

COMMENTARY

Bias and confounding in molecular epidemiological studies: special considerations

Paolo Vineis² and Anthony J.McMichael¹

Unit of Cancer Epidemiology, Ospedale S. Giovanni Battista e della città di Torino, via Santena 7, Torino, Italy and ¹Department of Epidemiology and Population Sciences, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

²To whom correspondence should be addressed
Email: terraci@molinette.unito.it

Introduction

Studies in the field of ‘molecular epidemiology’ are becoming common in the literature. This use of molecular-biological laboratory techniques in epidemiological research has some consequences that deserve specific attention, including the potential for some particular forms of bias. Most of the bias in epidemiological research, at large, occurs because of imperfect sampling or classification procedures that cause the data-set to misrepresent the true relationship that is being studied. While there has been a long history of methodological discussions about bias in mainstream (for example, questionnaire-based) epidemiology, the corresponding issues in relation to molecular measurements in epidemiology have not been systematically addressed.

Bias, or systematic error, occurs when there is a difference between what the estimator (e.g. the relative risk) is actually estimating within the acquired data and the true effect of interest within the source population. Systematic error is attributable to methodological aspects of the study design or analysis other than sampling variation, in particular the selection of subjects, the quality of information obtained, and the superimposed effects of variables other than the disease and study factor (1).

In the following, we propose (with examples) a categorization of bias, which is certainly not exhaustive but which can be a basis for further discussion.

1. Selection bias

- related to the selection of the index (diseased) or referent population (e.g. Berkson’s bias);
- related to the selection of the exposed or unexposed population (e.g. healthy volunteer effect);
- related to selective survival;
- related to changes in laboratory assay over time (e.g. Will Rogers phenomenon).

2. Information bias:

- related to detection of the disease (detection bias);
- related to exposure or disease misclassification (dilution bias).

3. Intra-individual variation.

4. Confounding.

Abbreviations: CHD, coronary heart disease; EPIC, European Prospective Investigation into Nutrition and Cancer; GST, glutathione *S*-transferase; PAH, polycyclic aromatic hydrocarbons; RCT, randomized controlled trial.

A short history of bias in epidemiology

Important discussions on bias took place as the concept of study design in modern non-infectious epidemiology was being refined. In the 1950s, the introduction of the randomized controlled trial (RCT) and its assigned role as a ‘gold standard’ for medical research led to the anticipation that both bias and confounding could be avoided in circumstances where randomization is feasible (as in studies of putative beneficial intervention). However, Cornfield in 1954 pointed out that the RCT is only ‘one of various inferential instruments’ useful for studying aetiological relationships (2).

In the 1970s, discussion intensified on the merits and limitations of different epidemiological study designs. Central to this discussion was the relative vulnerability of each study design to bias and uncontrollable confounding (confounding being clearly distinguishable from bias, as a problem of intermixed causal effects due to the non-random distribution of risk factors within the study population). In particular, controversy arose over aspects of case–control study design: whether the case–control study could be seen as a cohort study in reverse (Feinstein’s so called ‘trohoc’ design); the choice of controls (population versus hospital-based studies); the criteria for inclusion/exclusion of cases and controls; the criteria for matching, and so on (see, for example, the special issue of the *Journal of Chronic Disease*, **32**, 1979).

In addition to general discussions on the sources of bias, ‘taxonomies’ of bias were proposed by Murphy (3) and Sackett (4), among others. Murphy suggested eight classes of bias. Some of them are hard to understand today, and others are now designated differently. For example, although ‘selection’ bias is not clearly defined as such by Murphy, he gives a prototypical example. Wood in 1950, in a series of 233 cases of congenital heart disease treated in London, found that 5% had a specific disorder: ventricular septal defect. In contrast, estimates of 35 and 37% were reported from provincial centres. The explanation given by Wood was that ventricular septal defect is easy to recognize at the local level, whereas difficult diagnoses were selectively referred to London. Therefore, the London and provincial series were not comparable.

Sackett proposed a ‘catalogue’ of 35 types of bias, which reduce to six categories that entail bias in:

1. reviewing the literature on the field;
2. specifying and selecting the study sample;
3. carrying out the experimental manoeuvre;
4. measuring exposures and outcomes;
5. analysing the data; and
6. interpreting the analysis.

Of his 35 biases, Sackett discusses nine in detail. In particular, he evaluates their capacities to distort relative risk estimates. He concludes that the most distorting biases are ‘diagnostic suspicion bias’ (a form of detection bias, see below) and the various types of information bias in the measurement of exposure and outcome. Most categories used by Sackett can be subsumed under the modern concepts of selection or information bias.

This formulation of typologies of bias during the 1970s helped to define some of the most important sources of distortion in the design, analysis and interpretation of epidemiological studies. The subsequent period has been characterized by more formal and systematic definitions.

Modern formal definitions of bias

Modern definitions of bias tend to be more formal, i.e. they tend to reduce the different empirical examples to a limited number of categories that have a logical basis. According to Miettinen, bias refers to the validity of contrasts we make within epidemiological studies: 'the key to successful design of a non-experimental study in this area, as in general, is the emulation of experimentation' (5). The validity of the randomized controlled trial, proposed here as the 'gold standard', rests on three main features:

1. the use of a placebo, i.e. comparability of index and reference interventions;
2. the use of randomization, i.e. comparability of subsets of the study population; and
3. the use of blinding, i.e. comparability of information between compared subsets.

Hence, Miettinen derives a 3-fold classification of bias into comparison, selection and information bias (according to the prevailing type of design error). Similar formal classifications occur in other contemporary texts of epidemiology, such as Rothman's (6) and Hennekens and Buring's textbooks (7). According to Rothman, selection bias is 'a distortion of the effect measured, resulting from procedures used to select subjects that lead to an effect estimate among subjects included in the study different from the estimate obtainable from the entire population theoretically targeted for the study' (6). Selection bias, in cohort studies, depends on selection of the exposed/unexposed subjects on a basis that is not independent of the outcome, or, in case-control studies, on the selection of the diseased/healthy subjects on a basis not independent of exposure status. Similarly, we have information bias when the error of classification on one axis (exposure or outcome) is not independent of the classification on the other axis.

Consider, as an example, how the measurement of molecular biomarkers within epidemiological research may result in information bias. For example, in the measurement of oxidative damage to DNA, routine phenol-based DNA purification procedures can increase 8-hydroxydeoxyguanosine levels 20-fold in samples that are exposed to air following removal of the phenol (8). Such gross contamination would seriously bias an epidemiological study if subsets (batches) coming from different subgroups in the study population (e.g. exposed versus unexposed) undergo different technical procedures that result in different levels of error.

Common to many types of bias is the fact that the main aspects of study design (selection of subjects, collection of information) are not conducted independently of the a priori hypothesis: instead of a factual 'truth' we incur a 'logically true' relationship (a truism). This would happen, for example, in a cohort study that, through faulty selection procedures, recruited into the exposed group a 'cluster' of exposed cases, or when the unexposed group is recruited from among exceptionally healthy people (for example, blood donors). In both cases an association between exposure and disease is found as a logical, not an empirical, truth.

Major categories of bias and their relevance to molecular epidemiology

Selection bias

Berkson's bias. In 1946, Joseph Berkson published a paper (9) in which he raised a particular doubt about the validity of epidemiological research within hospital settings. The underlying idea was that the relative prevalence of disease x in a group of patients who are hospitalized for disease y is inherently biased when compared with the population served by the hospital. This phenomenon, a particular type of selection bias, reflects the way in which the probabilities of hospitalization combine in patients with more than one disease (if you have two diseases, your probability of being hospitalized is greater than the probability associated with either disease separately). Berkson's argument applies in particular to hospital-based case-control studies in which one or more risk factors (especially medications) are studied in relation to the risk of a specific disease. If, for example, obese people who have hypertension have a higher probability of being hospitalized than obese people without hypertension, a spurious association between obesity and anti-hypertensive drugs is observed. People with multiple diseases or conditions are over-represented in the hospital population, and this over-representation affects the distribution of risk factors as well.

Berkson's bias was long considered an epidemiological curiosity, until its reality was empirically demonstrated by Roberts *et al.* (10). They re-analysed household surveys of health utilization information. Information was gathered for eight clinical conditions and six medications from both hospitalized and non-hospitalized members of households. There were statistically significant differences in associations between drug use and specific diseases in community-based and hospital-based settings.

Berkson's bias is relevant to molecular epidemiology. Consider hospital-based case-control studies of disease risk in relation to metabolic polymorphisms: we can imagine at least three mechanisms by which Berkson's bias can occur. First, if a person is hospitalized for a specific reason, but has more than one pathological condition, it is possible that the concurrent disease is also associated with the genetic polymorphisms under investigation. Second, patients with a certain allele at the polymorphic locus under investigation can have adverse reactions to drugs and be hospitalized for this reason. Third, induction of an enzyme by treatment can influence the phenotypic indicator of genotype. For example, the administration of methotrexate can induce hydroxyfolate-reductase by gene amplification: therefore, if in a case-control study we include cancer patients among the controls we may have a distorted association between the disease under study and hydroxyfolate-reductase activity.

Selection of disease cases and survival effect. Other types of distorted estimates can occur as a consequence of case selection. For example, in investigations based on the collection of bladder biopsies and the subsequent analysis of $p53$ mutations, small biopsies (corresponding to early stages of disease) are usually omitted because the material is insufficient for DNA analysis. However, $p53$ mutations are more frequent in late stages; therefore, we are liable to overestimate the prevalence of mutations if we do not take care with case selection.

Selection bias can also result from selective survival. An example relates to the study of Alzheimer's disease. The apolipoprotein E e4 allele, a risk factor for Alzheimer's disease,

also increases the risk of atherosclerosis and coronary heart disease (CHD), especially in combination with smoking. Among people with the apo-E e4 allele, non-smokers are therefore more likely to survive to old age than are smokers. Thus, we are likely to find a deficit of smokers among Alzheimer's cases (typically aged over 75 years). While this observation would seem to suggest a protective effect of smoking, in fact it may be due to selective survival of non-smokers among persons at high risk of Alzheimer's disease, because of the intervention of CHD mortality (11,12).

Kelsey *et al.* have reported an example of selective survival recently: the frequency of the glutathione *S*-transferase (GST) class mu deletion was different among incident and prevalent (i.e. non-newly diagnosed) cases of breast cancer. Possibly, this distribution was related to a survival advantage associated with the polymorphism of GST mu (13).

Healthy worker effect. William Ogle, when studying death rates in different industries (1885), described two difficulties he encountered. The first was 'the considerable standard of muscular strength and vigour to be maintained' in order to keep on performing many tasks in the industry. If the individual's health or strength fell below this standard, he was compelled to move to a more suitable activity, or even retire. The second difficulty was that 'some occupations may repel, while others attract, the unfit at the age of starting work and, conversely, some occupations may be of necessity recruited from men of supernormal physical condition' (14).

Nearly 100 years after Ogle's observation, Fox and Collier in the UK (15) and McMichael in the USA (16) examined the same phenomenon in industrial cohorts in terms of standardized mortality ratios based on a general population referent. The term 'healthy worker effect' was coined in relation to the favourable mortality experience of rubber workers relative to that of the general population (16). More generally, the overall mortality experience of an employed population is typically better than that of the general population, at least in Western countries. For example, in a study described by Monson, the mortality rate per 1000 per year was 9.1 among White steelworkers, and 15.8 in the general population (non-White: 9.9 and 18.8, respectively) (17).

The most widely accepted explanation for the so-called healthy worker effect is selection of the workforce, either as a result of self-selection by the employee or selection by the employer (for more details see ref. 18). The unemployed section of the general population includes people with serious health conditions that hamper their ability to work. This is not necessarily true in the Third World, where manual workers may undergo more superficial pre-hiring assessment and may suffer from more severe consequences of workplace exposures.

A 'healthy volunteer effect' may occur in molecular epidemiological studies based on populations of volunteers, such as blood donors. Such populations tend to be self-selected on the basis of better lifestyles, so that the expected burden of cancer-related adducts, for example, may be lower than in the general population. This phenomenon will not create a problem of validity, if we only make internal comparisons within the volunteer population, although this would decrease the statistical power of the study.

The Will Rogers phenomenon. Will Rogers was a humorist-philosopher who described a geographic migration during the American economic depression of the 1930s. He said, 'when the Okies [the inhabitants of Oklahoma] left Oklahoma and

moved to California, they raised the average intellectual level in both states' (citation from ref. 19). Taste aside, the comment well describes a general phenomenon: for example, migration of an average soccer player from a very good to a poorly performing team will improve the performances of both teams. In medicine, the Will Rogers phenomenon refers to improvement over time in the classification of disease stages: if diagnostic sensitivity increases, metastases are recognized earlier, so that the distinction between early and late stages of cancer will improve. Survival rates rise in each group, without any change in individual outcomes, because the prognosis of those additional persons now correctly classified as late stage is worse than that for the remaining members of the early-stage group, but is better than that for the original members of the late-stage group (19).

The Will Rogers phenomenon can occur in studies of molecular epidemiology in several ways, not necessarily related to survival. For example, if diagnostic techniques improve (i.e. sensitivity and specificity increase in the course of time) this may introduce a 'drift' in longitudinal studies. Suppose we collect and freeze blood samples in a cohort of volunteers [such as the EPIC cohort on diet and cancer (20)], and then we conduct a nested case-control study on prospective cervical cancer cases and a sample of non-cases. We measure, for example, vitamin C in serum as a measure of exposure at the time of blood collection. Cancer cases diagnosed more recently will, because of diagnostic improvement, tend to belong to earlier stages than cases diagnosed years before (when blood was collected). Suppose vitamin C protects against cervical cancer only in the late events of its natural history, i.e. a short time before cancer onset. Then we may find erroneously that the association with vitamin C holds only for 'advanced' (invasive) stages of disease, simply because cancers diagnosed immediately after blood collection belonged more easily to 'late' clinical stages.

Information bias

Detection bias. 'Detection bias' is likely to be a common problem in case-control studies in which the risk factor investigated itself leads to increased diagnostic investigations and thus increases the probability that the disease is identified in that subset of persons. When the first studies on the relationship between mammographic patterns and the risk of breast cancer appeared, suggesting that benign breast disease could predispose to cancer, the objection was raised that the observed association could be attributed to 'detection bias', i.e. the greater probability that women with benign breast cancer had to undergo detailed examinations, including repeated mammograms, which would lead to an earlier diagnosis of cancer. This bias was empirically demonstrated by Silber and Horwitz (21) in a case-control study. They showed that the crude odds ratio for the association between benign breast disease and breast cancer was 2.6 (statistically significant). However, when the analysis was repeated within strata of diagnostic procedures, the association disappeared (OR = 0.9 for mammography patients, 0.8 for biopsy patients).

Detection bias can be considered as a form of information bias in that the probability of identifying the diseased people is conditional on the clinical information collected, which differs between categories of the risk factor (22).

In molecular epidemiological studies, this may happen, for example if molecular markers of early disease are prospectively analysed in a cohort. This will lead to easier detection of the

Table I. Association between birth weight or birth length and smoking exposure comparing results of questionnaire and biochemical or molecular methods (³²P-post-labelling DNA adducts) for exposure assessment (from ref. 25)

Parameter for smoking exposure	Birth weight		Birth length	
	R ²	P-value ^a	R ²	P-value ^a
Questionnaire	0.41	0.85	0.34	0.28
Biochemical	0.48	0.09	0.43	0.03
DNA adducts	0.52	0.025	0.47	0.01

^aSignificance of adding smoking parameter to basic model.

eventual clinical disease in those who test positively, even if the marker is not necessarily intermediate in the causal chain leading from exposure to disease. For example, some exposures, such as formaldehyde, can induce micronuclei in mucosal cells of the oral cavity; these, in turn, may lead to earlier detection of oral cancer through subsequent periodic examination of the workers who test positive. A similar phenomenon can happen with the identification of mutated oncogenes or tumour suppressor genes in exposed workers, well before the onset of clinical disease.

Dilution bias or random misclassification. It is usually claimed that random errors in the classification of the status of subjects (exposure or disease) lead to dilution of the effect of exposure, or bias 'toward the null' (23). This is not necessarily true (24), but is true in most cases, and it is a motivation for introducing better measurements in epidemiology (including molecular epidemiology).

An example of how the use of internal dose markers (in the field of reproduction) can reduce dilution bias is the study by Everson *et al.* (25) shown in Table I. Accurate dosimetry of fetal exposure is very important if we are interested in teratogenesis or transplacental carcinogenesis. Most teratogens act in a restricted time window (within the first 55–60 days of gestation) and traditional measures of exposure (questionnaires) may be too inaccurate to estimate exposure in such a restricted window. Everson *et al.* showed that while there was no statistically significant association between reproductive effects and smoking as assessed by questionnaire data, DNA adduct levels were clearly correlated with birth weight and length (Table I).

In general, more precise identification of risk factors and pathogenetic pathways via the use of molecular-biological measures should reduce misclassification: a well-known example is HPV detection in relation to the number of sexual partners in studies of the aetiology of cervical cancer. However, this is not necessarily true: for example, questionnaire measures of smoking are more accurate measures of cumulative exposure than various biomarkers (26).

Intra-individual variation

Biological markers are characterized by intra- and inter-individual variation, the magnitude of which depends on the type and significance of the marker. Inter-individual variation is a source of imprecision of the estimates (e.g. as measured by the confidence interval). Intra-individual variation poses different problems.

There are at least two types of intra-individual variation that are relevant to the measurement of biomarkers. Variation

Table II. Two types of intra-individual variation and their relevance to different categories of biomarkers

Biomarker category	Intra-individual variation over time	Biological sampling variation
Internal dose (blood)		
Hormones	Yes (diurnal variation)	No
Water soluble nutrients	Yes (short half-life)	No
Organochlorine	No (long half-life)	No
Biologically effective dose		
Peripheral white blood cells	Yes (half-life: weeks to months)	No
Exfoliated urothelial cells		
DNA adducts	Yes (half-life: months)	Yes
Early biological effects		
Lymphocyte metaphase		
Chromosome aberrations	More or less stable	?
Somatic cell mutations		
Glycophorin A	Probably low	No (?)
Intermediate markers		
Cervical dysplasia	Yes	Yes
Colonic hyperproliferation	Yes	Yes
Genetic susceptibility		
Genotype assay	No	No
Non-inducible phenotype	No	No
Inducible phenotype	Yes	No
Tumour markers	Yes	Yes

over time is characteristic, for example, of age-related or seasonal variables, such as the intake of fruit and vegetables. Variation over time may be a source of confounding if a determinant of the disease under investigation displays the same periodicity over time as does the study variable. For example, both fruit intake (i.e. the intake of vitamin C) and the diffusion of the influenza virus have a strong seasonality: when the virus is widespread, vitamin C intake is generally low. This particular relationship may be confounding and erroneously emphasize the protective role of the vitamin.

Another type of intra-individual variation is not related to time but to variation in biological sampling procedures within the individual. For example, hyperproliferation of colonic cells is extremely variable in different points of the colon mucosa. Therefore, measurements are strongly influenced by how and where the mucosa is sampled. It has been estimated (27) that 20% of the variability of the rectal mucosa proliferation index (measured by nuclear antigen immunohistochemistry) is due to inter-individual variation, 30% to the site of biopsy within the subject and 50% is due to inclusion of crypts (i.e. micro-anatomic location) within a biopsy. In other words, as much as 80% of variation is related to sampling. This implies not only a widening of confidence intervals of the estimates, but substantial misclassification of individuals.

Table II shows how the two types of intra-individual variation, over time and biological sampling procedure, can affect different types of biomarkers.

Confounding

Confounding, strictly, must be differentiated from bias (although some textbooks do not). Bias is a misrepresentation of the aetiological relationship because of defective study design and methods, whereas confounding is related to the distribution of relevant variables within the sampled population. A confounder is usually defined as a variable that is a risk factor for the disease and is associated with the exposure of interest in the study population (7). We can imagine several different types of confounders in molecular epidemiology, and

they may not be easily identified and adjusted for. For example, polyaromatic hydrocarbon (PAH)-DNA adducts integrate across several disparate environmental exposures: it is not possible to adjust out the components of adduct load that are not due to the index exposure (26).

It may also happen that the definition of confounder in molecular epidemiology is more subtle than in traditional epidemiology. Taioli and Garte have stressed that in the study of metabolic polymorphisms in relation to cancer risk a confounder can be an exogenous exposure that is associated not with another exogenous exposure but, rather, with gene expression or enzyme induction (E.Taioli and S.Garte, manuscript in preparation). For example, it has been noted that the CYP1A2 polymorphism is associated with variation in the risk of colon cancer; however, it has also been shown that the consumption of cruciferous and other vegetables induces the activity of the CYP1A2 enzyme (28). We know that vegetables reduce the risk of colon cancer (29) (perhaps for their content in anti-oxidants, acting via pathways unrelated to the CYP1A2 enzyme). Therefore, the association between CYP1A2 and colon cancer may be no more than a statistical association, confounded by dietary habits. In general, the use of inducible enzymes is problematic in case-control studies.

Conclusions

Epidemiology is largely a non-experimental discipline. Although this has often been perceived as a weakness, it is also an important strength in that observations are made on samples of free-living, real-world, populations. In fact, the limited opportunity for experiments in epidemiological research has led to a very critical theoretical understanding of the different types and sources of error.

Molecular epidemiology can be particularly prone to error because of the unknown and often unpredictable ways in which biochemical and molecular markers are associated with exposure, on the one side, and disease, on the other. The collection of examples of bias from field studies, and their organization into a theoretical framework, will help the evaluation of using molecular-biological markers in epidemiologic studies.

Acknowledgement

This work was partially supported by Associazione Italiana per le Ricerche sul Cancro.

References

- Kleinbaum,D., Kupper,L.L. and Morgenstern,H. (1982) *Epidemiologic Research*. Lifetime Learning, Belmont, CA, p. 186.
- Cornfield,J. (1954) Statistical relationships and proof in medicine. *Am. Statistician*, **8**, 19-21.
- Murphy,E.A. (1976) *The Logic of Medicine*. Johns Hopkins University Press, Baltimore, MA.
- Sackett,D.L. (1979) Bias in analytic research. *J. Chron. Dis.*, **32**, 51-63.
- Miettinen,O.S. (1985) *Theoretical Epidemiology*. John Wiley, New York, NY.
- Rothman,K.J. (1986) *Modern Epidemiology*. Little, Brown and Co., Boston, MA.
- Hennekens,C.H. and Buring,J.E. (1987) *Epidemiology in Medicine*. Little, Brown and Co., Boston, MA.
- Wiseman,H., Kaur,H. and Halliwell,B. (1995) DNA damage and cancer: measurement and mechanism. *Cancer Lett.*, **93**, 113-120.
- Berkson,J. (1946) Limitations of the application of 4-fold tables to hospital data. *Biomet. Bull.*, **2**, 47-53.
- Roberts,R.S., Spitzer,W.O., Delmore,T. and Sackett,D.L. (1978) An empirical demonstration of Berkson's bias. *J. Chron. Dis.*, **31**, 119-128.
- Riggs,J.E. (1993) Smoking and Alzheimer's disease: protective effect or differential bias? *Lancet*, **342**, 793-794.
- Hardy,J. and Roberts,G.W. (1993) Smoking and neurodegenerative disease [letter]. *Lancet*, **342**, 1238.
- Kelsey,T.K., Hankinson,S.E., Colditz,G.A. et al. (1997) Glutathione S-transferase class mu deletion polymorphism and breast cancer: results from prevalent versus incident cases. *Cancer Epidemiol. Biomarkers Prev.*, **6**, 511-516.
- Ogle,W. (1885) Letter to the Registrar-General on the mortality in the registration districts of England and Wales during the ten years 1871-80. Supplement to the 45th Annual Report of the Registrar General of Births, Deaths, and Marriages, in England, p. xxiii.
- Fox,A.J. and Collier,P.F. (1976) Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. *Br. J. Prev. Soc. Med.*, **30**, 225-230.
- McMichael,A.J. (1976) Standardized mortality ratios and the 'healthy worker effect': scratching beneath the surface. *J. Occup. Med.*, **18**, 165-168.
- Monson,R.R. (1980) *Occupational Epidemiology*. CRC Press, Boca Raton, FL.
- Arrighi,M. and Hertz-Picciotto,I. (1993) Definitions, sources, magnitude, effect modifiers and strategies of reduction of the Healthy Worker Effect. *J. Occup. Med.*, **5**, 890-891.
- Feinstein,A.R., Sosin,D.M. and Wells,C.K. (1985) Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N. Engl. J. Med.*, **312**, 1604-1608.
- Riboli,E. (1992) Nutrition and cancer. Background and rationale of the European Prospective Investigation into Cancer and Nutrition. *Ann. Oncol.*, **3**, 783-791.
- Silber,A.L.M. and Horwitz,R.L. (1986) Detection bias and relation of benign breast disease to breast cancer. *Lancet*, **i**, 638-640.
- Feinstein,A.R. (1979) Methodologic problems and standards in case-control research. *J. Chron. Dis.*, **32**, 35-41.
- Copeland,K.T., Checkoway,H., Holbrook,R.H. and McMichael,A.J. (1977) Bias due to misclassification in the estimate of relative risk. *Am. J. Epidemiol.*, **105**, 488-495.
- Dosemeci,M., Wacholder,S. and Lubin,J. (1990) Does non-differential misclassification of exposure always bias a true effect towards the null value? *Am. J. Epidemiol.*, **127**, 746-748.
- Everson,R.B., Randerath,E., Santella,R., Avitts,T.A., Weinstein,I.B. and Randerath,K. (1988) Quantitative association between DNA damage in human placenta and maternal smoking and birth weight. *J. Natl Cancer Inst.*, **80**, 567-576.
- Pearce,N., De Sanjose,S., Boffetta,P., Kogevinas,M., Saracci,R. and Savitz,D. (1995) Limitations of biomarkers of exposure in cancer epidemiology. *Epidemiology*, **6**, 190-194.
- Lyles,C.M., Sandler,R.S., Keku,T.O., Kupper,L.L., Millikan,R.C., Murray,S.C., Bangdiwala,S.I. and Ulshen,M.H. (1994) Reproducibility and variability of the rectal mucosal proliferation index using proliferating cell nuclear antigen immunohistochemistry. *Cancer Epidemiol. Biomarkers Prev.*, **3**, 597-605.
- Kall,M.A., Vang,O. and Clausen,J. (1996) Effects of dietary broccoli on human *in vivo* drug metabolizing enzymes: evaluation of caffeine, oestrone and chlorzoxazone metabolism. *Carcinogenesis*, **17**, 791-799.
- Potter,J. (1996) Nutrition and colorectal cancer. *Cancer Causes Control*, **7**, 127-146.

Received on March 2, 1998; revised on July 10, 1998; accepted on August 10, 1998