and systematic variation may be assessed by re-genotyping a sample in another laboratory. It is also likely that studies to identify the function of particular genotypes using non-cancer outcomes (e.g. DNA repair, hormone levels), which can aid the interpretation of subsequent cancer studies, will probably be given a higher priority.

Finally, in order to test the repeatability of findings, consortia of investigators working on particular cancer sites may need to be established. Similarly, for rarer cancer sites, the conduct of large studies may be most efficiently planned by coordination of multicentre studies, or through establishing consortia of independent investigators who start individual studies.

Acknowledgements

Many thanks to Paolo Boffetta, Pat Buffler and Nat Rothman for their helpful comments on earlier drafts of this paper.

References

1. Vogelstein,B. and Kinzler,KW. (1998) *The Genetic Basis of Human Cancer.* McGraw-Hill, New York, NY.
2. Bartsch,H., Nair,U., Risch,A., Rojas,M., Wikman,H. and Alexandrov,K. (2000) Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer Epidemiol. Biomarkers Prev.*, **1**, 3–28.
3. Strange,R.C. and Fryer,A.A. (1999) The glutathione *S*-transferases: influence of polymorphism on cancer susceptibility. *IARC Sci. Publ.*, **148**, 231–249.
4. Hirvonen,A. (1999) Polymorphic NATs and cancer predisposition. *IARC Sci. Publ.*, **148**, 251–270.
5. Berwick,M. and Vineis,P. (2000) Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. *J. Natl Cancer Inst.*, **92**, 874–897.
6. Cloos,J., Nieuwenhuis,E.J., Boomsma,D.I., Kuik,D.J., van der Sterre,M.L., Arwert,F., Snow,G.B. and Braakhuis,B.J. (1999) Inherited susceptibility to bleomycin-induced chromatid breaks in cultured peripheral blood lymphocytes. *J. Natl Cancer Inst*., **91**, 1125–1130.
7. Seker,H., Butkiewicz,D., Bowman,E.D., Rusin,M., Hedayati,M., Grossman,L. and Harris,C.C. (2001) Functional significance of xpd polymorphic variants: attenuated apoptosis in human lymphoblastoid cells with the xpd 312 asp/as genotype. *Cancer Res.*, **61**, 7430–7434.
8. Pianezza,M.L., Sellers,E.M. and Tyndale,R.F. (1998) Nicotine metabolism defect reduces smoking. *Nature*, **393**, 750.
9. Wu,X., Hudmon,K.S., Detry,M.A., Chamberlain,R.M. and Spitz,M.R. (2000) D2 dopamine receptor gene polymorphisms among African-Americans and Mexican-Americans: a lung cancer case-control study. *Cancer Epidemiol. Biomarkers Prev.*, **2000**, 1021–1026.
10. Sinha,R., Chow,W.H., Kulldorff,M., Denobile,J., Butler,J., Garcia-Closas,M., Weil,R., Hoover,R.N. and RothmanN. (1999) Well-done, grilled red meat increases the risk of colorectal adenomas. *Cancer Res*., **59**, 4320–4324.
11. Eriksson,C.J., Fukunaga,T., Sarkola,T. *et al*. (2001) Functional relevance of human ADH polymorphism. *Alcohol Clin. Exp. Res*., **5**, 157S–163S.
12. Rose,G. (2001) Sick individuals and sick populations. *Int. J. Epidemiol*., **30**, 427–434.
13. Lichtenstein,P., Holm,N.V., Verkasalo,P.K., Iliadou,A., Kaprio,J., Koskenvuo,M., Pukkala,E., Skytthe,A. and Hemminki,K. (2000) Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl J. Med*., **343**, 78–85.
14. Peto,R., Darby,S., Deo,H., Silcocks,P., Whitley,E. and Doll,R. (2000) 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.
15. Koopmans,J.R., Slutske,W.S., Heath,A.C., Neale,M.C. and Boomsma,D.I. (1999) The genetics of smoking initiation and quantity smoked in Dutch adolescent and young adult twins*. Behav. Genet.*, **29**, 383–393.
16. Lewontin,R. (2000) *The Triple Helix: Gene, Organism and Environment*. Harvard University Press, MA.
17. Kinzler,K.W. and Vogelstein,B. (1998) Colorectal tumours. In Vogelstein,B. and Kinzler,K.W. (eds) *The Genetic Basis of Human Cancer.* McGraw-Hill, New York, NY, pp. 565–590.
18. Eaden,J.A., Abrams,K.R. and Mayberry,J.F. (2001) The risk of colorectal cancer in ulcerative colitis: a meta-analysis. *Gut*, **48**, 526–35.
19. Shanahan,F. (2001) Relation between colitis and colon cancer. *Lancet*, **357**, 246–247.
20. Rothman,K.J. (1976) Causes. *Am. J. Epidemiol*., **104**, 587–592.
21. Brennan,P. (1999) Design and analysis issues in case-control studies addressing genetic susceptibility. *IARC Sci. Publ.*, **148**, 251–270.
22. Gillison,M.L., Koch,W.M., Capone,R.B. *et al.* (2000) Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J. Natl Cancer Inst*., **92**, 709–720.
23. Marcus,P.M., Vineis,P. and Rothman,N. (2000) NAT2 slow acetylation and bladder cancer risk: a meta-analysis of 22 case-control studies conducted in the general population. *Pharmacogenetics*, **10**, 115–122.
24. Green,J., Banks,E., Berrington,A., Darby,S., Deo,H. and Newton,R. (2000) *N*-Acetyltransferase 2 and bladder cancer: an overview and consideration of the evidence for gene–environment interaction. *Br. J. Cancer*, **83**, 412–417.
25. Egger,M. and Smith,G.D. (1998) Bias in location and selection of studies. *Br. Med. J.*, **316**, 61–66.
26. Macaskill,P., Walter,S.D. and Irwig,L. (2001) A comparison of methods to detect publication bias in meta-analysis. *Stat. Med*., **20**, 641–654.
27. Taioli,E. (1999) International collaborative study on genetic susceptibility to environmental carcinogens. *Cancer Epidemiol. Biomarkers Prev*., **8**, 727–728.
28. Rothman,N., Garcia-Closas,M., Stewart,W.T. and Lubin,J. (1999) The impact of misclassification in case-control studies of gene–environment interactions. *IARC Sci. Publ.*, **148**, 89–96
29. Thomas,D.C., Siemiatycki,J., Dewar,R., Robins,J., Goldberg,M. and Armstrong,B.G. (1985) The problem of multiple inference in studies designed to generate hypotheses. *Am. J. Epidemiol*., **122**, 1080–1095.
30. Sterne,J.A.C., Smith,G.D. and Cox,D.R. (2001) Sifting the evidence; what's wrong with significance tests? Another comment on the role of statistical methods. *Br. Med. J.*, **322**, 226–231.
31. Weinberg,C.R. and Umbach,D.M. (2000) Choosing a retrospective design to assess joint genetic and environmental contributions to risk. *Am. J. Epidemiol*., **152**, 197–203.
32. Khoury,M.J. and Flanders,W.D. (1996) Nontraditional epidemiologic approaches in the analysis of gene–environment interaction: case-control studies with no controls! *Am. J. Epidemiol*., **144**, 207–213.
33. Wacholder,S., Rothman,N. and Caporaso,N. (2000) Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. *J. Natl Cancer Inst*., **92**, 1151–1158.
34. Bennett,W.P., Alavanja,M.C., Blomeke,B., Vahakangas,K.H., Castren,K., Welsh,J.A., Bowman,E.D., Khan,M.A., Flieder,D.B. and Harris,C.C. (1999) Environmental tobacco smoke, genetic susceptibility, and risk of lung cancer in never-smoking women. *J. Natl Cancer Inst*., **91**, 2009–2014.
35. Yang,Q., Khoury,M.J. and Flanders,W.D. (1997) Sample size requirements in case-only designs to detect gene–environment interaction. *Am. J. Epidemiol*., **146**, 713–720.
36. Umbach,D.M. and Weinberg,C.R. (1997) Designing and analysing case-control studies to exploit independence of genotype and exposure. *Stat. Med*., **16**, 1731–1743.
37. DeMarini,D.M., Landi,S., Tian,D. *et al.* (2001) Lung tumor KRAS and TP53 mutations in nonsmokers reflect exposure to PAH-rich coal combustion emissions. *Cancer Res*., **61**, 6679–6681.
38. Garcia-Closas,M. and Lubin,J. (1999) Power and sample size calculations in cace-control studies of gene–environmental interaction: comments on different approaches. *Am. J. Epidemiol*., **149**, 689–692.

Received November 9, 2001; revised and accepted January 11, 2002