promoting responsible environmental, health, and safety decision-making

Epidemiology In Decision-Making

The Role of Epidemiology in Decision-Making

Legislatures and administrative agencies considering environmental and public health issues frequently must evaluate methods, analyses, and conclusions of epidemiological (EPI) and other scientific evidence which provides a basis for expert opinion studies. To provide guidance for decision-makers, The Annapolis Center convened a workshop in June 1998 of thirteen scientists, doctors, and lawyers with extensive training and experience in epidemiology, toxicology, pharmacology, and forensic use of scientific evidence.

Recognizing that sufficient literature for specialists already exists, the workshop group agreed to produce a primer for non-scientists who seek to understand epidemiologic studies. In that spirit, the group debated and eventually arrived at a number of observations concerning epidemiological studies. Those observations include the following:

- Today, epidemiology is formally understood as the study of the distribution and determinants of diseases in humans.

- No two EPI studies are identical, and no given study can be replicated exactly.

- Absolute risk is a more useful measure for legislators and policy-makers than relative risk because it shows the foreseeable impact of exposure to the risk factor.

- Nevertheless, in some situations, relative risk is more useful to courts and those required to evaluate causation.

- To epidemiologists, *association* means only that a risk factor and the disease occur together. It does not necessarily mean the factor *causes* the disease. Co-incidence is not proof of cause.

- Even if it finds an association between a suspect factor and disease, a poorly designed study is of questionable value to decision-makers because chance, bias, and confounding cannot be excluded as explanations for the association. However, some design errors may not be fatal, and the extent and direction of error can be estimated in some cases.

- In a well-designed study, the investigator will foresee the likely sources of bias and take steps to control them to the degree practicable.

- The meaning of *cause* in scientific inquiry often differs from the meaning of *cause* in legal proceedings. Accordingly, legal decision-makers must first determine that the evidentiary weight of relevant epidemiologic data is appropriate to the issues before them.

- Although EPI studies never prove causation, either generally or in a specific case, they can show cause to be more (or less) likely as a potential explanation for an observed association between risk factor and disease. Furthermore, EPI studies *per se* do not prove safety.

- **Determination of causation requires a weight-of-evidence approach that considers epidemiology, biologic mechanisms of action, relevant toxicology, and other factors.**

The group developed a series of cautions for users of EPI research. These cautions centered on:

- Scrutiny of investigator credentials whenever an EPI study is used in decision-making.
- Limitations in the applicability of EPI studies, given the parameters of design of a particular study.
- Reliance on databases and their reliability.
- Extrapolation of animal studies to the question of disease causation in humans.
- Reliance by EPI investigators upon research findings in other disciplines.
- Differing degrees of certainty in defining and measuring adverse effects following exposures of specified intensity and duration.
- Consistency of findings between and among studies.
- Expertise of peer reviewers and their ability to judge the quality of studies they review.
- Design of meta-analyses, especially the comprehensiveness of source information.

Lastly, the group developed the following suggestions for policy-makers, regulators, and courts for optimum use of EPI in legal proceedings:

- Consider the use of neutral, advisory experts in epidemiology and allied fields (*e.g.,* toxicology and statistics).
- Allow appropriately educated legislators and regulators to make final decisions on public-health policy that may have major economic impact. These decisions should be made only with explicit evaluation of the costs of such proposals measured against reduction of risk that may be achieved.
- Disclose the EPI and other evidence cited to justify proposed regulation, and consider application of Daubert to judicial review of the science upon which administrative agencies rely.
- Develop and apply uniform standards for the identification, characterization, and assessment of risks.
- Urge judges who function as "gate-keepers" of scientific evidence to make greater efforts to scrutinize the quality of research underlying an expert's opinion.

The Role of Epidemiology in Decision-Making

The Annapolis Center
1999

William Braithwaite, J.D.
Philip Cole, M.D., Dr. P.H.
Alvan R. Feinstein, M.D., M.A.C.P.
Michael Green, J.D.
Robert Hirsch, Ph.D.
Claire M. Lathers, Ph.D., F.C.P.
M. Stuart Madden, J.D.
M. Gerald Ott, Ph.D.
Gerhard K. Raabe, Dr.P.H., M.S.
Alan Charles Raul, J.D, M.P.A.
Ford Rowan, J.D.
Jack Snyder, M.D., J.D., Ph.D. (Workshop Chairman)
Thomas Starr, Ph.D.

Table of Contents

Introduction

Epidemiology is the study of which groups of people get which diseases and why. EPI is increasingly a source of the scientific evidence that legislatures, administrative agencies, and courts consider when they have to decide whether a substance is toxic enough to cause illness or death.

This primer on epidemiology was created for these decision-makers and their professional staffs – counsel to Committees of Congress and of State legislatures, agency legal staffs, and judicial clerks in federal and state courts. It may also be useful to others without special training in science, such as journalists and the public, who may have to judge the relevance and reliability of scientific or media reports about the risks of specific foods, medicines, activities, occupations, or environments.

Legislatures and administrative agencies considering environmental and public health issues frequently must evaluate methods, analyses, and conclusions of epidemiological (EPI) and other scientific evidence that provides a basis for expert opinions.

A typical case concerning the use of EPI data in litigation is the U.S. Supreme Court decision in General Electric Co. v. Joiner. In Joiner, a city electrician sued the manufacturer of the transformers he worked on, claiming that the polychlorinated biphenyls (PCBs) produced by the transformers caused his lung cancer. The trial court ruled that the opinions of Mr. Joiner's experts on causation were not admissible evidence because the EPI studies underlying these opinions were unreliable.

One of the studies, of cancer mortality among capacitor workers, involved a group of subjects too small to permit statistically significant estimates of risk. Another study of capacitor workers did show an increase in lung cancer deaths but did not examine PCB-exposure as a possible cause. A third study failed to rule out cigarette smoking as a possible cause.

The District Court granted summary judgment to General Electric, but the Court of Appeals disagreed with the lower court's assessment of the scientific evidence. General Electric appealed to the Supreme Court. In finding for General Electric, the Supreme Court relied on its own decision in Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 597 (1993). In Daubert, the Court construed Federal Rule of Evidence 702 (dealing with expert witness testimony) to require that federal trial courts admit scientific evidence only after determining that it is both relevant and reliable.

The rules of evidence in many states are modeled on the Federal Rules of Evidence. Consequently, trial courts in those states that decide to follow Daubert also take on the role of "gatekeeper". Legislatures and administrative agencies considering environmental and public health issues already find themselves having to judge the methods, analyses, and conclusions of EPI and other scientific studies upon which expert witnesses base their opinions. Yet the professional staffs on which these decision-makers routinely rely seldom have scientific training

or skills, or even a working familiarity with the terms, concepts, methods, and reliability criteria of EPI research.

In light of these developments in the law, The Annapolis Center convened a workshop in June of 1998 of thirteen scientists, doctors, and lawyers with extensive training and experience in epidemiology, toxicology, pharmacology, and the forensic use of scientific evidence. Recognizing that sufficient literature for specialists already exists, the workshop group agreed to produce a primer for non-scientists who seek to understand epidemiologic studies. This primer explains what EPI evidence is composed of, what its strengths and limitations are, and how it can best be used by legal decision-makers.

What is Epidemiology and Who Uses It?

In the early 20th century, "epidemiology" meant almost exclusively the study of epidemics of diseases such as smallpox, malaria, and typhoid fever, which are spread through infection by bacteria or viruses. In the United States and other advanced countries, most of the epidemic diseases of that time have now been eradicated or controlled. Since mid-century, it is chronic, rather than infectious diseases that have been studied the most.

Today, epidemiology is formally understood as the study of the distribution and determinants of diseases in humans.

A single event, such as an insect bite or ingestion of an air- or food-borne virus, can start an infection. Chronic diseases, in contrast, develop over time, often years (*e.g.,* various cancers and respiratory disorders), and they typically involve repeated exposure. They may be caused or influenced by agents found in foods and medicines, the home and the work place, or even in the natural environment.

Today, epidemiology is formally understood as the study of the distribution and determinants of diseases in humans. The two main branches of the subject correspond to the two elements of this definition.

Descriptive epidemiology tries to identify which groups of people get which diseases. Descriptive studies simply report the actual distribution of disease in different *populations.* Children exposed to environmental lead, workers exposed to particular job-site chemicals (*e.g.,* benzene, PCBs, asbestos), and users or consumers of the same food, prescription drug, or municipal water supply are examples of *populations.* Epidemiologists doing descriptive studies use census data, demographic information, death certificates, health and autopsy records, and other sources in their efforts to identify patterns of disease distribution.

Analytical EPI attempts to identify the reasons why certain groups develop certain diseases. Often, the investigator has used descriptive studies as the basis for hypotheses about what causes a disease or makes its onset more likely. An analytical study seeks to identify the specific agents or events that may be associated with the development of disease, and to assess the degree of risk, if any, that may result from exposure to the suspected hazards.

The most direct way to identify hazards that may cause or contribute to disease would be an experiment in which a group exposed to the suspect agent or event is compared to another group that has not been exposed. Medical ethics, however, does not permit this kind of experiment with human beings. Consequently, the analytical epidemiologist must gather data not by experiment but by observation. The researcher selects for study a sample of people who have allowed themselves to be exposed to the suspect agent (*e.g.,* cigarette smoke) or who have been exposed to an agent unknowingly (*e.g.,* radon). By studying the health histories of persons in the sample, the researcher hopes to discover *risk factors* that make onset of disease more likely.

Among the important concerns of legislators, regulatory agencies, and courts are the risk factors associated with particular diseases. Consequently, analytical, rather than descriptive, EPI studies are typically of greater interest to these decision-makers.

How Epidemiological Research Is Done

The investigator in an EPI study usually begins with a hypothesis about potential risk factors associated with a disease. The researcher chooses a population having the relevant characteristics of age, race, sex, health history, social and economic status, geographical distribution, and exposure to the suspect agent or condition. Using standard statistical methods, the researcher selects a sample of this population. After the relevant data from this sample are collected (by interviews, written questionnaires, medical examinations, or telephone surveys), they are tabulated and interpreted.

No two EPI studies are identical, and no given study can be replicated exactly.

Standards for interpreting data, judgments of relevance, and criteria of probative value may differ among investigators studying the same disease and risk factors. Standards, judgments, and criteria also may vary among different studies performed by the same investigator. Consequently, no two EPI studies are identical, and no given study can be replicated exactly. The uniqueness of each observational setting adds an element of complexity to assessment of the internal and external reliability of a study's conclusions.

Conventional science tests the reliability of conclusions by examining whether a number of different investigators doing the same experiment arrive at the same or similar findings. This kind of verification is not possible in epidemiology because its methods cannot include experiments that systematically expose people to suspected toxins. On the other hand, preventive measures, such as vaccines or new drugs, can be studied experimentally in *clinical trials,* where patients who are already ill are randomly assigned to receive either the treatment under study or a placebo (or the standard treatment, if any).

The Meaning of *Risk* in Epidemiology

Risk means likelihood and a *risk factor* is anything that increases the likelihood of disease. Since epidemiologists say that a person with a risk factor is *exposed* to the risk of developing the disease, risk factors are also called *exposures*. In a narrower sense, *exposure* means how great, how often, and how long was a person's contact with a risk factor (*e.g.,* cigarettes, asbestos, or lead).

Absolute risk often is a more useful measure for legislators and policy-makers than relative risk because it shows the foreseeable impact of exposure to the risk factor.

Assessment of the kind and magnitude of risk posed by a given exposure can be expressed as either an absolute or a relative risk. A statement of *absolute* risk indicates the percentage of the exposed population (*e.g.,* smokers) who will get the disease (*e.g.,* lung cancer). In public and private health decisions, absolute risk often is a more useful measure than relative risk because it shows the foreseeable impact, and therefore the probable health-care costs, of exposure to the risk factor.

Relative risk is more useful to decision-makers who must evaluate causation.

Relative risk, on the other hand, is more useful to decision-makers who must evaluate causation. A statement of *relative* risk tells how much more or less likely it is that people with the risk factor (*e.g.,* smokers) or suspect characteristic will get the disease, when compared with those not having the risk factor (*e.g.,* non-smokers). For example, "smokers are ten times more likely to get lung cancer than are non-smokers" is a statement of relative risk.

When an EPI study finds an association between a suspect agent and a disease, the investigator typically expresses the magnitude of the association (*i.e.,* the relative risk) by a number (*e.g.,* 2.4). Depending on the kind of study, this number is called the standardized mortality ratio (SMR), odds ratio, or relative risk.

To epidemiologists, *association* means only that a risk factor and the disease occur together. It does not mean the factor *causes* the disease.

Thus, the relative risk number compares the occurrence of the disease in two groups: one group of people exposed to the suspect agent and another group of people not exposed. A relative risk ratio of 1.0 means that these two groups were found to develop the disease at the same rate. That is, the suspect factor has not been shown to be a true *risk factor,* which is to say it has a *null effect.*

By contrast, a relative risk ratio exceeding 1.0 suggests that exposure increases the risk of getting the disease, and the higher this number, the greater the risk. Relative risk ratios therefore show how strong the association is between risk factor and a disease.

***C**o-incidence is not proof of cause.*

To epidemiologists, *association* means only that a risk factor and the disease occur together. It does not mean the factor *causes* the disease. A relative risk ratio

exceeding 1.0 shows that the risk factor occurs with the frequency or at the rate the ratio indicates. But co-incidence is not proof of cause. For example, we do not cause the sun to rise by getting up at the same time it does.

Even if it finds an association between a suspect factor and disease, a poorly designed study is of questionable value to decision-makers because chance, bias, and confounding cannot be excluded as explanations for the association. Some design errors, however, may not be fatal, and the extent and direction of error can be estimated in some cases.

Epidemiologists attribute a demonstrated association to chance, bias, confounding, and/or causality. In a well-designed study, the investigator tries to reduce the influence of the first three, leaving cause as the most likely explanation of the demonstrated association. Even if it finds an association between a suspect factor and disease, a poorly designed study is of questionable value to decision-makers because chance, bias, and confounding cannot be excluded as explanations for the association. Some design errors, however, may not be fatal, and the extent and direction of error can be estimated in some cases.

Chance and "*p* Values"

Epidemiologists investigate groups, not individuals, and all EPI research uses samples selected from larger populations. Since no sample ever completely mirrors the whole of which it is a part**,** some of the observed differences between two samples can always be the chance effect of sampling**.** Consequently, epidemiologists use "standard" statistical tests to estimate how much uncertainty there could be in their findings due to sampling effect.

Statistical significance tests in epidemiology have been devised to assess the compatibility of a set of data with the null hypothesis (H_o) that a population exposed to agent "x" and a population not so exposed do not differ in the incidence or prevalence of condition "y" (implying that agent "x" does not cause condition "y"). In the process of "accepting" or "rejecting" H_o, investigators can make one of four different decisions based on the result of the statistical test. First, they can accept H_o when it is actually true. Second, they can accept H_o when it is actually false. Third, they can reject H_o when it is actually true. Fourth, they can reject H_o when it is actually false.

Rejection of H_o when it is in fact true is called a type I error, and the probability of making this error is called alpha. Acceptance of H_o when it is in fact false is called a type II error, and the probability of making this error is called beta. Investigators typically select a value for alpha, known as the *p*-value, prior to evaluating their data. By contrast, the value for beta depends on how much the true situation deviates from H_o. The greater the true hypothesis deviates from H_o, the smaller the value of beta. If beta is the probability of making an incorrect decision when H_o is false, then 1 minus (ß) is the probability of making a correct decision when H_o is false. This probability is called the power of a statistical test, and power increases the more the true hypothesis deviates from H_o. Importantly, the power of a statistical test is also a function of the chosen *p*-value, the variance involved, and the sample size. The concept of power is extremely important in the interpretation of statistical tests.

The interpretation of statistical tests typically begins when a statistic (e.g., *t*, chi-square, etc.) that summarizes the evidence against the null hypothesis (H_o) is calculated and compared to the distribution of such statistics if H_o is true. If the calculated statistic is judged to be too unlikely under H_o, then H_o is rejected. Otherwise, H_o is accepted at a stated level of significance known as the *p*-value.

Obtaining a statistically significant difference (say, $(p) < .05$, $p < .01$) between exposed and unexposed populations indicates that differences as large as, or larger than, those observed may occur with "too small" a probability under H_o to be reasonably attributed solely to chance. Conversely, obtaining a test result that indicates no statistically significant difference (say, $(p) > .05$, $p > .01$) between exposed and unexposed populations implies that differences as large as, or larger than, those observed may occur under H_o with "too large" a probability for the investigator to rule out the null hypothesis. Epidemiologists typically specify what they mean by "too small" or "too large" by selecting a particular *p*-value.

Consequently, a difference observed between the results of two populations that is judged to be statistically significant at $(p) < .05$ means that differences as large as, or larger than, that observed in the study would occur by chance alone less than 5% of the time. Conversely, a

difference observed between the results of two populations that is judged not to be statistically significant at (p) > .05 means that differences as large as, or larger than, that observed in the study would occur by chance alone more than 5% of the time.

A null hypothesis is neither proved nor disproved by any statistic evaluated at an arbitrarily chosen level of significance (p-value). "Acceptance" and "rejection" of H_o are merely terms that relate to specific probability statements computed under the hypothetical condition that exposure to substance "x" does not cause condition "y."

Finally, statistical significance is not equivalent to practical, clinical, or biological significance. Statistical significance pertains only to the existence of a difference, not its magnitude. To judge the practical significance of a finding, estimated magnitudes of differences must be considered in light of all accumulated evidence known to the investigator. Conversely, real and important differences may be missed even when data do not yield a statistically significant difference at a conventional level of significance (usually 0.05). The failure to obtain a statistically significant difference does not prove a real difference does not exist; it only shows that the observed difference easily could be explained by chance alone. Small sample sizes, large population variability, and small but real differences can all decrease the ability to use statistics to distinguish a real difference from random processes. If the EPI investigation had a very low power, for example, important, true differences between populations may not have been detectable.

Bias and Confounding

What epidemiologists mean by bias is a distortion of the real relationship between risk factor and disease. Bias can result from mistakes in selecting the study population, in choosing the people in the sample ("selection bias"), or in classifying them as sick or well, exposed or not exposed. Clerical error, omissions in data collection, and imperfect or poorly performed tests for disease or exposure are common causes of misclassification.

In a well-designed study, the investigator will anticipate the likely sources of bias and will take steps to control them to the degree practicable.

Bias can also result from the recognized tendency of sick people to remember a non-existent exposure to the suspect agent ("recall bias") and from the tendency of investigators and interviewers to see what they want to see ("observer bias"). In a well-designed study, the investigator will anticipate the likely sources of bias and will take steps to control them to the degree practicable.

A "confounder" is another factor actually associated with both the putative risk factor and the disease, but not otherwise considered by the investigator. For example, we know that both smoking and workplace chemicals can be real or suspected causes of certain cancers. Suppose an investigator found a very strong association (risk ratio exceeding 3.0) between, say, PCBs and lung cancer, and the chosen *p* value was well under .05. Even if sources of bias in the study are well-controlled, the association would have little evidentiary value if the investigator failed to design the study to rule out smoking as a possible confounder. The trial court in Joiner found one of the studies cited by plaintiff's experts to have just such a defect.

Cause

Cause means something different in science from what it means in law. Different decision-makers have different standards for determining cause.

A variety of statistical, mathematical, and practical techniques are available to help epidemiologists minimize the effects of chance, bias, and confounding. When the investigator has used these techniques to the extent feasible, the most likely remaining explanation for a demonstrated association between risk factor and disease is cause. *Cause* means something different in science from what it means in law. Different decision-makers have different standards for determining cause.

Many scientists test conclusions by repeating experiments. When many different investigators make the same or similar findings, a conclusion is taken to have been proved true. The standard of proof is consensus of qualified opinions, verified by repeatable experiments.

Epidemiology is not an experimental science. Nor does it study individual cases in isolation. Its most meaningful results are statistical: the happening together of a suspect agent or event ("cause") and a known identifiable effect. Although EPI studies cannot prove causation, either generally or in a specific case, they can show cause to be more (or less) likely as a potential explanation for an observed association between risk factor and disease.

Epidemiology studies cannot prove causation, either generally or in a specific case, although they can show cause to be more (or less) likely as the explanation of a demonstrated association between risk factor and disease.

EPI studies cannot prove safety. Nevertheless, regulators and other legal policy-makers sometimes conclude from negative ("null effect"') studies and other evidence that a suspect agent does not really pose a significant risk. For example, epidemiologists cannot show that Bendectin will never cause birth defects, but the substantial body of studies that have been done has persuaded most people that very little risk exists. It was EPI studies combined with toxicological research that persuaded the Food and Drug Administration to remove sodium saccharin from its status as a suspected carcinogen.

Epidemiological studies cannot prove safety.

When an original, properly designed, and properly executed EPI study finds a strong association and lack of random effects, and animal studies show a positive correlation between dose and response, cause can be presumed as the most likely explanation. A detailed set of standards for assessing causation in individual studies can be found in a protocol known as the *Hill Criteria* (see bibliography). When several EPI studies find a strong association, and laboratory investigations support a known biological mechanism of action, or suggest a plausible one, the presumption is even stronger that the risk factor is a likely cause of the disease.

In the case of a single individual with the disease, one or the other of the two foregoing types of EPI evidence, a proven exposure, and the absence (or minimal effect) of any alternative cause

can, when taken together, provide sufficient evidence to satisfy tort law's *more probable than not* standard for proof of causation.

A thoughtful and useful examination of causality in all three of these contexts – the single study, the general case, and the individual plaintiff – can be found in Dr. Cole's article (see bibliography).

Cautions For Users of Epidemiological Research

Investigator Credentials Epidemiology is a recognized academic specialty, but epidemiologists are not examined, licensed, or certified by any government agency or professional organization. The user of an EPI study should consider the education and experience of the investigator, and compare these credentials to those with recognized expertise in the field. The American College of Epidemiology reviews contributions of epidemiologists and names as fellows of the College those with recognized accomplishments.

Study Design Epidemiology is observational, not experimental (exact replication of a given study is simply not possible) and the conclusions of a poorly designed study are of little value. Users of a particular study should scrutinize the research design, including use of (or failure to use) standard techniques for controlling the effects of chance, bias, and confounding. The criteria of data-interpretation also should be examined carefully.

Data Bases Some studies use published data bases, not all of which are equally reliable. Users of such studies should ask many of the same questions about these data bases that they ask about the study itself.

Animal Studies Even the most carefully performed animal studies cannot directly address either absolute or relative risk to human beings. Users should ask whether there are sufficient grounds for extrapolating from animals' doses and disease-responses to humans.

Allied Sciences An EPI study can be either corroborated or called into question by research in other sciences, especially toxicology, biochemistry, and pharmacology. Users should determine whether such research exists and if it does, consider its relevance to the epidemiological conclusions at issue.

Absolute Risk Legislatures, agencies, courts, and all other users of EPI research should always ask what is the absolute risk of harm, *i.e.,* how many exposed individuals became ill in a specified period of time versus how many of these individuals would have become ill if not exposed to the particular hazard. Where severity of harm varies with level of exposure, they should determine what degree of harm correlates with what dose-levels.

Consistently Strong Association In judging relative risk, the single most important factor is a consistently observed, strong association between risk factor and disease (relative risk exceeding 3.0). The inference of cause and effect is strengthened if the demonstrated statistical association is consistent with known or plausibly hypothesized biological mechanisms, and is observed in more than one study.

Peer Review Virtually all the EPI studies that legislators, regulatory agencies, and courts rely upon are published in peer-reviewed journals, but the mere fact of such

publication does not attest to a study's reliability in any major way. Peer review is essential to the research process and works well when a paper is personally reviewed by a specialist in the investigator's own field. However, time constraints for publication often cause the review to be done by someone less experienced or qualified, and the large number of journals increases the chances that a paper can get published somewhere eventually.

Meta-Analysis

Non-randomized observational studies (*i.e.,* those based on samples selected from exposed and unexposed populations) are sometimes aggregated for *meta-analysis*, a study of studies. If it discriminates the more and the less reliable among the underlying studies, and uses *sensitivity* analysis to identify the effects of one or a few studies with large samples but lower-quality research designs, a meta-analysis can help reduce random error. Meta-analysis is otherwise of limited value and does not take the place of a comprehensive review of the relevant literature.

Suggestions For the Best Use Of Epidemiology in the Law

Consider the use of neutral, advisory experts in epidemiology and its allied fields.

Legislators, regulatory agencies, and courts should, when practicable, consider the use of neutral, advisory experts in epidemiology and allied fields (*e.g.,* toxicology and statistics).

Public-health policy decisions that may have a major economic impact should be made by appropriately educated legislators and regulators.

Appropriately educated legislators and regulators should make final decisions on public-health policy that may have major economic impact. These decisions should be made only with explicit evaluation of the costs of such proposals measured against reduction of risk that may be achieved.

When a reliable risk assessment shows a large cost for a small reduction in risk, the decisions whether to legislate or regulate become how much, who pays, and what entities should bear the cost of regulatory compliance.

Fairness requires that agencies disclose the EPI and other evidence claimed to justify proposed regulation. The Daubert approach could be usefully applied in the context of judicial review of the science relied upon by federal agencies.

Suspect hazards should be shown to have a clear association with illness before regulation occurs. Yet agencies may sometimes judge that the protection of public health requires regulation, even when the evidence is not strong, much less conclusive.

Fairness requires that agencies disclose the EPI and other evidence claimed to justify proposed regulation, so that those who will bear the costs can understand the basis for decisions and be better prepared to specifically challenge rulemaking that they question.

Develop and apply uniform standards for the identification, characterization, and assessment of risks.

Consistent application of uniform standards for the identification, characterization and assessment of risks from suspect hazards should make regulation more efficient, encourage compliance, and facilitate changes in regulation when new scientific knowledge so requires (*e.g.,* the down-grading of saccharine from carcinogen status).

Use performance standards in preference to engineering standards.	When regulation is judged necessary, the agency should consider whether stating the public-health goal required to be achieved (a "performance" standard) is preferable to specifying the technology by which this goal must be achieved (an "engineering" or "command and control" standard).
Judges who function as "gate-keepers" of scientific evidence should make greater efforts to scrutinize the quality of research underlying an expert's opinion.	In litigation, EPI studies are not themselves admitted into evidence; they simply provide the basis for expert witness opinions. The tort law's ″more likely than not" standard of proof for causation correlates with a relative risk ratio exceeding 2.0, a ratio that means that a member of the exposed population is twice as likely to get sick as someone not exposed. Statistically, a 2.0 risk ratio is rather low, and could be accounted for by many factors other than a causal connection between suspect agent and disease. Courts taking on the "gate-keeper" duty under Daubert should carefully scrutinize the quality of research underlying an expert's opinion.

Suggestions for Further Study and Action

- To raise the quality of EPI research, improve study design and execution, and help reduce misinterpretation of research results by the public and other non-specialist users, one or more professional groups of epidemiologists should develop minimum standards for the credibility of research (especially exposure assessment) and standard definitions of terms of art.

- Such a group could also fruitfully investigate, at both the theoretical and the practical level, fundamental issues common to science and the law, e.g., what constitutes "evidence," what "proof" is, the meaning of "cause," etc.

- An inter-disciplinary group, such as the one assembled for the June 1998 Annapolis workshop, could produce a non-technical set of recommendations and check-lists for judging the quality of EPI research, addressed to legislators, regulatory agencies, and courts (the main audience for this primer).

- The results of the initiatives described above could be used to help educate the public and news media.

- These standards, recommendations, and checklists could also, in a further educational effort, be tailored to the different processes of particular regulatory agencies that rely upon epidemiological research, e.g., FDA, EPA, and OSHA.

- Further consideration should be given to whether having the parties in tort litigation share the cost of securing the opinion of a neutral advisory panel of experts on the relevance and reliability of particular EPI studies, could help the disputants assess the merits of the claims at issue, and thereby foster resolution without the time and expense required for judicial decision.

Biographies of Epidemiology Workshop Participants

William Braithwaite, J.D. is a Tutor at St. John's College, in Annapolis, Maryland. At Loyola Law School, in Chicago, he taught Professional Ethics, Evidence, Remedies, Torts, and other courses from 1979-95. Prior to that he practiced law in Chicago.

Philip Cole, MD, DrPH is a professor of epidemiology at the University of Alabama at Birmingham. Dr. Cole's major interests lie in chemical and hormonal carcinogenesis and in issues of causation in epidemiology. He has published nearly 200 papers in these areas.

Alvan R. Feinstein, MD, MACP is Sterling Professor of Medicine and Epidemiology at the Yale University School of Medicine, where he is also Director of the Clinical Epidemiology Unit and the Robert Wood Johnson Clinical Scholars Program. In his research on care of patients, he has developed new clinical investigative techniques and clinical epidemiological approaches and methods that have been reported in several books.

Michael D. Green, J.D. is Professor of Law at the University of Iowa. He teaches Products Liability, Mass Torts, and Complex Litigation. He is the author and co-author of several books including the Reference Guide on Epidemiology in the Federal Judicial Center's Reference Manual on Scientific Evidence, a work prepared for the federal judiciary. In addition, he has written a number of articles in the area of Products Liability, Toxic Substances Litigation, and the use of scientific evidence as proof in legal cases.

M. Stuart Madden, J.D. is Charles A. Frueauff Research Professor and Distinguished Professor of Law at Pace University School of Law. Professor Madden is the author or co-author of several books and numerous articles on Torts, Environmental Torts and Products Liability subject matters. He is an elected member of the American Law Institute.

M. Gerald Ott, Ph.D. is Director of Epidemiology at BASF Corporation. Dr. Ott has conducted occupational health studies over a period spanning nearly 30 years. Previously, he was a commissioned officer in the U.S. Public Health Service assigned to the National Center for Health Statistics. He has published numerous studies examining the relationships between occupational exposure to a variety of substances and health outcomes ranging from cancer to targeted clinical endpoints. He has also published widely on approaches to linking industrial hygiene and health outcome data.

Gerhard K. Raabe, Dr.P.H., M.S. is Director, Medical Information and Health Risk Assessment for Mobil Business Resources Corp., Global Medical Services. Prior to joining Mobil, he was a Senior Research Scientist for New York State responsible for the Epidemiology Statistical Resources Section attached to Columbia University. He has been a consultant and author in research methods, occupational epidemiology, cancer classification, and health effects of gasoline and benzene, ethical behavior for epidemiologists and medical information systems. He is a Fellow of the American College of Epidemiology.

Alan Charles Raul, J.D., M.P.A. is a partner in the Washington, D.C. office of the international law firm Sidley & Austin. His practice involves litigation, advocacy and counseling in connection with federal government regulation, enforcement and investigations. Mr. Raul has served as General Counsel of the U.S. Department of Agriculture, and as General Counsel of the Office of Management and Budget in the Executive Office of the President. From 1986-1988, Mr. Raul served as Associate Counsel to the President.

Thomas B. Starr, Ph.D. is a principal in the Health Sciences Division of ENVIRON International Corporation. His research has focused on means for explicitly incorporating knowledge of toxic mechanisms into the quantitative risk assessment process, and improving epidemiologic methods for assessing effects of chemical exposure on worker health. He has published over 80 scientific papers and given hundreds of scientific presentations. Dr. Starr holds an adjunct faculty appointment in the Department of Environmental Sciences and Engineering in the School of Public Health at the University of North Carolina, Chapel Hill.

Annapolis Center Board Members Participating in the Workshop

Robert Hirsch, Ph.D. is a physicist and an engineer. He is a consultant with Advanced Power Technologies with considerable experience in virtually all aspects of energy in government and industry. He currently serves as Chairman of the Board on Energy and Environmental Systems at the National Academy of Sciences.

Claire Lathers, Ph.D., F.C.P. is the Chief Scientific Officer for Barr Laboratories. She teaches employees at Barr Laboratories about the clinical pharmacological aspects of drugs that the company will dose. Previously she served as President and Dean of the Albany College of Pharmacy. She has achieved international recognition for her work in the two areas of cardiovascular autonomic dysfunction associated with space flight and with sudden death in persons with epilepsy.

Ford Rowan, J.D., is an expert in crisis management. He is a lawyer with a decade of experience as a network television reporter who has successfully managed dozens of health, environmental, safety and financial issues for corporate clients. Rowan is a former NBC news correspondent and host of "International Edition," a weekly program on public TV. He is the principal author of "Crisis Prevention, Management, and Communication". He has written dozens of articles on such varied topics as news ethics and information technology.

Jack Snyder, M.D., J.D., Ph.D. is a physician-attorney with training and experience in pharmacology, toxicology, pathology, and occupational medicine. He is currently regional director for SmithKline Beecham Clinical Laboratories. Previously, Dr. Snyder taught occupational medicine, toxicology, pathology, and health law at Thomas Jefferson University. In addition, he is a frequent lecturer, advisor and consultant to corporate, academic, legal and governmental bodies in matters involving legal medicine, forensic science, laboratory medicine, toxic torts, workers' compensation, hazardous waste, occupational disease, disaster planning, and

adverse drug reactions. Dr. Snyder served as the chairperson for this Annapolis Center workshop.

Bibliography

Cole, Philip, *Causality in Epidemiology, Health Policy, and Law*, The Environmental Law Reporter, 27(6): 10279-10285, 1997.

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 597 (1993).

Federal Focus, Inc., *Principles for Evaluating Epidemiologic Data in Regulatory Risk Assessment,* 1996.

Graham, John D., *The Role of Epidemiology in Regulatory Risk Assessment.* Elsevier, 1995.

Hill, Austin B., *The Environment and Disease: Association or Causation?* Proc. Royal Soc. Med., Vol. 58, p 295, 1965.

Raul, Alan Charles, "Junk Science In, Junk Policy Out: Science and Administrative Law", New York Academy of Sciences (August 1998).

Raul, Alan Charles, "Scientific Risk Assessment and Judicial Review: Deferring to Uncertainty", Shepard's Expert and Scientific Evidence Quarterly (Summer 1994).

Raul, Alan Charles and Smith, Stephen F., *Judicial Oversight Can Restrain Regulators' Use of Junk Science*. Washington Legal Foundation (January 8, 1999).

Taubes, Gary, *Epidemiology Faces Its Limits*. The American Association for the Advancement of Science, *Science* (July 14, 1995, Vol. 269).

Wong, Otto and Raabe, Gerhard K., *Applications of Meta-Analysis in Reviewing Occupational Cohort Studies.* Occupational and Environmental Medicine 1996; 53:793-800.

About The Annapolis Center

The Annapolis Center supports and promotes responsible environmental, health, and safety decision-making.

The Center evaluates risk and cost-benefit analysis both to assist the public in understanding hazards and the relative risks they may present and to identify areas for emphasis in research and policy. The Center's *Annapolis Accords* provide vehicles to evaluate the quality of science underlying risk analysis and the quality of the policy foundation supporting risk management, as well as cost-benefit analysis. The Annapolis Center is a non-profit, 501(c)3 educational organization.

VADM Harold Koenig, MC, USN (Ret.)
Chair and President

The Annapolis Center
111 Forbes Street, Suite 200
Annapolis, Maryland 21401

410-268-3302
410-268-4953 (fax)
www.annapoliscenter.org