Epidemiology: Problems in the Study of Cancers of Low Incidence and the Need for Collaboration

Many cancers exist for which there are currently no clear hypotheses that could be pursued in epidemiologic investigations. These include carcinomas of the kidney, some parts of the digestive system, non-Hodgkin’s lymphomas, tumors of the nervous system, and many sarcomas, whereas the same lack of etiologic factors is also evident in many common cancers other than those clearly related to smoking or occupational carcinogens. Even so, the incidence of most of them shows substantial international variation. Although part of this variation may be due to international differences in diagnostic or reporting practices, it is nevertheless important, if we adhere to the environmental hypothesis as an explanation for most of the discrepancies, that we attempt to identify the relevant environmental factors.

The question arises as to how an epidemiologist obtains appropriate hypotheses and how he secures support for an exploratory study which may be criticized because he does not have a clear hypothesis to test.

Classically, the epidemiologist will proceed by literature review, consultation with clinical and basic science colleagues, and by a search for leads from clinical practice or laboratory investigations, especially if there are appropriate animal models. Descriptive national and international epidemiologic data will be explored over recent and longer time periods. To seek other clues, these data may be analyzed further for cohort differences. Thus the importance of good reliable data is emphasized. Thinking and posing the right question are essential. But in the absence of an appropriate hypothesis, epidemiologists are driven to perform exploratory studies in search of information on any factor that may be applicable to the condition under investigation. Usually this will involve the case-control or retrospective approach, but occasionally access to existing cohorts with identified information collected before the onset of the disease might justify either a case-control approach within these data or a follow-up cohort study.

Unfortunately, in a case-control study a long time will be required for a study of a rare condition. Furthermore, the nature of the data sought and ideas may change with time so that the investigator may be tempted to include additional requirements for information within the instrument used. If this process continues extensively, the end result may be a data set consisting of information obtained on varying numbers of individuals that may be almost impossible to interpret.

Further, problems may arise when one considers sources of cases and controls. How important is it that they should be representative of the population from which they are drawn? In the past many epidemiologists apparently regarded exploratory studies as justified, albeit they were based on highly selected populations sometimes picked up through a referral center. It may be questioned whether one can justify a study of prevalent cases or new ones coming to a referral center, since the results apply only to patients surviving long enough to be available for study or attending the center for therapy. Thus the extent to which they are representative of all cases cannot be determined.

It is important, even in a rare condition, that unselective series are studied, preferably all those that occur in a defined area and time period. Otherwise, whatever the leads, further studies will be needed to confirm the general applicability of the results. An additional advantage of such an approach is that attributable risk estimates can be calculated. That is the extent to which a factor, identified as contributing significantly to the occurrence of the disease, is sufficiently important in terms of the numbers of cases induced by it to justify further investigation and preventive measures.

Editor’s note: Periodically the Journal publishes solicited guest editorials as a means of transmitting to investigators in cancer research the essence of current work in a special field of study. The Board of Editors welcomes suggestions for future editorials that succinctly summarize current work toward a clearly defined hypothesis regarding the causes or cure of cancer.
What about the controls? Is it ever justifiable to use patients with malignant disease as controls for patients with other malignant conditions? Presumably not, if most cancers are environmentally induced, because these individuals cannot represent the population from which cases are drawn. What about available numbers? Although a number of epidemiologists have increased the control-to-case ratio to compensate in rare conditions for deficiencies of cases, it is not possible if invalid controls are used to compensate for deficiencies in design by this process. Surely, in an investigation of a rare condition, appropriate care must be taken to select controls representative of the population from which the cases are drawn. This is not to imply that controls always have to represent the normal population, even though this should usually be the rule in exploratory studies. Close matching may be required if one is seeking to refine the conclusions from previous studies to avoid confounding (6), but unfortunately the opportunities for such precision are rare.

Thus we must make certain we do not overlook elementary precautions of design in the need to ensure economies in exploratory studies which in the past resulted in reports raising more questions than they answer.

What about the analysis? Why do authors so often ignore the matching they have painstakingly done initially, even though the precision of analysis is increased by the retention of the original matching (7), whereas to ignore it could be incorrect (8)? Surely we should not refrain from attempting to unravel multifactorial causes, particularly since models for multivariate analysis are applicable to case-control studies (9).

What about the report? Is it justifiable at this time of information overload for editors to devote space to reports of studies criticized by reviewers for what now seem to be elementary errors of design but which may be based on plans finalized many years before? Editors should neither condone elementary defects in analysis nor permit authors to fail to consider the implications of admitted deficiencies of design. However, albeit some conclusions of authors may be questioned because deficiencies may give rise to many difficulties in interpretation, they still may give other workers the clue to the appropriate investigation. Thus publication is justified at least in a brief form. After all, a study does not live in science until it has been published and can be criticized; editors and reviewers should be careful that their own biases and dogma do not prevent from seeing the light of day what, to their surprise, could be important leads. But authors should be careful not to submit for publication studies, however inconclusive, without at least pointing out a tentative working hypothesis that can be evaluated by themselves or others. A recent example is a study of prostate carcinoma (10). Although the work was exploratory and largely negative, it did suggest to the authors a more appropriate age (late middle age) for investigation of possible etiologic factors than they had themselves pursued. As Lowe (3) reminded us, a good scientific hypothesis is one that can be disproved, and thus can lead to the substitution of a more satisfactory hypothesis, and so on, until "truth" is finally obtained.

If so many questions can be directed to studies that have been or may be published, does the epidemiologist have a place in attempting to develop hypotheses that he may evaluate in man, or his laboratory colleagues may evaluate in animals before further investigations are contemplated in man? Surely, the answer must be yes. But the epidemiologist in designing studies and the granting agency in funding him do have an obligation to the scientific community to ensure that all available knowledge is taken into account. The International Agency for Research on Cancer is attempting to make certain that appropriate consideration is given to the multidisciplinary approach in investigations, as shown by a conference report which considered environmental causes of cancer (11). There is also an increasing interest in multidisciplinary workshops to consider more focused attacks on problems.

Other solutions involve a degree of cooperation and collaboration, possibly foreign to many investigators, and a subjugation of personal aims that scientists, who think of themselves as independent investigators, may find unacceptable. However, lessons have been learned by clinicians in relation to the need to conduct multicenter-controlled collaborative clinical trials; it is time that epidemiologists realized a similar need. With the use of the resources of several centers and with a carefully thought-out common design and instrument, studies of a rare condition will result in more rapid answers than are possible from one institution. This will avoid the tendency to change design or the nature of the data sources as time passes. Such studies need not be onerous for the investigator, because his contribution will be relatively slight, but he will have the intellectual satisfaction of participating in a joint project and achieving answers that he could not get on his own. This approach will be more expensive for the granting agency, but will be worthwhile because it will be more "cost-effective" than alternative approaches. A recent example of such collaboration has been the investigation of the epidemiology of Hodgkin's disease by workers in New Orleans and Los Angeles (12) and is also exemplified by current collaborative studies of breast and bladder cancer in Canada.

A further approach is to make sure that unselective cases are studied with the use of data now being accumulated in cancer registries in many parts of the world. The opportunities, now arising in the Program of Cancer Surveillance, Epidemiology, and End Results Reporting of the U.S. National Cancer Institute, should be used for studies within individual component parts of this program and also by collaboration between centers to guarantee that rare conditions are appropriately studied.

There is nothing wrong with our methodology, either case-control or cohort. It is our inability to
use the data, ideas, and leads already available to us to investigate the etiology of the conditions we are interested in, using modern methods of analysis to dissociate the contribution of the various factors which together cause the cancer. Indeed, it is possible for many conditions that the clues are already available. To enable us to solve the problems, we need appropriate detective work and alert minds for the unexpected. An example of an unexpected outcome from an exploratory study has just come from the Boston Collaborative Drug Surveillance Program (13). Unfortunately, two studies designed to confirm a possible association of reserpine and breast cancer (14, 15) suffered from some of the deficiencies I have discussed so that the reality of the association and possibly more important its significance remain in doubt.

An "ex cathedra" statement sometimes made is that the differences between those that have cancer and those that do not may not be sufficiently great in the Western World for detection of causative factors with the use of standard epidemiologic techniques. This is a belief to which I do not subscribe, though I accept that perhaps in many circumstances our methodology has not been good enough, until now, to enable us to detect the differences that must exist. We have to remember it may be as important to find out why 9 of 10 heavy smokers do not get lung cancer as to identify the 1 of 10 who does. These are the sorts of clues we should follow with as much energy as our searches for factors in individuals who actually develop the disease.

REFERENCES

(5) MIETTINEN OS: Proportion of disease caused or prevented by a given exposure, trait or intervention. Am J Epidemiol 99: 325-332, 1974
(14) ARMSTRONG B, STEVENS N, DOLL R: Retrospective study of the association between use of Rauwolfia derivatives and breast cancer in English women. Lancet 2: 672-675, 1974