Controlling the Avoidable Causes of Cancer: Needs and Opportunities for Etiologic Research
Jonathan M. Samet
School of Hygiene and Public Health, The Johns Hopkins University, Baltimore, Maryland

This meeting of the President’s Cancer Panel was designed to provide an overview of known and suspect causes of cancer and to indicate those that might be considered avoidable. Two complex concepts are inherent in this charge: cause and avoidability. Risk factors for cancer are designated as causal when the evidence from observational and laboratory research is judged sufficient in relation to criteria for causality; the extent to which cancers of specific sites can be avoided is best estimated by the attributable risk statistic, which incorporates both the exposure pattern and the relative risk for the cancer-causing agent. A research agenda on avoidable causes of cancer should then address both the risks associated with the agents that cause cancer and the pattern of exposure to the agents. Presentations at the meeting highlighted gaps in the evidence on the risks associated with various known and potential causes of cancer and on the patterns of exposure across the diverse groups within the population. In spite of these gaps, presenters emphasized that the evidence is already sufficient to justify intervention for many agents and that action need not be delayed for the well-characterized causes of cancer. In addition to research recommendations offered by presenters for specific causal agents, the scientific basis for cancer prevention might be generally strengthened by new research strategies directed at developing new tools for exposure assessment, for investigating the risks of mixtures, and for population surveillance. — Environ Health Perspect 103(Suppl 8):307–311 (1995)

Key words: cancer etiology, cancer control, risk assessment, epidemiology

Introduction
This conference sought to provide an overview of known and suspect causes of cancer and to indicate those that might be considered avoidable. The presentations covered the full array of environmental agents, including such widely present exposures as air pollution, water pollution, infectious agents, and ionizing and nonionizing radiation. They focused more specifically on hormones and medications, occupational exposures, and lifestyle-associated risk factors such as diet, tobacco, and alcohol. Gender and ethnicity were considered as factors that may influence exposure patterns and determine responses to environmental agents. The emerging field of cancer genetics was also covered; this field brings the promise of a heightened understanding of mechanisms of cancer causation and even the possibility of identifying persons at risk from environmental exposures on the basis of genetic susceptibility.

These wide-ranging presentations point to myriad opportunities for future research and to ways that the new evidence can be used for cancer prevention and control. This paper considers the needs for new etiologic information on avoidable causes of cancer and the most immediate opportunities for research, highlighting key issues raised during the meeting. The coverage is selective, given the broad scope of the presentations.

A number of general themes are apparent. First, there have been numerous successes using both epidemiologic and laboratory-based approaches to identify avoidable causes of cancer—such historically significant examples as Percivall Potts’ finding of scrotal cancer in chimney sweeps based on astute observation, and identification of smoking as a cause of lung cancer based on straightforward application of the case–control design. These and other examples of successes involved exposures to single agents or environments.

A second theme is that we have had less success in investigating more complex exposures, particularly those like diet that are correlates of lifestyle choices. For such exposures, the signal-to-noise ratio may be low. Biomarkers of exposure, dose, susceptibility, and response may enhance the sensitivity of epidemiologic studies. Large study populations represent another solution to the signal-to-noise problem.

Several papers reflected on a third theme—that population is not homogeneous by race, ethnicity, and gender, by socioeconomic status, or by susceptibility to cancer-causing agents. In a broadly conceived scheme of the causal pathway that leads to cancer, primary determinants of patterns of environmental exposure include race, ethnicity, gender, income, and education (Figure 1). These sociodemographic factors may increase both the likelihood of exposure to cancer-causing agents and the doses of these agents. To strengthen the scientific basis for prevention, papers were unanimous in indicating that research must reflect the diversity of the population.

Finally, some of our successes in identifying avoidable causes of cancer have shown that this identification is only a first step in cancer control. Cigarette smoking remains an exemplary challenge. Over 40 years after the publication of convincing

Figure 1. Cancer risk reflects environmental exposures and genetic susceptibility.
evidence that cigarette smoking causes lung cancer, approximately 25% of adults in the United States are smokers, and adolescents continue to start smoking, even with knowledge of the future risks of smoking-caused disease. We have learned that tobacco control requires a broad-based approach that incorporates strategies for education, strategies for assisting the individual smoker to quit, and strategies to change the social milieu in ways that foster smoking prevention and cessation (1).

**Causes and Attributable and Preventable Risks**

The phrase "Avoidable Causes of Cancer" incorporates two complex concepts: that of causality and that of avoidability. The concept of causality and the identification of causal relationships have long engaged not only philosophers but scientists involved in both experimental and observational research (2). More modern systems for gauging evidence and determining causality date to the Koch-Henle postulates. The currently applied criteria for determining the causality of exposure-disease associations include the strength of the association and its consistency among studies, the presence of a dose-response relationship of agent with disease, proper temporality of the association, and coherence and plausibility of the association with other biomedical evidence. These criteria have proven effective for agents like cigarette smoking that are powerful risk factors for disease (3).

They have not been so readily applied to agents with weaker effects and, indeed, they may be inappropriate for this purpose (4,5). Even the application of these criteria requires a subjective determination of whether they have been met. Consequently, some suspect causes of cancer quickly have become cloaked in controversy as ambiguous evidence is interpreted as showing or not showing a causal relationship.

In describing the burden of disease that might be avoided, epidemiologists use a quantity referred to as attributable risk (4). Attributable risk indicates the burden of disease that could be avoided if exposure to the agent of concern were fully prevented. One form of attributable risk, population attributable risk (PAR), describes the proportion of disease in a population associated with exposure to an agent. The PAR can be calculated as:

\[
\text{PAR} = \frac{(RR - 1)(P)}{1 + (RR - 1)(P)},
\]

where RR is the relative risk of the disease, comparing persons exposed to the agent with the nonexposed, and P is the proportion of the population exposed to the risk factor. The PAR increases with both the value of RR and of P. The equation shows that a rare exposure associated with high RR for those exposed may have little impact on PAR for the population in general. For example, an agent that increases cancer risk 10-fold for a small segment of the population—for example, 0.001%—has a PAR of only 0.009%. This example might represent a strong occupational cause of disease affecting a group of workers. By contrast, the PAR for an agent with a RR of 1.10 and an exposure prevalence of 10% is 1%.

Attributable risk estimates should be interpreted as representing the theoretical maximum number of cases avoidable if the exposure of concern could be fully removed, i.e., P reduced to zero. In practice, however, it may not be feasible to prevent exposure completely, and the number of cases preventable by intervention may be only some fraction of the theoretical estimate of PAR. Moreover, interventions may not be intended to eliminate exposures but to reduce those exceeding a threshold level of acceptability. Radon, for example, has been estimated to cause about 14,000 lung cancer deaths per year in the United States (6,7). These cases could be avoided only if levels of radon in residences were reduced to the background level in outdoor air, an unachievable goal. Lubin and Boice (8) show that reducing the levels in homes having concentrations above the guideline value of the U.S. Environmental Protection Agency (U.S. EPA) removes only one-third of the total burden of lung cancer attributable to radon. Larger gains can be made only by reducing concentrations in homes below the guideline value as well. By contrast, the entire burden of lung cancer caused by cigarettes could be eliminated with complete control of smoking.

Many cancers of current public health concern are likely to have multiple causes. For diseases caused by multiple agents, the total burden of disease theoretically preventable may exceed the observed number of cases or 100% if there are synergistic patterns of joint action as assessed on an additive scale. For example, radon and cigarette smoking are synergistic in causing lung cancer (9). Estimates of radon-attributable lung cancer cases can be conceptualized as to include those caused by radon in never smokers, those caused by radon in smokers, and those caused jointly by radon and smoking in smokers. This subtlety of the attributable risk statistic is not widely appreciated and, for example, there is a widespread misconception that the lung cancer cases that might be caused by radon include only the 10 to 15% not directly attributed to smoking.

Although simplistic in its mathematical formulation of complex biologic phenomenon, the PAR makes clear that cancer can be avoided by either reducing RR or P. As a basis for avoiding cancer, we must identify causal agents, those factors with RR above unity for which the evidence warrants designation of the association as causal. Cases of cancer can be avoided by reducing the value of RR (e.g., through reduction of the level of exposure or by use of a chemopreventive agent) and also by reducing the value of P (e.g., by eliminating the exposure). The development of markers of susceptibility may allow us to partition the population into two groups, those with increased or higher RR because of susceptibility and those with lower RR or no increased risk who are not susceptible (Figure 2). Preventive strategies would then target those with increased risk on the basis of susceptibility.

To date, epidemiologic and laboratory-based research have successfully identified many risk factors for cancer. There are abundant opportunities for cancer control using this evidence. However, the principal risk factors for some of the most common cancers (e.g., colon cancer and prostate cancer) have yet to be identified, and some of the identified risk factors cannot be readily modified (e.g., reproductive history and risk for breast and uterine cancer). Papers at this meeting emphasized that
new research methods bringing laboratory-based approaches into the population context, so-called molecular epidemiology, offer stronger data in support of cancer control initiatives.

The New Tools

A number of new research tools have the promise of deepening our understanding of the causes of cancer. These tools include the rapidly advancing techniques of modern molecular and cellular biology, which are already providing markers of genetic susceptibility for some cancer sites, and biomarkers of exposure and dose, which can sharpen the characterization of risks in population studies. This blending of markers of susceptibility, exposure, dose, and response into the population context has been called molecular epidemiology.

The tools also include powerful new study designs. Relative risks are anticipated to be relatively modest for many potential risk factors of current concern and quantification of the risks may be compromised by exposure misclassification and confounding. Large studies are needed to address such risks; studies have now been conducted that demonstrate the feasibility of collecting information from large groups and then following the subjects for disease incidence and mortality. In the Nurses Health Study, for example, information on risk factors was obtained by mailed questionnaire on enrollment from over 100,000 nurses who were then followed actively and passively (10).

The subsequent collection of dietary information using a food frequency approach makes this study a unique source of information on diet and cancer. Nested design approaches that involve sampling within larger study populations have also been developed (11). These designs facilitate the informative application of more complicated, invasive, or expensive approaches for assessing exposures or outcomes.

New tools are also available for exposure assessment (12,13). The conceptual basis for exposure assessment has become increasingly formalized and the concept of exposure broadened to encompass the full range of media through which exposure occurs. Exposure assessment technology suitable for epidemiologic research has now been developed for many environmental agents. For example, radon can be readily measured in indoor air using relatively inexpensive passive monitoring devices; thousands of measurements will be made in the case-control studies now in progress throughout the world (14).

Biomarkers are indicators of exposure, dose, outcome, or susceptibility measured in biologic specimens (15,16). Such biomarkers have the promise of heightening the validity of indicators used in cancer research. For smoking and lung cancer, for example, levels of cotinine, a nicotine metabolite, provide an index of immediate exposure, and levels of some carcinogens and adducts reflecting their binding to DNA can be measured (17). Epidemiologic investigations have also incorporated genotyping for putative susceptibility genes.

The Score Card in 1994

The predominant avoidable causes of cancer fall under the broad groupings of tobacco, alcohol, diet and nutrition, hormones and medications, occupational exposures, environmental exposures, radiation, and infectious agents. Gender, ethnicity, and environment are also factors, as are genetic susceptibility and interaction of genetic susceptibility with the environment. Table 1 summarizes my impression of the status of the evidence in each of these areas. I have arbitrarily graded on a four-level scale—from 0 to +++—for the evidence on both components of the PAR, the relative risk, and the exposure prevalence (P). While some might dispute these summary designations, the table provides a perspective on the relative status of the evidence for these factors.

Tobacco smoking is at the highest level of understanding for both the RR and P. The extensive epidemiologic research on tobacco smoking and health provides a rich database on the RR for the many cancer sites causally linked to tobacco smoking (18,19). Patterns of tobacco smoking are also amply documented through diverse population surveys (19). The population burden of smoking-attributable death is regularly tracked and reported; a software system has been developed for this purpose (20).

By contrast, the burden of cancer associated with diet remains uncertain and controversial, although diet was considered to account for a substantial proportion of cancer cases (35%) in the 1981 report of Doll and Peto (21). Dr. Walter Willett reaffirmed this view of the importance of diet and suggested that most of the attributed cases could, in fact, be prevented (23). However, there are abundant hypotheses concerning diet and cancer that reflect the myriad components of the diet that might act to increase or decrease cancer risk and there is still little coherence among the RR estimates for any particular nutrient for a specific cancer site. There have been notable successes in identifying agents in the diet that cause cancer, but the most prominent examples reflect exposures, such as aflatoxin, that have an impact in developing countries.

Table 1 makes clear that a substantial research agenda remains to be met if we are to target resources optimally to prevent avoidable cases of cancer. Research is still needed to quantify the RRs associated with the many potential causes of cancer. We also need evidence from the full array of research approaches used to understand the causes of cancer in man: epidemiologic studies, animal studies, and more basic laboratory approaches.

In general, we know less about patterns of exposure to the many agents that cause cancer. National and other surveys provide information on smoking, alcohol, and diet, although the dietary data may not be sufficiently detailed for all dietary components of potential interest with regard to cancer. Less information is available on exposure patterns for other broad groups of causal agents (Table 1).

The score card is particularly poor for research on risk to special populations. Relative risk values among special populations can be expected to differ from those of white populations and to vary because of patterns of exposure and gene–environment interactions. To date, this anticipated heterogeneity remains largely unexplored.

Research Opportunities and Approaches

An extensive research agenda directed at cancer etiology is needed to complete the many gaps in research pointed out in Table 1. In spite of these gaps, it is apparent that the evidence is already sufficient to justify intervention for many of the categories in...
Table 1 and that action need not be delayed for well-characterized causes of cancer.

Recommendations were offered for each of the broad groups of factors discussed here. In addition to the recommendations made for specific causal agents, the scientific basis for cancer prevention might be generally strengthened by new research strategies directed toward developing new tools for exposure assessment and investigating the risks of mixtures, and toward population surveillance.

Any strategy to prevent the occurrence of avoidable cases of cancer must reflect population patterns of exposure to the agents of concern. Surveys at the local, regional, and national levels now provide insight into exposure distributions for some agents such as cigarette smoking. However, more comprehensive information that characterizes exposures in relation to potential determinants of susceptibility and to gender, ethnicity, and other sociodemographic factors are needed to set priorities and to develop appropriately targeted interventions. Data on key risk factors should span the diverse groups within the population and provide a description of exposure distributions as well as determinants of exposure. Information on exposure distributions might be used to direct interventions toward those at the highest risk and toward those exposures that might be most effectively reduced. Persons at high risk on the basis of genotype might be given the highest priority. A holistic approach toward exposure assessment would also serve to profile cancer risk more generally and would facilitate consideration of combined exposures that may synergistically increase risk.

Investigation of the health risks of combinations of factors has proved equally vexing to scientists using epidemiologic approaches and those using laboratory approaches. Yet, many of the causes of cancer are complex mixtures, like diet and nutrition in general, or tobacco smoking. The combined effects of the multiple agents that cause particular cancers have also challenged our investigative capabilities. Some mixtures, like tobacco smoke, have been considered as though they were a single agent from the public health perspective, and the findings of epidemiologic research on these mixtures have adequately served the purpose of cancer control. By contrast, for some mixtures like diet, the effects of individual components must be characterized for effective cancer control. Observational studies can be used to investigate mixtures, but exposure misclassification and limited statistical power may compromise their findings. Large study populations are one solution to these problems of investigating complex mixtures. However, there may be inadequate opportunity for replication, and exposure misclassification and numbers of cancer cases may still limit the power of seemingly large studies.

Innovative approaches for investigating the effects of complex mixtures have not been forthcoming (23). Biomarkers may improve exposure estimation and more efficient designs may strengthen observational studies. Mixtures of presumed chemopreventive agents can be assessed in randomized trials and randomization strategies can be used to permit assessment of the effects of individual mixtures or combinations of agents. Neither observational studies nor clinical trials alone will be sufficient, and a stronger context for interpreting the effects of mixtures must be found in heightened understanding of the mechanisms by which the components of mixtures work in concert to cause cancer.

Finally, I concur with Dr. Hoda Anton-Culver, who recommends developing population-based approaches for surveillance of risk factors for cancer (24). Public health surveillance incorporates ongoing, systematic data collection with analysis and interpretation of the health data (25). The data serve the purpose of prevention and control, and surveillance can support the iteration of intervention and evaluation. To date, we have missed the opportunity for surveillance of the changing causes of cancer afforded by population-based registries such as those in the National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) Program. Such population-based cancer registries should be recognized as cohort studies of dynamic populations. Supplementing ongoing case ascertainment with information on known and putative causes of cancer gained from samples of the covered population would constitute a surveillance system for tracking changes in risk factors and associated changes in cancer occurrence.

Conclusions

Research on the etiology of cancer has had some notable successes and provides a framework for implementing an agenda of research and intervention for cancer control. However, research strategies to identify avoidable causes of cancer continue to be constrained by the methodologic difficulties posed by the low signal-to-noise ratio for many lifestyle-associated risk factors and the complexity of the environmental exposures of current concern, many of which are complex mixtures. The emerging field of cancer genetics offers a new set of approaches and promises a heightened understanding of mechanisms of cancer causation, and even the possibility of identifying those at risk from environmental exposures on the basis of genetic susceptibility. In addition to the recommendations made for specific causal agents, the scientific basis for cancer prevention might be generally strengthened by new research strategies directed at developing new tools for exposure assessment and investigating the risks of mixtures, as well as at population surveillance.

REFERENCES

1. National Cancer Institute. The SEER Program Code Manual, Revised Edition. Bethesda, MD: National Institutes of Health, 1992.
2. Evans A. Causation and Disease: A Chronological Journey. New York: Plenum Medical Book Company, 1993.
3. DHEW. Smoking and Health. Report of the Advisory Committee to the Surgeon General. Washington, D.C.: U.S. Department of Health, Education and Welfare, 1964.
4. Rothman KJ. Modern Epidemiology. Boston: Little, Brown and Company, 1986.
5. DHHS. The Health Consequences of Involuntary Smoking: A Report of the Surgeon General. Washington: U.S. Department of Health and Human Services, 1986.
6. U.S. EPA. Technical Support Document for the 1992 Citizen’s Guide to Radon. EPA/400-R-92-011. Washington: U.S. Environmental Protection Agency, 1992.
7. Lubin JH, Boice JD, Edling C, Boice JD Jr, Edling C, Hornung RW, Howe G, Kunz E, Kusiak RA, Morrison HJ, Radford EP, Samet JM, Timarch M, Woodward A, Yao XS, Pierce DA, Radon and lung cancer risk: a joint analysis of 11 underground miners studies. Bethesda, MD: National Institutes of Health, 1994.
8. Lubin JH, Boice JD. Estimating radon-induced lung cancer in the U.S. Health Phys 57:417–427 (1989).
9. National Research Council. Health Risks of Radon and Other Internally Deposited Alpha-emitters: BEIR IV. Washington: National Academy Press, 1988.
10. Hennekens C, Speizer F, Rosner B. Use of permanent hair dyes and cancer among registered nurses. Lancet 1:1390–1393 (1979).
11. Prentice RL. A case-cohort for epidemiologic cohort studies and disease prevention trials. Biometrika 73:1–11 (1986).
12. Armstrong B, White E, Saracci R. Principles of Exposure Measurement in Epidemiology. New York: Oxford University Press, 1992.
13. Samet JM, Stolwijk J, Rose SL. Summary: International Workshop on Residential Radon Epidemiology. Health Phys 60:223–227 (1991).
14. Hulka BS, Wilcosky TC, Griffith J. Biological Markers in Epidemiology. New York: Oxford Press, 1990.
15. Links J, Kensler T, Groopman J. Biomarkers and mechanistic approaches in environmental epidemiology. Annu Rev Public Health 16:83–103 (1994).
16. World Health Organization IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Lyon: International Agency for Research on Cancer, 1986.
17. DHHS. Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General. DHHS Publ No (CDC) 89–8411. Washington: U.S. Department of Health and Human Services, 1989.
18. Centers for Disease Control. Smoking-attributable mortality and years of potential life lost—United States, 1984. MMWR 36:693–697 (1987).
19. Doll R, Peto R. The Causes of Cancer. Oxford: Oxford University Press, 1981.
20. Willett WC. Diet, nutrition, and avoidable cancer. Environ Health Perspect 103(Suppl 8):165–170 (1995).
21. Samet J, Speizer F. Introduction and recommendations: Working Group on Indoor Air and Other Complex Mixtures. Environ Health Perspect 101:143–147 (1994).
22. Anton-Culver H. Cancer prevention strategies: the use of cancer prevention research registries. Environ Health Perspect 103(Suppl 8):237–239 (1995).
23. Thacker S, Berkelman R. History of public health surveillance. In: Public Health Surveillance (Halperin W, Barker EJ, Monson R, eds). New York: Van Nostrand Reinhold, 1992.