REPORT OF DAVID OZONOFF, MD, MPH

PROFESSOR OF ENVIRONMENTAL HEALTH

CHAIR EMERITUS, DEPARTMENT OF ENVIRONMENTAL HEALTH

BOSTON UNIVERSITY SCHOOL OF PUBLIC HEALTH
715 ALBANY STREET
BOSTON, MA 02118

I. INTRODUCTION AND SUMMARY OF OPINIONS

My name is David Michael Ozonoff. As detailed below in my *Curriculum Vitae*, attached as Exhibit A, I am a physician and a Professor in the Department of Environmental Health, Boston University's School of Public Health and Chair Emeritus of the Department, which I founded and guided for 26 years. My business address is the Department of Environmental Health in Boston University's School of Public Health, 715 Albany Street, Boston, MA 02118.

I have been asked by the Collins Law Firm of Naperville, Illinois and Varga, Berger, Ledsky, Hayes and Casey of Chicago, Illinois to offer an opinion whether the environmental contamination of the Class Area by chlorinated ethylene solvents including tetrachloroethylene (PCE) constitute a public health risk to the affected population.

I have reviewed the environmental reports, analytical data and information provided to residents and plaintiffs' counsel by Defendant and the Wisconsin Department of Natural Resources. In addition I have reviewed the scientific literature on the chlorinated ethylene organic solvents, primarily PCE, trichloroethylene (TCE) and vinyl chloride (VC). I have myself conducted studies on chlorinated ethylenes in the course of my professional career, stretching over a period of decades prior to being consulted in this case.

Reports indicate that a substantial contamination by chlorinated ethylene solvents of soil, groundwater and soil vapor occurred at the Madison-Kipp Corporation (MKC) facility located at 201 Waubesa Street, beginning decades ago and continuing until at least 1989, resulting from improper management and disposal of chlorinated ethylene solvents; that this contamination found its way into the groundwater, soil, soil vapor and indoor air at homes in the vicinity of the

MKC facility and that this contamination has resulted in exposures through inhalation of chlorinated ethylene solvents (primarily PCE) to residents of these homes.

Data provided to me indicate that the concentrations of the chlorinated ethylene organic solvents in the indoor air to which residents have been, are currently, and in the future could be exposed present an imminent and substantial long term health danger. In particular, I have arrived at the following opinions on the basis of my research, education, training and experience and the facts of the case as obtained in the documents referred to above:

It is my opinion, within a reasonable degree of medical certainty, that exposures to PCE in the residential environment presents a public health risk to Class Area residents[1]. This risk is related to exposures to PCE and its degradation products via inhalation through indoor and ambient air.

It is my opinion, within a reasonable degree of medical certainty, that the weight-of-the-evidence favors the proposition that exposure to PCE in the residential environment of Class Area members presents an increased and unacceptable risk of cancer to those exposed under the usual circumstances of living and working in a contaminated environment such as in Madison, Wisconsin.

[1] The Class Area comprises those portions of Madison, Wisconsin identified in Judge Crabb's class certification order.

II. BACKGROUND AND QUALIFICATIONS TO OFFER OPINIONS RELATED TO PUBLIC HEALTH RISKS OF CHLORINATED ETHYLENE SOLVENT EXPOSURE

I am an epidemiologist, physician, and government-funded researcher specializing in the study of diseases caused by exposure to toxic chemicals and other environmental agents. I served as the Chair of the Department of Environmental Health at the Boston University School of Public Health for 26 years before retiring from that position in 2003. As Chair, I oversaw 14 other professors and was responsible for the research and teaching programs in the field of environmental health for all doctoral students enrolled in our Department's doctoral program and all Masters students. In addition to my responsibilities as Chair of the Department, I taught epidemiology and subjects related to the causes of illness and adverse effects from toxic chemicals to medical, doctoral and Masters candidates virtually every year since 1977, including courses in *Environmental Epidemiology*, *Cancer Toxicology* and *Toxicology and Epidemiology of the Chlorinated Ethylenes*. After retiring from the Chairmanship, I resumed my work as Professor of Public Health and continued to conduct my heavy research agenda, teach my doctoral students, give lectures and carry out the usual administrative and committee responsibilities of a full time senior faculty member. In January of this year, I reduced my work load from full-time to part-time.

In addition to my teaching and supervisory responsibilities, I have been actively engaged in basic scientific research for the past 44 years. I am currently involved as the principal investigator or co-principal investigator of two government-sponsored studies, and have been involved in large and complex environmental epidemiology studies, including one in which I am supervising, with a colleague, an epidemiological research project on the health effects of PCE

exposure that has been ongoing since 1988 funded by the National Institutes of Health and the Commonwealth of Massachusetts.

I am a licensed medical doctor (M.D.) and I am an active practitioner, researcher, and teacher in the field of public health. Public health is a discipline that investigates the health of, and suggests preventive measures and remedial treatments for, populations rather than individuals. I received my MD from Cornell University Medical College in 1967, obtained my post-doctoral Masters Degree in Public Health from The Johns Hopkins University – School of Hygiene and Public Health in 1968 (where I studied epidemiology and epidemiology-related subjects and where in 2001, I was honored to be selected a member of the Johns Hopkins Society of Scholars), received a license to practice medicine as a physician from the Commonwealth of Massachusetts in 1973, held an appointment in the Department of Radiology at the Peter Bent Brigham Hospital (one of the teaching hospitals affiliated with the Harvard Medical School, now part of Brigham and Women's Hospital) from 1971 to 1977, was appointed Medical Director of the Boston Environmental Hazards Center of the Veterans Administration (VA) Medical Center in Boston in 1994 – 1999; and held appointments on the staffs of the Neurology and Medical Services of the Veterans Administration (VA) Medical Center in Boston during that period. In 2001, I was elected a Fellow of the *Collegium Ramazzini*, a select body limited to 160 members of the world's occupational and environmental health experts. At Commencement for the year 2012, I received Boston University School of Public Health's Faculty Career Award for Research and Scholarship. I was the first elected Chair of the School's Faculty Senate and currently represent the School on the University's Faculty Council.

In addition to supervising government-funded projects and serving as an advisor to many federal, state, and international agencies (including the Centers for Disease Control and

Prevention (CDC), US EPA, the National Institutes of Health (NIH), and the World Health Organization (WHO), I have been privileged with membership in eleven professional societies (including a term as the President of the Massachusetts Public Health Association, the public health counterpart to the state's medical society). I have served on several panels of the National Research Council/National Academy of Sciences on water contamination and currently chair the panel, advisory to US EPA, on research in water security.

I am co-Editor in Chief of a peer reviewed scientific journal, *Environmental Health*. I serve or have served as an editor or referee to many peer reviewed journals devoted to scientific- or health-related issues (including *JAMA*, *Science*, and *The New England Journal of Medicine*) and have published over 100 articles, editorials, chapters, or comments, including many in the world's most highly regarded scientific journals. I have been asked to write editorials or commentaries on scientific topics on five separate occasions by *The Lancet*, one of the world's premier medical journals. I also have presented many dozens of papers at national and international meetings.

In the course of my own research, teaching and service work I have become familiar with the scientific issues underlying human exposure to hazardous chemicals as well as the means, methods and acceptable practices used by scientists to draw conclusions. I am co-author of a textbook chapter on environmental hazards in both editions of a major textbook of primary care.[2]

I am co-principal investigator of a large epidemiological study of reproductive, developmental and cancer effects from environmental organic solvent exposures with chlorinated

---

[2] Ozonoff, D., et al., "Health Problems Reported by Residents of a Neighborhood Contaminated by a Hazardous Waste Facility." Amer. J. Ind. Med. 11:581-597, 1987; Ozonoff, D. and Aschengrau, A. "Community exposures to toxic substances," Ch. 27 in Paul, M., editor, Occupational and Environmental Reproductive Hazards, Williams and Wilkins, Baltimore, MD, 1993; Ozonoff, D. "Environmental Health," Ch.9 in Textbook on Internal Medicine and Primary Care, ed. by J.Noble, Little-Brown, Boston, 1987; Pepper, L., Ozonoff, D. "Environmental Health," in Textbook of Internal Medicine and Primary Care, Second Edition, ed. by J.Noble and G. Modest, Mosby, Philadelphia, 1996.

ethylenes, the compounds found in Madison, Wisconsin. I have published numerous scientific publications in the peer-reviewed literature concerning the health risks related to exposure to the chlorinated ethylenes, TCE or PCE.[3] I am overall Program Director of a $10 plus million dollar NIH-funded research center looking into the health effects from chemicals, a center that encompasses epidemiological, toxicological and ecological studies in three institutions and with seven senior principal investigators. I have published and given papers at national and international meetings on the subject of scientific method and the uses of science outside of science.[4]

---

[3] Aschengrau, A., Ozonoff, D. Paulu, C., Coogan, P. Vazina, R., Heeren, T., Zhang, Y., "Cancer risk and tetrachloroethylene (PCE) contaminated drinking water in Massachusetts," Archives of Environmental Health, 48:284-292, 1993; Aschengrau, A. and Ozonoff, D. Upper Cape Cancer Incidence Study. Final Report. Massachusetts Department of Public Health, Boston, January 9, 1992, 700 pp.; Byers V.S., Levin A.S., Ozonoff D.M., Baldwin R.W., "Association between Clinical Symptoms and Lymphocyte Abnormalities in a Population with Chronic Domestic Exposure to Industrial Solvent-contaminated Domestic Water Supply and a High Incidence of Leukaemia," Cancer Immunology and Immunotherapy 27:77-81, 1988; Aschengrau, A., Ozonoff, D., Coogan, P., Vezina, R. Heeren T., Zhang, Y. "Cancer risk and residential proximity to cranberry bog cultivation in Massachusetts," AM J Pub Hlth, 86:1289-1296, 1996; Aschengrau, A, Paulu C, Ozonoff D, "Tetracholoroethylene-contaminated drinking water and the risk of breast cancer," Environ Health Perspect, 106(suppl4):947-953, 1998; Paulu C., Aschengrau A, Ozonoff D, "Tetrachloroethylene-contaminated drinking water in Massachusetts and the risk of colon-rectum, lung, and other cancers," Environmental Health Perspectives, 107:265-271, 1999; Aschengrau A, Rogers S. Ozonoff D, "Perchloroethylene-contaminated drinking water and the risk of breast cancer; additional results from Cape Cod, Massachusetts, USA," Environ Health Perspect. 111:167-74, 2003; Vieira V. Aschengrau A, Ozonoff D. "Impact of tetrachloroethylene-contaminated drinking water on the risk of breast cancer: using a close model to assess exposure in a case-control study," Environ Health. Feb 25;4(1):3, 2005; Spence LA, Aschengrau A, Gallagher LE, Webster TF, Heeren TC, Ozonoff DM, "Evaluation of the Webler-Brown model for estimating tetrachloroethylene exposure from vinyl-lined asbestos-cement pipes," Environ Health. Jun 2;7:24, 2008; Aschengrau A, Weinberg J, Rogers S, Gallagher L, Winter M, Vieira V, Webster T, Ozonoff D, "Prenatal exposure to tetrachloroethylene-contaminated drinking water and the risk of adverse birth outcomes," Environ Health Perspect. 116:814-20, 2008; Gallagher LG, Vieira VM, Ozonoff DM, Webster TF and Aschengrau A, "Risk of breast cancer following exposure to tetrachloroethylene-contaminated drinking water in Cape Cod, Massachustts: reanalysis of a case-control study using a modified exposure assessment," Environmental Health, 10:47 (21 May 2011); Aschengrau, A., Weinberg, J., Janulewicz, P., Romano, M., Gallagher, L., Winter, M., Martin, B., Viera, V., Webster, T., White, R., and Ozonoff, D., "Affinity for risky behaviors following prenatal and early childhood exposure to Tetrachloroethylene (PCE)-contaminated drinking water: a retrospective cohort study," Environmental Health 2011, 10:102; Aschengrau A, Weinberg JM, Janulewicz PA, Romano ME, Gallagher LG, Winter MR, Martin BR, Vieira VM, Webster TF, White RF ,Ozonoff DM, Occurrence of mental illness following prenatal and early childhood exposure to tetrachloroethylene (PCE)-contaminated drinking water: a retrospective cohort study, Environmental Health 2012, 11:2 (20 January 2012).

[4] I have had a longstanding and deep interest in fundamental epistemological questions surrounding explanation in medicine and biology. This began during my undergraduate years where I studied the philosophy of science and mathematics, continued in medical school where my senior thesis examined the logical status of functional explanations in medicine, and in seminars at Johns Hopkins in the history and philosophy of medicine and public

The basic method used in reaching my opinions in this case – the weight-of-the-evidence methodology – is well-accepted by other scientists and constitutes the methodology used by most of them in seeking to determine risks of disease in human beings. The weight-of-the-evidence methodology is not only the same one that I and my BU colleagues use in teaching graduate students in epidemiology and environmental health, the same one I use in my work for government agencies, and the same one that I have used in my research, but is the same methodology used by state, federal, and international agencies (including the International Agency on Cancer Research (IARC), the US EPA, and the National Toxicology Program of the US Public Health Service).

---

health. One of my earliest publications was on the subject of scientific explanation (Ozonoff, D. "An Attack on 'A Defense of Vitalism,'" *J. Theoretical Biology,* 24:l2l, 1969). On my arrival at Boston University School of Public Health I originated the course in The History and Philosophy of Public Health and taught it for several years. My scholarly publications since have frequently visited various aspects of science and scientific explanation as they relate to legal and regulatory matters. See, for some examples, Krimsky, S. and Ozonoff, D. "Recombinant DNA Research: The Scope and Limits of Regulation," *American Journal of Public Health,* 69:l252-l259, 1979; Boden, L., Miyares, J. R., Ozonoff, D., "Science and Persuasion: Environmental Disease in U. S. Courts," *Social Science and Medicine.* 27:1019-1029, 1988; Ozonoff, D. Review of "Asbestos: Medical and Legal Aspects," by Barry Castleman. *Am. J. Ind. Med.* 12:113-115, 1988; Ozonoff, D. "Medical and Legal Causation," in Landrigan P.J. and Selikoff, I.J., eds., *Occupational Health in the 1990s: Developing a Platform for Disease Prevention,* Ann. NY Acad Sci, 572:23-26, 1989; Ozonoff, D. "The Discovery of Occupational Disease by the Workman's Compensation System in the l930s," paper presented at the *Annual Meeting of the American Association for the Advancement of Science*, May 1984; Ozonoff D, "Woburn Hazardous Waste Case," paper presented at the Annual Meeting of the American Public Health Association, New Orleans, October 1987; and, on matters of methodology, Ozonoff, D. and Wartenberg, D. "Toxic Exposures in a Community Setting: The Epidemiological Approach," in Groopman J, and Skipper P, eds., *Molecular Dosimetry and Human Cancer; Analytical, Epidemiological and Social Considerations*, CRC Press, Boca Raton, FL, 1991; Ozonoff, David, "Conceptions and Misconceptions about Human Health Impact Analysis," *Environmental Impact Assessment Review*, 14:499-516, 1994; Ozonoff, D, "A Fish Out of Water: Scientists in Court," Workshop on Scientific Evidence in Court, National Academy of Sciences, September 6, 2000, Washington, DC; Ozonoff, D, "Is a Legal Cause-in-Fact in Fact a Cause?," Paper delivered to the Robert S. Cohen Forum on Science Studies, Boston Colloquium for the Philosophy of Science, Boston, MA, October 16, 2000; Ozonoff D "Science and Justice: Clueless in the Courtroom," College-wide lecture, Dartmouth College, November 9, 2000; Ozonoff D, "Is ethical theory *of practical use in conducting community health studies?", Environmental Epidemiolology and Toxicology 2:67-73, 2000; Ozonoff D,* "Superfund Basic Research Program: A Model for Contemporary Research Programs: Guest Editorial," *Environmental Health Perspectives 111:A140-A141, 2003; Ozonoff D,* "On being careful what we wish for: Some Difficulties with Operationalizing the Precautionary Principle," *European Journal of Oncology, in press; and Int J Occup Med Environ Health. 2004;17(1):35-41. Review; Clapp, R, Ozonoff D,* "Environment and health: vital intersection or contested territory?" *Am J Law Med. 2004;30(2-3):189-215; Ozonoff D,* "Epistemology in the courtroom: a little "knowledge" is a dangerous thing," *Am J Public Health. 2005;95 Suppl 1:S13-5; Ozonoff D,* "Legal causation and responsibility for causing harm," *Am J Public Health. 2005;95 Suppl 1:S35-8;* Boden L, Ozonoff D, "Litigation generated science: why should we care?" Environ Hlth Perspect. Jan;116(1):117-22, 2008

III. THE SCIENTIFIC METHOD

A. Overview of the Sections on Scientific Method (Sections III, IV, V, VI)

While I am not a lawyer or a legal expert, I have made an effort to familiarize myself with the purpose, requirements and use of expert opinion in the legal setting so as to be better able to provide helpful material to the Court; to reduce the burden on both parties to inquire via the discovery process about the nature and basis for my opinions; and to attempt to provide the Court with information to allow an easier decision as to whether additional fact-finding via a Daubert-type hearing is needed. This is the reason for the length and detail of the Report, in which I have taken considerable pains to set out, as clearly as I can (acknowledging that my understanding of these matters is from the perspective of a practicing researcher and scientist rather than a lawyer or jurist) the reasoning and methods I used to arrive at my opinions. Given my current understanding of what is required, I have not only set out my opinions and their bases, but the method and reasoning I used to arrive at them.

Thus, Section (III) deals in very general terms with Scientific Method; Section IV with general questions about Causality; Section V on how opinions about causality are arrived at; and Section VI, how these principles were applied by me to arrive at my opinions.

Section VII contains my opinions on the health risks from chlorinated ethylene solvent exposure to residents of the Class Area and the bases of my opinions.

B. Vigorous Disagreements Among Scientists Are The Rule, Not The Exception

The science that demonstrates that human exposure to chlorinated ethylene solvents like PCE and TCE threaten health is hardly controversial, as I will show with citations to the scientific and regulatory literature. However that does not mean that opposing scientists in an adversary process will not have disagreements.

Laypersons often assume that scientists rarely disagree because science is capable of "objective" confirmation and validation. In fact, however, the history of science shows that disagreements among scientists are ubiquitous and that discord among schools of scientific thought is often both bitter and prolonged[5]. As a panel of the National Academies of Science/National Research Council (NAS/NRC) noted:

> It is disquieting to many nonscientists that scientific experts representing different interests can disagree markedly. There is an implicit assumption that disagreement among scientists should be rare because science is capable of objective, if not always experimental, verification. In fact, however, differences are common in science, although the arguments are spread out over many [different] research papers and long time spans and are usually couched in careful, if not polite, language.[6]

Another NAS/NRC Report expanded upon this observation in discussing the scientific judgments that EPA would need to make in gauging the risks that a drinking water contaminant might present. Noting that EPA was charged, by statute with using the "best available, peer-reviewed science and supporting studies conducted in accordance with sound and objective scientific practices…" this panel convened by the National Research Council further observed that:

> The use of peer-reviewed science will not, however, guarantee agreement among all practices that might be affected by a [judgment that] a chemical … is [on a list of concern] since scientists often weigh the different strands of evidence and supporting data differently …. Disagreements on some [judgments] are to be expected and do not necessarily indicate that they are unsound. Rather, the

---

[5] Kuhn, TS, The Structure of Scientific Revolutions 7 (2d ed. 1970).

[6] NRC, Commission On Geosciences, Environment, And Resources, Setting Priorities For Drinking Water Contaminants 17-18 (National Research Council, Washington, DC, 1999). See also, Jasanoff W, Science at the Bar: Law, Science, and Technology in America Harvard U. Press, 1995; Hacking I, The Social Construction of What?, Harvard University Press, 1999; Faigman D, Legal Alchemy: The Use and Misuse of Science in the Law, WH Freeman 1999; and Pickering A (editor), Science as Practice and Culture, University of Chicago Press, 1992, especially Part I (Positions).

soundness of the judgments will have to be decided on the more-or-less usual way of reasoned and supported argument among the contending and interested parties.[7]

Scientists recognize that disagreements are not only inevitable but a sign of the strength and vitality of a given discipline. Courts are coming round to the same view. In fact, the US Supreme Court recognized this phenomenon in its *Daubert* opinion, stating that if the history of science proves anything to a certainty it is that "there are no certainties in science."

The chief reason disagreements among scientists in the courtroom seem so stark is because time and space are compressed and the nuances of language erased in the adversary process. While disagreements in science are commonplace, these disagreements are over *applications* of scientific reasoning and the judgments and the conclusions drawn from such applications. With the *Daubert* decision the focus has correctly shifted to scientific method and reasoning itself, allowing scientists and jurists to see more clearly the similarities and not just the differences between the requirements of the laboratory and the courtroom.

C. The Task of Scientific Practice and the Law: Finding the Truth

1. Appreciating the Fact that "Facts" are Explanations

Any discussion of scientific method is best done in the context of science as it is actually practiced.[8] Both the scientific method and legal processes attempt to find the truth ("the facts") by means of a defined set of rules and procedures. (Here I am using the word "fact" in the sense

---

[7] NRC, Committee on Drinking Water Contaminants, Water Science and Technology Board, Board on Environmental Studies and Toxicology, Division on Earth and Life Studies, *Classifying Drinking Water Contaminants for Regulatory Consideration*, National Academy Press, Washington, 2001, p. 49.

[8] The Supreme Court concluded that Rule 702 limited expert testimony to opinions that are the product of a scientific thinking process. (See Berger, M., "Evidentiary Framework," Federal Judicial Center Reference Manual on Scientific Evidence, 1994.) As expressed by Judge Kosinski in the Daubert remand, "...we read the Supreme Court as instructing us to determine whether the analysis undergirding the experts' testimony falls within the range of accepted standards governing how scientists conduct their research and reach their conclusions."

used when a jury or other "trier of fact" arrives at a decision about "what is going on" or "what happened," *i.e.,* an *explanation*.)[9] It might be expected, therefore, that there will be great similarities between science and law as they each set out to decide between alternative sets of facts or explanations.

The key questions are: how does a scientist answer scientific questions (one type of which is "what causes what?") and how does a scientist determine whether the answers obtained are scientifically valid facts?

Facts are "decided" by the scientific method, and although that process is not a trial, it too has its rules of evidence and procedure. As in the Law, scientists have subsidiary facts and ultimate facts. In the biological and environmental sciences, the subsidiary facts may consist of empirical results from a scientist's own research, or the laboratory results from environmental testing. Subsidiary facts may also consist of inferences, interpretations, deductions, or theories that are pertinent to the investigation or an individual's problem that can be gathered from the scientist's own experience, gleaned from orally-transmitted experiences of the scientist's mentors and colleagues, or drawn from the scientific literature. The ultimate facts are the conclusions the scientist or physician reaches about her object of study or an individual who has consulted her.

The bridge between subsidiary facts and ultimate facts is scientific reasoning. Scientists select from the universe of subsidiary facts those they judge relevant to the question they are attempting to answer. Scientists make this selection by evaluating work done by other scientists or from their own experience, and by determining whether a given theory, datum, inference or report, is useful and accurate. Thus a scientist asks him- or herself a series of questions,

---

[9] By fixing "facts" as explanations I am glossing over some important distinctions that have concerned legal scholars for many decades. These needn't worry us for this restricted purpose, but become important in other contexts. For an illuminating discussion see Hart HLA and Honoré A, Causation in the Law, Oxford University Press, 1959.

including questions such as: Does he believe this or that theory? What weight does she give to this study or that experiment? How confident is he about the data?

Judgment is always and *necessarily,* exercised in deciding what additional evidence bears on the results of an experiment or diagnosis of a patient, underlining the similarity to using an expert in court. Indeed, disagreements in the scientific literature are often expressed in the different ways scientists construct supporting arguments from the same body of data or the same literature. The phenomenon of "dueling experts," so characteristic of a trial, is also present in science, but in a less visible form, distributed across separate publications in the literature.[10]

2. The Mechanisms that Scientists Use to Ensure that Scientists Are "Honest" – and to Ensure That Science Remains Credible and Useful

Scientists have considerable leeway – but not complete freedom – in selecting useful supporting results from scientific research results and scientific literature. Likewise, there are limits to what an expert can say in court. It is not true that *anything* goes. The *Daubert* decision instructs lower federal courts to ensure that the testimony experts give is based on sound scientific reasoning, which is to say the same type of reasoning that scientists use when they are working in the real world of science. That is, scientists who testify in court are required to use the same sort of thinking, the same sort of standards, the same sort of "rigor," and the same types of data and knowledge and inferences they use in normal scientific practice. As such, the *Daubert* decision raises the question of what keeps things within accepted bounds in normal scientific practice. Although there are codified rules of logic and procedure, there is also

---

[10] It would be convenient if scientists could bifurcate the scientific process into a purely empirical operation followed by an interpretive one. This is a subject much discussed by philosophers of science, who are struck by the "theory laden-ness" of scientific language (see Hanson N, Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science, Cambridge Univ. Press, Cambridge, 1961 and the works by Jasanoff, Hacking and Pickering cited above). This analytical problem is not confined to the expert "fact" witness, but applies also to a lay fact witness who must explain how his account is correct in the face of another lay fact witness who testified to a different perception of what they saw.

considerable latitude in the workings that Thomas Kuhn, one of the late 20th century's most influential philosophers of science, called "normal science." Indeed without this latitude science would not be able to adapt and progress.

There are three self-regulating mechanisms that operate formally or informally to keep scientific practice within accepted bounds – and keep scientists "honest." I will refer to them as (a) critical thinking (a habit of thought), (b) peer pressure (a product of social relations), and (c) peer review (a formal or semi-formal procedure).

a. Critical thinking

Scientists are taught by precept and example to take a "hard look" at their own work and the work of colleagues. In this respect, scientists resemble judges more than attorneys. An attorney's job is to be a "zealous advocate" for his clients, *i.e.,* to take and espouse a deliberately one-sided view of the world. Good trial attorneys, however, will also try to look at things from their adversary's point of view, to anticipate difficulties and problems in their own case, and to evaluate how trial judges and appellate courts will view the claims he makes, the evidence he presents, and the argument he advances. This is precisely the kind of self-critical thinking good scientists routinely engage in with respect to their own and other's science.

Like all virtues, critical thinking can be carried to excess and become an inappropriate and even paralyzing skepticism. Put more simply, "the perfect is the enemy of the good." Mervyn Susser, a well-known scholar of epidemiology and the former editor of the *American Journal of Public Health*, gives some interesting and informative examples and counter-examples of this phenomenon:

> "Undue skepticism... can be as dangerous as credulity to scientific progress and the improvement of health. Only judgment can prevent the hypercritical rejection of useful results. Rigorous and well-founded criticism of the work of Skodak and Skeels (on the mental

> performance of children removed from their families) deferred general recognition of the substantial effect of social milieu on intellectual development. Karl Pearson's[11] valid 1906 critique of Almroth Wright's data on the effectiveness of his typhoid vaccine did not deter the British Army from using it; had they been deterred, disaster might well have followed for the British and French forces in the trench warfare of World War I...The large-scale trials of the Salk polio vaccine in the 1950s also were strongly criticized. Here too, had not the criticism been counter-weighted by the substance of the results, the virtual eradication of poliomyelitis would have been delayed for perhaps five years, at the certain cost of crippling and fatal attacks....
>
> "For more than a decade, the [cigarette] industry was able to ward off its public health enemies. It could draw for its defense on such redoubtable critics of studies of the effects of smoking as Sir Ronald Fisher and Joseph Berkson [two of the most famous statisticians of the twentieth century]. Few now doubt that their strictures were ill-judged."[12]

This habit of skepticism and relentless critique has become such a reflex with epidemiologists, that the late Marvin Schneiderman, former chief statistician of the National Cancer Institute, once half-jokingly defined "epidemiology" as "the practice of criticizing other epidemiologists." The arena of the adversary process, in particular, tends to convert a normal habit of critical thinking into a relentless skepticism. Attorneys have a tendency to view facts or opinions that do not help their case as deliberate attempts to deceive or defeat them, and, indeed, can magnify an appropriate questioning to a cynical skepticism. [13]

b. Peer pressure

Scientists spend many years learning their discipline, and today a great deal of scientific research simply cannot be done without the funding of government agencies, the support of large

---

[11] Karl Pearson (1857-1936) was one of the founders of modern statistics.

[12] Susser, Mervyn, Causal Thinking in the Health Sciences: Concepts and Strategies of Epidemiology, Oxford Univ. Press, New York 1973, pp 141-2. See Cerf, C. and Navasky, V., The Experts Speak: The Definitive Compendium Of Authoritative Misinformation (2d ed. 1998) (providing numerous examples of faulty assessments and predictions by great scientists, inventors, and other experts).

[13] Judge Kosinski, in the Daubert remand, quotes Judge Frank Johnson approvingly: "the examination of a scientific study by a cadre of lawyers is not the same as its examination by others trained in the field of science or medicine."

institutions, and the collaboration of colleagues. A reputation for doing "bad science" is hardly a ticket to obtaining grants, to having one's research published, or to winning tenure or academic promotion – in short, to having a successful and remunerative career. An NIH (National Institutes of Health) grant review, for example, explicitly takes into account the experience and reputation of the proposing scientist as perceived by the reviewer, usually a colleague in the same field. Thus evidence that a scientist has regularly received government research funding, has been promoted to high academic rank, has been asked to serve on advisory bodies, and performs peer review duties for scientific publications is some evidence that his or her work is not regarded as "outside the bounds" by the community of scientists.

Peer pressure is therefore a powerful deterrent to legitimate scientists not to violate the scientific method. Conversely, peer pressure is an incentive for professional behavior, an incentive that affects some more than others, operating most directly for people whose careers are devoted academic research. But it is also true that the majority of independent consultants and applied scientists value their reputations as much as most lawyers and jurists.

c. Peer review

Formal review and evaluation by scientific colleagues of research results or proposed research is widely practiced by funding agencies, independent commissions, and scientific publications. At its best, it provides the helpful "other pair of eyes" that can catch lapses in logic or method that escape those too close to their own work. Although peer review neither guarantees that meritorious work will be reported nor that work that does not meet scientific standards will not, it constitutes the most visible and formal mechanism designed to ensure that the scientific method is followed.

Like peer pressure, peer review is not foolproof. The more "cutting edge" the work the more likely it is to be rejected.[14] At the same time, it is common knowledge that a good deal of substandard or uninformative work finds its way into the peer reviewed literature. Nor is everything in so-called peer-reviewed journals actually peer reviewed, or everything not in this form published without peer review. For example, editorials, invited papers, and papers by editors may or may not be peer reviewed, even in a journal whose general policy is to peer review unsolicited articles. On the other hand, many book chapters, state and federal reports or articles in stand-alone publications are extensively reviewed before publication. In the case of public agency reports, particularly, the more controversial the topic, the more likely extensive peer review will be undertaken, whether or not it appears in a "peer reviewed" publication.

None of the three forces that are employed to keep scientific practice within the bounds of what the science-of-the-day deems acceptable are formalized, except for peer-review. Even peer review is practiced inconsistently, to varying extents, and with little means to evaluate its efficacy. Science has no magic method to draw boundaries around what is acceptable and what is not, even on its own terms.[15]

As with most things, there will be instances which are clearly "beyond the pale" and others which are clearly mainstream science. But the "gray area" in between is substantial and contains a proportion of propositions that are truly scientific advances and also those that will be judged with time as lacking in scientific foundation.

---

[14] Gans, J, Shepherd B, "How are the mighty fallen: rejected classic articles by leading economists," J Econ Perspectives, 8:165-179, 1984

[15] This is called the Demarcation Problem in the philosophy of science and it remains without an agreed upon solution.

IV. THE METHOD OF INQUIRING INTO THE CAUSES OF ENVIRONMENTAL DISEASES

A. The Experimental and Observational Sciences Are Used to Provide Data to Make Judgments About Causality

Can chlorinated ethylene solvents cause cancer in human beings? Scientific practice is taken up with more than exploring questions of causation, but this is a central question in many legal cases. Many epidemiologists now eschew the language of "cause," preferring the more neutral sounding "risk factor." Risk factors are elements that affect the chances (technically, the probability) of developing a disease. In this instance, if exposure to chlorinated ethylene solvents increases the risk of cancer, it would be classified as a risk factor by epidemiologists. In this Report I will use the more commonly understood term, cause.[16]

What does "A causes B" mean to a scientist? Although much ink has been spilled in discussing philosophical aspects of scientific causality, most scientists have adopted a pragmatic approach whose formal articulation goes back at least to John Stuart Mill's famous "Method of Difference."[17] Briefly, Mill's Method holds that A causes B if, all else being held constant, a change in A is accompanied by a subsequent change in B.[18] The formal method to detect such an occurrence is an Experiment, whereby:

---

[16] The relationship between statistical "causes," "causal necessity," and "causal regularities," is the subject of much discussion amongst philosophers of science and among some scientists (especially physicists). Those issues, although of importance, are not relevant to this discussion.

[17]. Actually Susser (op.cit.) discusses four of Mill's "Canons", including the Methods of Agreement, Residues, and Concomitant Variation. One can interpret the others as variations on the Method of Difference, however, so I have elected not to elaborate on them. HLA Hart and Anthony Honoré discuss Mill at some length in their 440 page book on Causation in the Law, noting that there are several areas where the legal context requires a departure from Mills' views. Those contexts do not affect this instance, and, in any event, I am here discussing the spontaneous concepts of most scientists.

[18] This of course does not mean that *nothing else* can produce a change in B.

- all things are held constant except A and B,[19]
- A is varied, and
- B observed.

Not all sciences can utilize a strictly experimental method, however. Some sciences must content themselves with making observations of the real world and deducing scientific fact by applying reasoning and principles from experimental sciences or logic and mathematics. Astronomy is such a science. So is geology. And so also is epidemiology. Astronomers cannot manipulate distant stars and planets experimentally, but they can apply the methods and results of terrestrial physics along with mathematical theories like quantum mechanics or relativity theory to make inferences about the interiors of stars or the structure of other galaxies. Another observational science is geology. In one of its subdisciplines, seismology, scientists *observe* earthquakes; they certainly do not stage city-sized experiments on the factors which cause earthquakes. The inability to conduct full-scale experiments does not connote the inability to do good science, nor that the science involved is inherently "error prone" or less reliable.

Scientists extrapolate from laboratory scale experiments to make scientifically defensible statements about the origins of a "black hole" in space or the causes of earthquakes on our planet. There may be disagreement among experts as to the aptness of a *particular* extrapolation or inference, but generally there is no disagreement that the process of applying events or principles observed on the scale of the laboratory bench to events occurring on the scale of a geographic

---

[19] One of the criticisms of some modern philosophers is that "all other things" can never be kept equal. However, the belief in their ability to select those ancillary things that "count" and those that do not is one of the bedrocks of the usual practice of scientists. The difficulty that this presents for an adequate account of the meaning of causation is serious, however, because we are usually not aware of all the differences, nor do we always know which ones count and which ones don't.

region is scientifically defensible, and indeed something similar is the norm in virtually all observational sciences.

In the biological sciences, in general, and in the public health field, in particular, inferences for one group of humans are regularly drawn from epidemiological studies from another group of humans. Inferences about humans are also made on the basis of observations on animals or test-tube experimentation. Indeed, the scientific reasonableness of drawing inferences from animals to humans provides the principal justification for the decision of the NIH to devote hundreds of millions of dollars to animal research.

Thus, while any *particular* inference may be arguable (and certainly may be the basis of a dispute between the parties in a lawsuit), the *method and reasoning* are not subject to debate, and it is method and reasoning which are the subject of the *Daubert* opinion.[20]

In general there are three sources of information on the effects of toxic exposures in human beings: (a) case reports, (b) toxicological research (including both animal studies and chemical/structural research), and (c) epidemiological studies.

1. The use of case reports regarding the effects of toxic exposures in human beings

A case report, *i.e.,* a report in the medical or scientific literature of a single case or series of cases, is one of the most important sources of information scientists have on effects of toxic substances, and often the only source of information. Reports of cases of accidental poisonings or suicides provide information, such as detailed clinical observations or autopsy data not

---

[20] This is clearly consistent with judicial practice. For example, Judge Becker wrote in Paoli II, page 117: "While it may be true that defendant can offer tests and experiments that do not support the findings of plaintiffs' experts, the defendant cannot deny that animal studies are routinely relied upon by the scientific community in assessing the carcinogenic effects of chemicals on humans. Even defendant's own expert acknowledges that animal experiment studies are built on 'prudent presumptions,' although he concludes that they should not be admitted." See also, Henefin M, Goldstein B, "Reference Guide on Toxicology," In: Reference Manual on Scientific Evidence, Federal Judicial Center: "the responses of laboratory animals are useful predictors of toxic responses in humans" is a central tenet of toxicology. (p. 185).

obtainable by any other route. Moreover they constitute important and obvious "natural experiments," experiments where the relationship between the exposure and effect is usually clear. The use of case reports in medicine is longstanding and important, as evidenced by the continued appearance of such reports in the literature.[21] Indeed the *logic* of a case report is similar to that of a more formal epidemiological study.

Recently Thun and Sinks [22] reviewed the role of cancer clusters in discovering chemicals that caused cancer in humans:

> There are well-known instances in which the investigation of an unusual cancer cluster has led to the identification of a previously unrecognized human carcinogen. All of the examples listed in Table 1 involved clusters of a rare type of cancer in people with prolonged, high-intensity exposure to industrial or medical carcinogens. Each was recognized as extraordinary by an alert clinician and reported to public health and medical officials for evaluation. Although such examples are rare, even in occupational settings, they illustrate how some cancer clusters can provide new scientific information about the causes and prevention of cancers. . . . These women were exposed to ionizing radiation from radium present in the luminous paint when they used their lips to form a sharp tip on the paintbrush. Other clusters involved pleural mesothelioma among asbestos workers in London and angiosarcoma of the liver among chemical workers exposed to vinyl chloride monomer.[23]

It is of note that vinyl chloride is considered one of the classic examples of a cancer causing agent discovered through a case description by an alert clinician.

2. The use of toxicological research reports to understand the effects of toxic exposures in human beings

Toxicological research (including both animal studies and chemical/structural correlations), along with epidemiology, is one of the two other sources of information which provides much of the basis for scientific judgments relating toxic exposures to health effects.

---

[21] The Lancet, for example, one of the world's leading medical journals, contains a Case Report every week, as does the New England Journal of Medicine.

[22] Thun M, Sinks T, "Understanding cancer clusters," CA Cancer J Clin 2004; 54:273-280

[23] Creech JL Jr, Johnson MN. Angiosarcoma of liver in the manufacture of polyvinyl chloride. J Occup Med 1974; 16: 150–151

Indeed the first evidence, using animals, that vinyl chloride could cause cancer predates the appearance in human cases referred to above by three years.[24] The animal data for vinyl chloride is not only qualitatively informative but also quantitatively.[25]

Toxicology is an *experimental* science, while epidemiology is an *observational* science.[26] The advantages of being able to conduct an experiment is obvious, because Mill's Method of Difference depends upon observing the result on B of a change in A, *other factors held constant.* The essence of an Experiment is the control of all factors, except for A and B. This kind of control allows the scientist to ask quite precise questions about explicitly defined A's and B's, and get relatively unambiguous answers[27].

3. The use of epidemiological studies regarding the effects of toxic exposures in human beings

Epidemiological studies are observations of "natural experiments" that are occurring in the real world. The idea is to find situations which are almost like laboratory experiments, observe them, obtain as much information as possible from them, and then interpret the results. At the heart of a natural experiment in epidemiology is almost always a comparison between groups, for example, a group exposed to a chemical and one not exposed. The ideal situation

---

[24] Viola PL, Bigotti A, Caputo A, "Oncogenic response of rat skin, lungs and bones to vinyl chloride, Cancer Res 3:516-522, 1971; See also, Melnic k RL, "Carcinogenicity and mechanistic insights on the behavior of epoxides and epoxide-forming chemicals," Ann N Y Acad Sci. 982:177-89,2002

[25] Sanner T, Dybing E, "Comparison of carcinogen hazard characterisation based on animal studies and epidemiology," Basic Clin Pharmacol Toxicol. 96:66-70, 2005

[26] I speak here of research toxicology and epidemiology. The field of clinical toxicology, by contrast, is an observational science, taking as its subject the diagnosis and treatment of individuals; and clinical epidemiology is often experimental, involving randomized clinical trials. This semantic inversion when each is qualified with the word "clinical" presents no conceptual difficulties.

[27] Whether complete control is practically possible varies, of course, but the principle should be clear. To the extent the answers are ambiguous, another experiment can be designed to resolve the ambiguity.

would be to have the groups in the real world the same in all relevant respects (*i.e.*, comparable) except for the variable under study (*e.g.*, exposure to chlorinated ethylenes).

Unfortunately such natural groupings are rarely comparable, and special techniques must be used to account for known differences. However, not all sources of non-comparability are known.[28] Providing they are not a necessary accompaniment of the variable being investigated, these residual factors fall by chance in the two groups being compared. The result is that there are usually differences solely attributable to the random way these factors are distributed between groups in the particular study. The "chance" fluctuations in apparently otherwise similar populations require an epidemiologist to use statistical tools to evaluate the role of "noise" that might be obscuring an underlying "signal."

Observing some unintended or "natural" experiment in the real world, which is the essence of observational sciences like epidemiology, has the enormous advantage that it involves human beings living under conditions similar to ones found by complainants in a civil suit. Evidence from epidemiology has confirmed the evidence from animal studies and case reports that chlorinated ethylenes are human carcinogens.[29]

Toxicological experiments and epidemiological studies each have characteristic strengths and weaknesses.[30] In view of the fact that different scientific disciplines have disparate strengths and weaknesses, and given the propensity of scientists to disagree, the key question for courts is

---

[28] This is a deterministic view of disease causation. One could also take a probabilistic view, in which case scientists would have to discuss sample error from some assumed super-population of identical study settings. This alternative view does not affect any of the points made.

[29] Boffetta P, Matisane L, Mundt KA, Dell LD, "Meta-analysis of studies of occupational exposure to vinyl chloride in relation to cancer mortality," Scand J Work Environ Health. 29:220-9, 2003

[30] Ozonoff, David, "Conceptions and Misconceptions about Human Health Impact Analysis," Environmental Impact Assessment Review, 14:499-516, 1994.

to understand how scientists decide which studies, data, experiments and articles to use and rely on and for what purposes, *i.e.,* how do they *interpret and apply* the results of scientific studies?

B. Interpreting Scientific Studies

It is well known that when different scientists interpret the same studies they do not always reach the same conclusion. How and why do scientists interpret the "same" basic facts, the same set of numbers, the same research report, in different ways? Two aspects of scientific interpretation are relevant to this discussion. In the literature of scientific methodologies they are commonly referred to as internal and external validity.

Internal validity refers to a judgment about the extent to which the experiment or study produces valid information on its own terms, *i.e.,* the extent to which it is *internally* valid. Thus, for internal validity the crucial question to be answered is *not*, "If TCE causes birth defects in Wistar rats does it also do so in humans?" but *rather* "Did the experiment validly show that TCE causes birth defects in Wistar rats?"

External validity, on the other hand, does refer to a judgment about the extent to which the internally valid results of an experiment or study can be generalized to other situations, and to which ones. Thus, for external validity the crucial question to be answered is not "did the experimental evidence adequately demonstrate that TCE causes birth defects in Wistar rats?" but rather "If TCE *does* cause birth defects in Wistar rats, does it *also* do so in humans?"

1. Internal validity: How good is the study in its own terms?

At the heart of a case report, a toxicological experiment, or an epidemiological study lies a comparison.

- Case reports are usually made to call the attention of the medical community to an "interesting" observation (compared to "the usual"), such as a rare disease in the context of an unusual exposure (e.g., an unusual cancer in a worker exposed to PCE). The comparison is with previous or usual experience.
- In an experiment, the comparison is between the different states of B, when A is varied.
- In an epidemiological study, it is the analogous comparison in the "natural" or unintended experiment that is being observed.

Once an event is observed or an experimental comparison is obtained it remains to explain or interpret the observation or result in the expected or compared entities, whether a difference or a lack of a difference.

Take as an example a study comparing the health outcome of two distinct groups of human beings, one group comprised of those workers in a factory who were exposed to a chemical used in the production process, and the other group consisting of all members of the general population, most (but perhaps not all) of whom were not exposed to the chemical. Suppose the workers have more disease than the general population (the following analysis works just as well in the case of no increased disease). There are three generic reasons such a difference (or lack of difference) might be observed, referred to as "bias," "chance," and real effect. They are conceptually independent, but not mutually exclusive forces, *i.e.,* all, some, or none can operate simultaneously. Each must be evaluated to extract a valid message ("the real picture" or the "true signal") from the study.

a. Evaluating the Role of "Bias" in Evaluating the Internal Validity of a Research Study:

Another term for "bias" is "systematic error." This differs somewhat from the common usage of the word (lack of objectivity, independence, and/or neutrality), and in epidemiology the word has been refined and qualified to encompass a wide variety of sources of systematic error, each given a name.[31] For example, epidemiologists talk of various kinds of "information bias" ("recall bias," "observation bias," "differential or non-differential misclassification bias," etc.) or types of "selection bias," "confounding bias," etc. All biases have as their underlying mechanisms factors which make the compared groups different in more ways than just the variable being studied. Because the object of an experiment or study is to isolate one element (exposure to the chemical in my example), one must estimate the effect of the uncontrolled differences on the comparison.

A common source of potential bias in an epidemiological study is "confounding," and I illustrate bias with this example. Suppose scientists were comparing cancer rates in two groups. As in all epidemiological studies, this comparison is of the nature of an experiment, but one that is "handed to us" by nature (*i.e.,* circumstance), not one of our own devising. Thus scientists are unable to control everything they might like to control in this comparison.[32] It might be, for example, that the workers in this hypothetical instance are considerably younger than the general population, and because cancer risks rise with age, they would be expected to have less cancer than the comparison group, all other things being equal. If this difference were not somehow accounted for, a comparison between the two groups would be misleading. The same non-

---

[31] See Bailey LA, Gordis L, Green M, "Reference guide on epidemiology," Reference Manual on Scientific Evidence, Federal Judicial Center, 1994, pp 121-180 , especially p. 132.

[32] Epidemiologists depend on Nature to be their "research assistants" and Nature is not usually very tidy or cooperative. Thus it is normal and natural for there to be "loose ends" sticking out of epidemiological studies, loose ends that other epidemiologists usually cannot resist pulling. As with most things, designing an informative study is difficult. Criticizing one is easy.

comparability could influence a comparison in the opposite way if the workers were on average older than the general population (say, if they were a group of retirees).

The most important means of coping with bias is to recognize it. An important part of the training and practice of an epidemiologist is to recognize and account for the effects of the inevitable non-comparability found in observational studies. Once recognized, an epidemiologist can often gauge the impact of a source of bias on the results and adjust interpretations accordingly. Sometimes the data themselves can be "adjusted" ("controlled") to eliminate the non-comparability in the two groups for certain factors like age or sex.

b. Understanding the Meaning of "Statistical Significance"

Not all sources of non-comparability are known. Providing they are not a necessary accompaniment of the variable being investigated, these residual factors are distributed by chance between the two groups being compared. The result is that there are usually differences solely attributable to the random way these factors are distributed between groups in the particular study. The "chance" of fluctuations in apparently otherwise similar populations requires an epidemiologist to use special tools to discern the true meaning from the chaos of disparate data – to "see" the true picture amidst a welter of images, or to "hear" the true, underlying "signal" in the midst of the noise produced by these variations. The mathematical tools used for these purposes involve statistical analysis.

The main use for statistics in epidemiology, for the purposes of this discussion, is to evaluate the role these random allocations of other factors ("chance") might have played in the results.[33] Statistical methods *do not* prove that chance is the source of a difference (or lack of

[33] Statistical methods are used for other purposes, as well, of course, such as modeling or estimation, but for our purposes the role in evaluating chance is most relevant.

difference). These methods only provide information on how likely it is that chance *could* have played a part *if there were no bias and no true effec*t. The meaning of "statistically significant" is that the likelihood that chance *could* have produced the observed results *if there were no bias and no real effect* is less than some arbitrarily predetermined level, such as 5% ("p<.05").[34]

For these reasons and other reasons, it is absolutely false – and, indeed, a serious interpretive error – to assert that a result that is not "*statistically significant*" means the results must be due to chance. And for these reasons, prominent epidemiologists eschew "statistical significance," believing that it is not a *sine qua non* of "good science" and maintaining that it is "neither necessary nor appropriate as a requirement for drawing inferences from epidemiologic data."[35]

These views are hardly mine alone, nor are they new. Instead, they are representative of the views of many others, including Sir Austin Bradford Hill, one of the 20th century's most influential statisticians, and some of the most highly regarded epidemiologists in this country, such as Dr. Kenneth Rothman (the co-author of the most sophisticated current textbook on epidemiology and founding Editor-in-Chief of the journal, EPIDEMIOLOGY) as well as other epidemiologists.[36]

---

[34] The original source of the 5% criterion is lost in time. It apparently came from the original applications of statistical methods to agricultural experiments and expressed a cost-benefit statement about the expense of redoing a large trial involving a whole growing season and plots of various seeds and fertilizers. Its use for public health purposes might thus be questioned. It is interesting to note that in other sciences, notably, physics, another common criterion for "statistical significance" is not 5% but 10%. In any event, virtually every elementary statistics text warns the student of the highly arbitrary nature of the figure.

[35] Amicus Brief filed in the US Supreme Court in Daubert vs. Merrell Dow by Dr. Kenneth Rothman, et al.

[36] In their amicus brief in Daubert, Profs. Rothman and Weiss alerted the US Supreme Court to the fact "that there is a large community of respected epidemiologists that rejects the single-minded focus on significance testing that is expressly mandated by the trial court and implicitly adopted by the appellate court, without the benefit of expert testimony or analysis memorialized in its opinion. As an example, Epidemiology [an epidemiology journal of which

Hill, for example, put it this way in 1965:

> "No formal tests of significance can answer those questions. ("Is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?") Such tests can, and should, remind us of the effects that the play of chance can create, and they will instruct us in the likely magnitude of those effects. Beyond that they contribute nothing to the 'proof' of our hypothesis." … "I wonder whether the pendulum, has not swung too far -- not only with the attentive pupils, but with the statisticians themselves. … Fortunately I believe we have not yet gone so far as our friends in the USA where, I am told, some editors of journals will return an article because tests of significance have not been applied …."[37]

Similarly, in an amicus brief to the US Supreme Court in the *Daubert* case, Professors Rothman and others, stated: "Significance testing, however, is neither necessary nor appropriate as a requirement for drawing inferences from epidemiologic data." [38] The amicus brief continued:

> "The notion that only when data demonstrate "statistical significance" do epidemiologists draw inferences about observed associations between suspected risk factors and medical conditions is mistaken. Significance testing is nothing more than a statistical technique that attempts to evaluate what is called "chance" as a possible explanation for a set of observations, and classify the observations "significant" or "not significant" based on the likelihood of observing them if there were no relationship between the suspected cause and effect. Testing for significance, however, is often mistaken for a sine qua non of scientific inference. . . . Scientific inference is the practice of evaluating theories. As such, it is a thoughtful process, requiring thoughtful

---

amicus Dr. Rothman was then editor] discourages publication of articles that rely on significance testing." Amicus Brief filed in the US Supreme Court in Daubert vs. Merrell Dow by Dr. Kenneth Rothman, et al.

The views of other prominent epidemiologists, toxicologists, biostatisticians, and other scientists who believe that statistical significance is a problematic test of scientific validity are discussed in Joseph L. Fleiss, Significance Tests Have a Role in Epidemiologic Research: Reactions to A.M. Walker, 76 Am. J. Pub. Health 559, 559-60 (1986); Steven N. Goodman & Richard Royall, Evidence and Scientific Research, 78 Am. J. Pub. Health 1568, 1568-74 (1988); Charles Poole, Beyond the Confidence Interval, 77 Am. J. Pub. Health 195, 195-99 (1987); W. Douglas Thompson, Statistical Criteria in the Interpretation of Epidemiologic Data, 77 Am. J. Pub. Health 191, 191-94 (1987); Alexander M. Walker, Reporting the Results of Epidemiologic Studies, 76 Am. J. Pub. Health 556, 556-58 (1986).

[37] Austin Bradford Hill, The Environment and Disease - Association or Causation?, Proceedings of the Royal Society of Medicine (1965) 58:296 at p. 299.

[38] Rothman et al., "Summary of Argument" section of their amicus brief in Daubert.

> evaluations of possible explanations for what is being observed. Significance testing, on the other hand, is merely a statistical tool that is frequently, but inappropriately, utilized in the process of developing inferences."

One amicus has observed:

> "With the focus on statistical significance, if chance seems to be a plausible explanation, then other theories are too readily discarded, regardless of how tenable they may be. As a result, effective new treatments have often been overlooked because their effects were judged to be "not significant," despite an indication of efficacy in the data. Conversely, if "significance" seekers find that the results of a study are calculated as improbable on the basis of chance, then chance is often rejected as an explanation when alternative explanations are even less tenable."[39]

The outcomes of statistical tests are strongly influenced by the size of the study population. For small populations, very large observed differences, of substantial *public health* significance, may still not be *statistically* significant.[40] That is to say, a large effect that a scientist would take seriously from the public health point of view cannot be differentiated on its

---

[39] Rothman's amicus brief in Daubert, citing K. Rothman, Significance Testing, 105 Annals of Internal Medicine 445, 445-46 (1986) (citations omitted). According to the Rothman Weiss amicus brief:

> A better approach to evaluating the error in scientific measurement is the use of "confidence intervals." A confidence interval is a range of possible values for a parameter that is consistent with the observed data within specified limits. The process of calculating a confidence interval within the chosen limits is know as "interval estimation." See K. Rothman, Significance Testing at 119.
>
> An important advantage of interval estimation is that it: "do[es] not require irrelevant null hypothesis to be set up nor [does it] force a decision about 'significance' to be made -- the estimates can be presented and evaluated by statistical and other criteria, by the researcher or the reader. In addition the estimates of one investigation can be compared with others. While it is often the case that different measurements or methods of investigation or theoretical approaches lead to 'different' results, this is not a disadvantage; these differences reflect important theoretical differences about the meaning of the research and the conclusions to be drawn from it. And it is precisely those differences which are obscured by simply reporting the significance level of the results. "

Rothman/Weiss amicus brief in Daubert, quoting L. Atkins and D. Jarrett, The Significance of "Significance Tests," in J. Irvine and I. Miles (eds.) Demystifying Social Statistics (1979).

[40] A detailed example showing how results can be of public health significance but not statistical significance can be found in Ozonoff, David, "Conceptions and Misconceptions about Human Health Impact Analysis," Environmental Impact Assessment Review, 14:499-516, 1994.

face from chance. Either chance or a real causal influence (or bias) could be responsible for the worrisome effect. Conversely, in large populations, very slight and substantively meaningless differences can be "statistically significant."[41]

Moreover, statistical methods are sometimes mistakenly viewed as standard, agreed-upon, and mechanical procedures. Scientists even allow computers to do them, seemingly without human intervention. But as any statistician knows, there is a great deal of judgment in deciding which tests to use in which circumstances, which tests are valid in those circumstances, and what they do and do not mean. Less well recognized is that statistics itself, like all active disciplines, is a field in ferment and change. Thus not all statisticians will agree on the propriety of even commonly used tests.[42] In his book, *Scientific Inference*, Michael Oakes has written:

> "It is a common complaint of the scientist that his subject is in a state of crisis, but it is comparatively rare to find an appreciation of the fact that the discipline of statistics is similarly strife-torn. The typical reader of statistics textbooks could be forgiven for thinking that the logic and role of statistical inference are unproblematic and that the acquisition of suitable significance-testing recipes is all that is required of him."[43]

---

[41] For example, a difference of 1/8" in height between east coast children and west coast children will be statistically significant if very large numbers of children on both coasts are measured.

[42] A good example is the Fisher Exact Test, commonly used for small tables frequently encountered in environmental epidemiology. Certain well known statistical programs even force the user to employ this test if several table cells contain expected values of less than five, even though it has been known for years that the test is inappropriate. See D'Agostino R, Chase W, Belanger A, "The appropriateness of some common procedures for testing the equality of two independent binomial populations," Am Statistician 42:198-202, 1988, and references therein.

[43] Oakes, Michael, Statistical Inference, Epidemiological Resources Inc., Chestnut Hill, MA, 1990. Oakes then goes on to quote a review (by Dusoir) of a statistics text in a technical journal:

> "A more fundamental criticism is that the book, as almost all other elementary statistics texts, presents statistics as if it were a body of coherent technical knowledge, like the principles of oscilloscope operation. In fact statistics is a collection of warring factions, with deep disagreements over fundamentals, and it seems dishonest not to point this out."

When used, statistical methods are meant to help scientists *evaluate* the possible role of chance.[44] Scientists must evaluate the possibility of a concurrent *real* effect separately.

c. *Understanding the Relative Role of a Real Effect (or the Absence of a Real Effect) in Evaluating the Internal Validity of a Research Study*

The most important reason for a difference between two groups, however, is an actual effect or influence of the variable being studied (exposure to a chemical at work in my example), *i.e.,* that "A *does* cause B." As discussed in greater detail below, scientists recognize that "causation" should *not* be regarded as an experimental or epidemiological result, but rather as a judgment made about the experimental or epidemiological data. [45]

It is apparently not always appreciated that causation inherently involves a judgment.[46] There is a tendency to believe that somehow "causation" is not a subjective judgment or interpretation but an actual, real, objective, discoverable, and measurable property of a relationship that can be demonstrated empirically, as if some associations had readable labels on

---

[44] As expressed by the epidemiologist Kenneth Rothman in his Daubert amicus brief:

> "The result of using significance testing as a criterion for decision making is that the focus is changed from the information presented by the observations themselves to conjecture about the role chance could have played in bringing about those observations." [emphasis in original]. Quoted by Berger M, cited above (op. cit.). Rothman is the author of a standard text, Modern Epidemiology and Editor in Chief of the journal Epidemiology.

[45] As professors Rothman and Greenland explain, at p. 22 of their textbook:

> "Perhaps the most important common thread that emerges from the debated philosophies [of scientific causation] is Hume's legacy that proof is impossible in empiric science. This simple fact is especially important to epidemiologists, who often face the criticism that proof is impossible in epidemiology, with the implication that it is possible in other scientific disciplines. Such criticism may stem from a view that experiments are the definitive source of scientific knowledge. Such a view is mistaken…Even the most careful and detailed mechanistic dissection of individual events cannot provide more than associations… ."

[46] See Federal Judicial Center Reference Manual on Scientific Evidence (1994) at p. 157 ("causation is a judgment issue for epidemiologists and others interpreting the epidemiological data."). See also the extended discussion of this point in K. Rothman & S. Greenland, Causation and Causal Inference," in: K. Rothman and S. Greenland, Modern Epidemiology (Second ed. 1996) at pp. 7-28.

them that said 'causal' and all that scientists need is the right instrument to read the label.[47] Although some scientists may be loathe to admit it, and although many lawyers and judges may not believe it, there is simply no magic formula or easy checklist for making scientific judgments.[48]

I say this even though for example in the case of the chlorinated ethylene, there is little judgmental variation between experts as to their carcinogenicity in humans. The scientific consensus on VC's carcinogenicity is still a consensus of judgment.

d. Summary of the Means by Which the Internal Validity of Studies Is Evaluated

Evaluating internal validity requires the assessment of the roles played by bias, chance, and real effect. Each can operate, sometimes reinforcing other factors, sometimes offsetting them. There is often disagreement among experts, stemming from differing weights each places on the influence of bias, chance and real effect. Such differences in science are common, both in and out of court. The fact that two scientists have different judgments about how much weight to

---

[47] Thus Judge Kosinski, in the Daubert remand, writes of the plaintiff's case that it does not "attempt to show causation directly; instead, they rely on experts who present circumstantial proof of causation." Of course there is no such thing as a "direct" proof of causation.

[48] Professors Rothman and Greenland are not alone in their view that judgment – not a checklist – is a scientist's essential tool in inferring causation. Indeed, that perspective is shared by a number of the nation's leading epidemiologists and other scientists, historians of science, and philosophers of science. Thus, another amicus brief tendered to the US. Supreme Court in the Daubert case by Harvard professors Stephen Jay Gould (Zoology, Geology, and History of Science), Gerald Holton (Physics and History of Science), Everett Mendelsohn (History of Science), and Kathleen Joy Propert (Biostatistics), Columbia University professor Ronald Bayer (Sociomedical Sciences), and NYU professor Dorothy Nelkin (Sociology and Law) explained that "'[c]onclusiveness in inferring causality -- in epidemiology as with the study of all free-living human beings – is a desire more often than an accomplishment.'" Amicus Brief of Bayer, Gould, etc., quoting Mervyn Susser, Rules for Inference in Epidemiology, 6 Regulatory Toxicology and Pharmacology 116, 127 (1986). These scholars went on to observe that "[a]s a consequence, those who seek in science the immutable truth they find lacking in the law are apt to be disappointed." (Ibid.) Furthermore, "One notable similarity [between law and epidemiology] is the dependence of both fields upon subjective judgments. … In the end, a quality which lawyers should understand – judiciousness – matters more than any. Scientists use both deductive and inductive inference to sustain the momentum of a continuing process of research. … The courts of law, and the courts of application, use inference to reach decisions about what action to take. Those decisions cannot rest on certitudes, most especially when population risks are converted into individual risks." (Ibid., quoting Susser, op. cit., at p. 128 (my italics)).

give a study does not demonstrate that either has failed to use scientifically acceptable reasoning, but only that the ultimate opinion about the weight to accord a study is inherently part of the subjective judgment process of scientists.

An evaluation of internal validity helps a scientist in deciding how much to rely on the specific results of a particular experiment or study. It does not tell a scientist how much to extend that result to contexts or situations different than the one studied in the particular study, i.e., how much to generalize the result. A separate evaluation for external validity is needed.

2. External Validity: To What Extent Can Valid, Reliable, And Useful Generalizations Be Drawn From The Results Of A Particular Study?

Scientists observe and experiment in order to generalize, that is, to explain as much of the world as possible. Generalization is the source of science's fascination, power of explanation, and practical importance in the world outside the community of scientists. The limits and extent of the generalization that can be drawn from a given study constitute the study's external validity. For present purposes, the question is whether research results and conclusions developed in one context (e.g., an exposure to one organic solvent, or the exposure of a mouse to TCE at a particular dose) can be generalized to cover other contexts (e.g., an exposure to another, closely related, organic solvent, or the exposure of a human at a different dose). Defining and constraining generalizations is an active process of forming opinions about studies.

There are no fixed, definite, and generally agreed upon rules about how – and how far – to generalize. Each study must be evaluated in a specific context.

V. DEVELOPING AN OPINION ABOUT CAUSATION

A. Arriving at an Explanation: Assembling the Picture

Clinical observations and case reports, epidemiological and animal studies, and toxicological experiments are like the pieces of a picture puzzle, albeit with the difference that the pieces are being fit into a picture that is being formed in the mind of the scientist on the basis of, and at the same time, that the individual pieces are being discovered and taking shape, and with additional caveats that some existing pieces may not fit (and thus may not be used) and that not all of the pieces that might be needed to fill are available for placement in the picture when the scientist completes the process, let alone when he or she starts the process. All in all, fitting the pieces into a scientific picture is a fluid, dynamic, and difficult process.

Depending upon a scientist's judgment of the internal validity (or inherent quality) of a particular study, an individual "piece" may be clear and well defined, or fuzzy and indefinite. Depending upon a scientist's judgment of external validity of a particular study, he or she may decide that an individual piece forms a large and central part of the picture, or is just a small piece on the periphery of the picture, or not even part of the picture at all. In addition, a scientist's experience, expertise and basic judgment are involved. The objective for the scientist is to take the available picture pieces, judge their internal and external validity, and assemble a picture. The goal is a coherent, sensible, comprehensive, and "elegant" picture of "reality," i.e., a picture that represents his or her decision about "what is happening."

Constructing an explanation involves putting together the scientific evidence into a coherent picture. Clinical observations, toxicology, and epidemiology provide the puzzle pieces, but the parts do not always fit together smoothly or without gaps. Each puzzle piece represents or registers different aspects of the total picture, with results that show only a portion of the whole. Scientists are sometimes in the position of the three blind men and the elephant, one

feeling the long, tufted tail and concluding he was encountering a sleek and agile zebra-like creature, one feeling the muscular and prehensile trunk and concluding he had before him a very large snake, the third confronted with the massive body and believing he has hold of a rhinoceros.[49]

As already noted, interpreting a scientific study for use in assembling a coherent picture requires the use of critical thinking to weigh the various factors that might be responsible for the observed association. This includes evaluating the part played by bias, chance, and real effect, together and separately, and judgments on what generalizations are valid. In such a complex process and with practical matters of consequence at stake, it is not surprising that differences of opinion develop. It is also not surprising that such differences are highlighted and, indeed, magnified by the adversary process. But even when so magnified, such disagreements are not merely *artifacts* of the adversary process, but essential features of science as it is routinely practiced. The differences are not evidence of flawed scientific reasoning or methodology any more than a dissenting opinion in a legal decision proves the majority opinion has not followed accepted legal methodology or reasoning, or the subsequent reversal of the case by a higher court proves the opposite.

For many purposes a partial picture is sufficient. The simple knowledge that an animal is very heavy suffices to understand that if it sits on you the result will be disastrous. Thus, as a practical matter, if you are walking on a game trail in elephant country you need not know the internal anatomical details of the pachyderm that can crush you, only the fact that you can be crushed. Similarly, for example, in the field of toxicology, although scientists are still in

---

[49] One must of course ask what objective the scientist is concerned with. The last blind man may not have had valid grounds for concluding he had examined a rhinoceros, but still valid grounds for declining to crawl beneath the object of inquiry.

disagreement about the precise mechanism of asbestos carcinogenesis, no reputable scientist harbors doubts that asbestos is a carcinogen. The judgment that some chlorinated ethylenes are human carcinogens is similarly based on abundant evidence from animal studies, case reports and human epidemiology that cancer is a result of exposure, although the details may still be the subject of argument.

In sum, scientists may (and often do) disagree about which pieces are internally valid (which ones can be used in putting together a picture), disagree about which pieces are externally valid (relevant and suitable for fitting into the picture), and disagree about where each internally and externally valid piece should go, that is, just how to assemble the relevant pieces of the puzzle. What scientists do not disagree about, though, is that they routinely select pieces and assemble such pictures and call the end product of this process of selection and assembly an Explanation.

B. Tying It All Together: The Weight-of-the-Evidence Methodology

1. Reminder: Causation is not a factual datum but a judgment based on data

As noted and documented by authorities above, scientists agree that "causation" should *not* be regarded as an experimental or epidemiological result, but rather as a *"judgment"* made about the experimental or epidemiological data. How, then, are causal judgments made? Below, I give an account of how causal judgments are made in *practice* and I address a common misunderstanding among non-epidemiologists as to actual practice versus idealized renditions of what actually occurs. To introduce the discussion, I can do no better than quote Professors Rothman and Greenland, whose epidemiology text, now in its third edition (2008), is widely considered the most complete and sophisticated of its kind:

> If a set of necessary and sufficient causal criteria could be used to distinguish causal from noncausal relations in epidemiologic studies, the job of the scientist would be eased considerably. With such criteria, all the concerns about the logic or lack thereof in causal inference could be forgotten: It would only be necessary to consult the checklist of criteria to see if a relation were causal. We know from philosophy that a set of sufficient criteria does *not* exist. Nevertheless, lists of causal criteria have become popular, possibly because they seem to provide a road map through complicated territory.[50]

2. There are no fixed criteria or checklists for determining causality

One of the more popular "checklists" of "causal criteria" in the legal context is the so-called "Hill criteria," named for Sir Austin Bradford Hill, the biostatistician who first identified and summarized nine items that he carefully described as "viewpoints," not criteria.[51] Various truncated versions appear in many places, including the Federal Judicial Center's (FJC) *Reference Manual on Scientific Evidence*, where the criteria are attributed not to Hill but to Henle and Koch.[52] There is no doubt the "viewpoints" Hill enumerated (or their variants from other lists) can be useful as a framework to organize thinking, although it is interesting they are rarely used explicitly in the epidemiologic literature. In essence they are a summary of the kinds

---

[50] Rothman and Greenland, *op cit*., at p. 24.

[51] Sir Austin Bradford Hill, The Environment and Disease: Association or Causation?, Proceedings of the Royal Society of Medicine 295, 299 (1965), reprinted in Evolution of Epidemiologic Ideas: Annotated Readings on Concepts and Methods at p. 16 (Sander Greenland, ed., 1987). Hill's list of viewpoints also appeared in the eighth edition of his Principles of Medical Statistics, Oxford University Press, New York, 1966.

[52] Reference Manual, pp. 161ff. The Manual correctly characterizes the list as factors which might "guide the epidemiologist in making a judgment about causation." The attribution to Henle and Koch ("two infectious disease researchers") is curious. When Henle first enunciated far different criteria for contagious diseases in 1847, Koch was only 4 years old. Koch's later (1882) version is intimately tied to his development of the pure culture technique. Clearly the *logic* used by both Henle and Koch derives from the same source as Mill's Canons, but the Hill list is more obviously related to the latter than to Henle or Koch. See Rothman and Greenland, op. cit., p. 24.

of factors frequently considered, often only implicitly, when epidemiologists make causal judgments.

Unfortunately, the use and claimed importance of Hill's "criteria" by non-epidemiologists can reach absurd and misleading heights, as when it is declared that individual elements of the list (which, again, exist in many variants) *must* be fulfilled for a causal judgment to be made, or worse yet, that they must *all* be fulfilled. Indeed it is a serious misuse of Hill's viewpoints and of the scientific method to use the Hill considerations as a mechanical checklist. Neither EPA nor the International Agency for Research on Cancer (IARC) nor most epidemiologists use them in this fashion, or even use them explicitly at all. EPA and most epidemiologists (including myself) instead use the "weight-of-the-evidence" approach (see below). However, the Hill viewpoints are worth considering in terms of what they tell us about causal judgments and the factors that go into making them, as well as the ways these factors should be interpreted in the real world.

Hill's original list of characteristics of associations that are causal is as follows (I have added explanatory comments). Quotations are from Hill.[53]

a. Strength of the association

The stronger the association the more likely it is to be causal. The reasoning is that unknown confounders (*i.e.,* unknown factors leading to non-comparability) are unlikely to be powerful enough to explain a strong association. The Reference Manual adds, "The use of the strength of the association as a factor does not reflect a belief that weaker effects are rarer phenomena than stronger effects."[54]

---

[53] Op. Cit

[54] Reference Manual, p. 161, citing the first edition of Rothman's text; see also, Rothman and Greenland, op. cit.: "To some extent, this is a reasonable argument, but, as Hill himself acknowledged, the fact that an association is

b. Consistency

Is the association also seen in other studies using different designs, different populations, and different investigators? It is by no means required that the same effect be seen in *every* study. The situation is not unlike a doctor confronting a set of x-rays, one or more views of which (but not all) show a tumor. She does not disregard this evidence because not all views reveal the mass. In consistent associations, it may be less likely that the association is the result of some unseen design flaw or chance occurrence and this factor is cited repeatedly by defendants in legal cases as an argument against a causal relationship for various toxic agents, and on occasion been elevated to a criterion of cardinal importance. It is also given some emphasis by the *Reference Manual* (at p. 162). Rothman and Greenland, however, disagree, and I (and many others) agree with their reasoning:

> "Lack of consistency, however, does not rule out a causal association because some effects are produced by their causes only under unusual circumstances. More precisely, the effect of a causal agent cannot occur unless the complementary component causes act or have already acted to complete a sufficient cause [i.e., the "background conditions" that are also needed in any particular instance]. These conditions will not always be met....Consistency is apparent only after all the relevant details of a causal mechanism are understood, which is to say very seldom. Furthermore, even studies of exactly the same phenomena can be expected to yield different results simply because they differ in their methods and random errors. Consistency serves only to rule out hypotheses that the association is attributable to some factor that varies across studies."[55]

c. Specificity

If the effect seen is limited to certain kinds of workers and certain diseases (e.g., cancer at a specific site), this may be a strong argument for causation. Thus the disease mesothelioma is seen almost exclusively in asbestos workers. On the other hand, scientists know today of many

---

weak does not rule out a causal connection...a strong association is neither necessary nor sufficient for absence of causality." (pp. 24-25).

[55] Rothman and Greenland, op cit. at p. 25.

agents that cause a variety of common diseases (e.g., cigarettes cause lung cancer, emphysema and other diseases). Hill himself was very cautious about the use of this criterion: "We must not...over-emphasize the importance of the characteristic...In modern times the prospective investigations of smoking and cancer of the lung have been criticized for not showing specificity – in other words the death-rate of smokers is higher than the death-rate of non-smokers from a number of causes of death." The *Reference Manual* is even more cautious: "…epidemiologists no longer require that the effect of exposure to an agent be specific for a single disease." (p. 163). Rothman and Greenland are more scathing: "…specificity does not confer greater validity to any causal inference regarding the exposure effect. Hill's discussion of this criterion for inference is replete with reservations, but even so, the criterion is useless and misleading." (p. 25).

d. Relationship with time

Did the "cause" precede the "effect"? While this is a logical necessity for causation, many epidemiological study designs do not allow easy verification of temporal sequence, especially in long latency diseases where the exposure to a toxin and manifestation of a disease are decades apart. The *Reference Manual* accepts proper temporal sequence as a *sine qua non* for causal effect (p. 162), as do Rothman and Greenland, although even here, they make the following qualification:

> "This criterion is inarguable, insofar as any claimed observation of causation must involve the putative cause C preceding the putative effect D. It does *not*, however, follow that a reverse time order is evidence against the hypothesis that C can cause D. Rather, observations in which C followed D merely show that C could not have caused D in these instances; they provide no evidence for or against the hypothesis that C can cause D in those instances in which it precedes D." (p. 25).

e. Biological gradient (dose-response relationship)

If scientists increase exposure do they also see an increase in risk? When this is demonstrated it can be a persuasive argument for causality. Indeed, many researchers put a great deal of emphasis on this point. However, if this relationship is not found, it is *not* a persuasive argument against causality.[56] The *Reference Manual* goes further: "a dose-response relationship is not necessary to infer causation." Rothman and Greenland concur: "…the existence of [an increasing risk with increasing exposure] is neither necessary nor sufficient for a causal relation. A nonmonotonic relation [i.e., no "dose response"] only refutes those causal hypotheses specific enough to predict a monotonic dose-response curve." (p. 26). For example, the dose-response relationship may be more complex than a simple increasing function of dose. Other factors may also affect the relationship, including misclassification of exposure, a common source of bias in occupational and environmental epidemiology. There is a substantial body of technical literature on this question.

f. Biological plausibility

As Hill remarks, "It will be helpful if the causation we suspect is *biologically* plausible though this is a feature we cannot demand." [emphasis in original].[57]

Hill notes that what is biologically plausible depends on the contemporary state of knowledge "In other words, the association recorded may be one new to science or medicine and must not therefore be too readily dismissed as implausible or even impossible."[58]

---

[56] "The lack of a dose-response relationship is fairly weak evidence against causality. The measure of exposure may be misclassified, there may be a threshold necessary for the exposure to cause the disease, there may be bias in the measure of exposure. The presence of a dose-response relationship is relatively strong evidence for causality." Monson R, Occupational Epidemiology, 2nd Edition, CRC Press, Boca Raton, FL 1990

[57] A similar comment is made by Rothman and Greenland, op. cit., p. 26.

g. Coherence of the evidence

While a precise knowledge of the mechanism is thus not required, it is important that the alleged causal association not conflict with generally known facts about the disease, *i.e.,* the association should have coherence with those facts. Knowledge from animal and test-tube experiments clearly fit in here, but *"while such laboratory evidence can enormously strengthen the hypothesis of causation and may even determine the actual causative agent, the lack of such evidence cannot nullify the epidemiological observations in man."* [emphasis added]. This one of Hill's original factors is not included in the *Reference Manual*, possibly because the distinction with their "consistency" factor is ambiguous.

h. "The Experiment"

It may be possible to find instances where some intervention was followed by a decrease in disease, thereby providing another kind of "natural experiment." Such evidence, when available can be very persuasive. It would seem logical that the reverse, too, could be appropriately persuasive, as when a mother is given Bendectin during the relevant period of her pregnancy and then gives birth to a child with limb reduction deformities. While this does not *prove* that Bendectin is a teratogen, it certainly is plausibly an element in a causation judgment and is pertinent to a causation judgment in an *individual*.[59]

---

[58] Hill writes, concerning the modern appreciation of typhus as a disease spread from rats to humans via fleas, "It was lack of biological knowledge in the 19th [century] that led a prize essayist writing on the value and the fallacy of statistics to conclude, amongst other 'absurd' associations, that it could be no more ridiculous for the stranger who passed the night in the steerage of an emigrant ship to ascribe the typhus, which he here contracted, to the vermin with which bodies of the sick might be infected." Hill, op. cit..

[59] This was rejected by the District Court and by Judge Kosinski in the Daubert remand, however. As Hill's writings show, the reasoning is not scientifically incorrect in and of itself, and taken together with other evidence could well be determinative in the usual course of medical practice. See discussion of Dr. Palmer's testimony in the Daubert remand. This factor is not included in the Reference Manual. Similarly, Rothman and Greenland point out that

i. Reasoning by analogy

Analogy can be a valuable heuristic device. Hill comments, "With the known effects of the drug thalidomide and the disease rubella we would be ready to accept slighter but similar evidence with another drug or another viral disease in pregnancy."[60] This factor is not included in the *Reference Manual*. Rothman and Greenland accord it less weight than would I, but their logic is clear: "Whatever insight might be derived from analogy is handicapped by the inventive imagination of scientists who can find analogies everywhere. At best, analogy provides a source of more elaborate hypotheses about the associations under study; absence of such analogies only reflects lack of imagination or experience, not falsity of the hypothesis."

j. Hill's own formulation of his "viewpoints"

Hill summed up his concept of the value of his "viewpoints" as follows:

> "Clearly none of these nine viewpoints can bring indisputable evidence for or against a cause-and-effect hypothesis and equally none can be required as a *sine qua non*. What they can do, with greater or less strength, is to help us to answer the fundamental question – *is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?*" (emphasis in original).

Thus, Hill warned against a "hard and fast" checklist approach to science. As Sander Greenland, Professor of Epidemiology at UCLA and co-author of the most widely respected textbook on epidemiology has noted:

"[l]ogically…experimental evidence is not a criterion, but a test of the causal hypothesis…" and its interpretation is often difficult (p. 27).

[60] In the Daubert remand Judge Kosinski argued that evidence about Bendectin's similarity to thalidomide was not admissible because it only showed the "possibility of causation," which did not meet the standard of "preponderance of the evidence." This, however, is a serious misunderstanding of the role this evidence plays in the demonstration of causation. It should properly have been seen as an element in establishing an overall picture that explains ("provides a coherent scientific context for") the evidence.

> It is unfortunate that [Hill's] list or similar ones have been presented in textbooks as "criteria" for inferring causality of associations, often in such a manner as to imply that all the conditions are necessary. A careful reading of Hill shows that he did not intend to offer a list of necessary conditions; on the contrary, . . . he warned against laying down "hard and fast rules of evidence that must be obeyed before we accept cause and effect."[61]

Hill never intended to have his "viewpoints" replace common sense and judgment, but merely to aid them. They must be used judiciously.[62] "Causation" is not just a residual effect after bias and chance have been accounted for, but an *independent* factor that must be evaluated. Causation can still operate in the face of chance and bias, each of which are artifacts of the study design. By contrast, causation is an attribute of the real world.

Instead of inflexible checklists, *Hill himself* emphasized the role of judgment in applying his "viewpoints." Indeed, Hill stressed that judgment should guide the application of each of his "viewpoints."[63] For example, Hill warned that when scientists weigh the "strength," or "relative risk," of an association (the first of his "viewpoints"), they should take care to avoid "dismissing a cause-and-effect hypothesis merely on the grounds that the observed association appears to be slight."[64]

Hill not only applied his overall flexible philosophy regarding judiciousness to how causal judgments should be made but also to how each of his "viewpoints" should be assessed and applied. Thus, Hill thought it would be a mistake to overemphasize the "consistency of the relationship." He cautioned researchers against weighing too heavily the "consistency of the

---

[61] Sander Greenland, Preface to Austin Bradford Hill, The Environment and Disease: Association or Causation?, 58 Proceedings of the Royal Society of Medicine (1965) 295 at p. 299 (1965), reprinted in Evolution of Epidemiologic Ideas: Annotated Readings on Concepts and Methods 15, 19 (S. Greenland, ed., 1987) (my italics).

[62] The Reference Manual adds an additional consideration, "Have alternative explanations been ruled out?" (p. 163). Other writers do not include this on the list as it is usually considered a separate matter.

[63] Hill, Environment and Disease, in Evolution of Epidemiologic Ideas at p. 16.

[64] *Ibid.*

relationship" (the second of his viewpoints) and to remember that "there will be occasions when repetition is absent or impossible and yet we should not hesitate to draw conclusions."[65] Hill similarly admonished scientists against placing too much reliance on the importance of specificity (his third "viewpoint") because, as he explained, "[o]ne-to-one relationships are not frequent. … [I]f specificity exists we may be able to draw conclusions without hesitation; if it is not apparent, we are not thereby necessarily left sitting irresolutely on the fence."[66]

Likewise, when it came to "plausibility" (Hill's sixth viewpoint), he cautioned against requiring that this factor be met in every case for the simple reason that "[w]hat is biologically plausible depends upon the biological knowledge of the day. … [T]he association we observe may be one new to science or medicine and we must not dismiss it too light-heartedly as just too odd …." [67]

3. Federal, state, and international health research organizations rely on the weight-of the-evidence approach

In light of what Hill actually said (as opposed to what some think or wish he said), it is hardly surprising the EPA, many state and international health research agencies, and most practicing epidemiologists utilize what is called the "weight-of-the-evidence" method or approach, rather than a mechanical checklist approach that Hill rejected.

Although there is no fixed definition, the essence of the "weight-of-the-evidence" approach usually uses different types of data, evaluated together. This may include toxicology and chemical/structural studies, epidemiological studies and animal studies.

As the one state EPA explained:

---

[65] *Ibid.* at p. 17.

[66] *Ibid.* at p. 17.

[67] *Ibid.* at p. 18.

> "Both IARC and the EPA rank chemicals as to the weight of evidence that indicates their likelihood for causing cancer. The National Toxicology Program, United States Public Health Service, DHHS also annually lists chemicals known or reasonably anticipated to cause cancer. The Department accepts the expertise of all these groups and regulates water pollutants named in any pertinent subgroup of their lists as carcinogens. (The reason the lists are not identical to one another is because each group evaluates the weight of evidence for their priority chemicals, and there are differences in their priorities.)"[68]

The EPA weight-of-the-evidence approach is outlined in that agency's *Guidelines for Carcinogen Risk Assessment*.[69] As the EPA explained, in order to evaluate the "Overall Weight of Evidence for Human Carcinogenicity," researchers should evaluate:

> "all relevant information … to determine if the designation of the overall weight of the evidence needs to be modified. Relevant factors to be included [in the evaluation are]:
>
> - tumor data from human … studies;
> - tumor data from … animal studies;
> - structure-activity relationships;
> - short-term test findings;
> - results of appropriate physiological, biochemical, and toxicological observations; and
> - comparative metabolism and pharmacokinetic studies."[70]

The EPA uses this methodology in evaluating water and air contamination:

---

[68] Pennsylvania Guidelines for Development of Criteria for Toxic Substances and Guidelines for Development of Human Health Based Criteria, 25 Penn Code Section 16.33; Supp. 306 (May 2000) (emphasis added). See also, e.g., Washington State Water Quality Standards, Wash. Admin. Code 193-201A-020 (April, 2000) (relying on the EPA's weight-of-the-evidence approach); Texas Ecological Risk Assessment and Development of Ecological Protective Concentration Levels, Tex. Admin. Code, title 30, section 350.77 (Dec. 1999) (relying on a weight-of-the-evidence approach); Wyoming Water Quality rules and Regulations, Wyoming Administrative Code Environmental Water Quality, Ch. 17 Section 4 (April, 2000) (relying on the EPA's weight-of-the-evidence approach); Ohio Surface Water Quality Standards, Ohio Administrative Code, Section 3745-1-38 (April, 1999) (relying on the EPA's weight-of-the-evidence approach); Indiana Water Quality Standards, Indiana Administrative Code, Title 327, Rule 2-1.5-14 (March, 2000) (relying on the EPA's weight-of-the-evidence approach).

[69] 51 Fed. Reg. 33992 (1986)

[70] 51 Fed. Reg. at p. 34000.

> "The best available toxicity data on the adverse health effects of a chemical and the best data on bioaccumulation factors shall be used when developing human health Tier I criteria or Tier II values. The best available toxicity data shall include data from well-conducted epidemiologic and/or animal studies which provide, in the case of carcinogens, an adequate weight of evidence of potential human carcinogenicity and, in the case of noncarcinogens, a dose- response relationship involving critical effects biologically relevant to humans. Such information should be obtained from the EPA Integrated Risk Information System (IRIS) database, the scientific literature, and other informational databases, studies and/or reports containing adverse health effects data of adequate quality for use in this procedure." 71

International agencies use the same methodology:

> "The [weight-of-the-evidence] approach takes into account the cumulative weight of the many studies that address the question of injury or the likelihood of injury to living organisms. If, taken together, the amount and consistency of evidence across a wide range of circumstances and/or toxic substances are judged sufficient to indicate the reality or a strong probability of a linkage between certain substances or classes of substances and injury, a conclusion of causal relationship can be made. This conclusion is made on the basis of common sense, logic, and experience, as well as formal science.
>
> "The Commission notes that the definition is not based on arbitrary rules or formulae but is consistent with the use of the term in law and science. The question to be answered is "how and when do we know there is sufficient or accumulated knowledge so that a reasonable person will conclude that policy makers should act." (my emphasis)[72]

And this is exactly the approach used by public health scientists when making judgments about causality, and is the approach that has resulted in the judgment by almost all authorities, government and non-governmental entities and regulatory agencies that exposure to chlorinated ethylenes can be harmful to human health.

---

[71] Methodologies for Development of Human Health Criteria and Values, CFR Pt. 132, App. C on Water Quality Guidance (current through May 2, 2000).

[72] See Massachusetts Weight of the Evidence Workgroup, Draft Report: A Weight of the evidence approach for evaluating ecological risks," November 1995 (summarizing the methodology adopted by the International Joint Commission on the Great Lakes.

VI. OPINIONS ABOUT THE PUBLIC HEALTH RISKS ASSOCIATED WITH CONTAMINATION OF THE RESIDENTIAL ENVIRONMENT BY THE CHLORINATED ETHYLENES, PCE, TCE AND VC.

It is my opinion, within a reasonable degree of medical certainty, that exposure to PCE in the residential environment presents a public health risk to the Class Area. This risk is related to exposures to PCE and its degradation products.

A. Exposure to chlorinated solvents with special reference to PCE, TCE and VC

A solvent is any substance used to dissolve another substance. The most common solvent is water, but many substances, especially those with oils or oily substances as part of their make-up do not dissolve well in water. Thus another class of solvents, often called organic solvents, is frequently used. Organic solvents are compounds of carbon combined with hydrogen and often other elements (frequently chlorine) in various ways. These solvents dissolve fatty and oily substances easily, and in turn, dissolve in fatty and oily substances as well. It is partly this property which allows organic solvents to pass easily across what is normally a natural protective barrier between the brain (central nervous system) and the rest of the body. Some solvents dissolve relatively well both in water and fat. For the same reason, exposures of the mother to organic solvents during pregnancy results in exposure of the developing central nervous system (CNS) of the fetus.[73] In addition, the skin has a substantial fatty component, allowing these fat-soluble materials to be easily absorbed through the skin.[74]

---

[73] Welch, LS, "Organic Solvents," Ch. 19 in Paul, M., editor, Occupational and Environmental Reproductive Hazards, Williams and Wilkins, Baltimore, MD, 1993

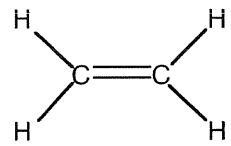
[74] Poet TS, Corley RA, Thrall KD, Edwards JA, Tanojo H, Weitz KK, Hui X, Maibach HI, Wester RC, "Assessment of the percutaneous absorption of trichloroethylene in rats and humans using MS/MS real-time breath analysis and physiologically based pharmacokinetic modeling," *Toxicol Sci*, 56:61-72, 2000

Many organic solvents are quite volatile (evaporate easily) as well, making exposure through inhalation possible. Such solvents, when present in groundwater can volatilize, get into the soil pores above the water table (the air spaces between the soil particles), and from there enter homes through foundations, crawlspaces and slabs.

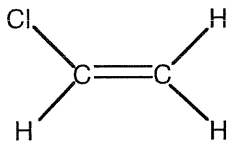
TCE, PCE and VC are typical chlorinated organic solvents in all these regards. Thus contamination of groundwater can result in exposure of the occupants through inhalation of indoor air.

B. Chlorinated ethylene organic solvents (PCE, TCE, VC and their relatives)

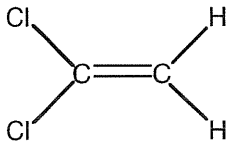
As noted above, chlorinated organic solvents are compounds composed of carbon, hydrogen and chlorine. A small subclass of chlorinated solvents, the chlorinated ethylenes, have been especially important. There are a total of seven members of this group, of which PCE, TCE and VC are three. They are all built on a common plan, a variation on the underlying structure of the hydrocarbon known as ethylene. Ethylene is composed of two carbon atoms, rigidly connected by a double bond. Each carbon has two more places to connect other atoms. If all four are composed of hydrogen, the result is the parent compound, ethylene (upper left, figure below). Successively replacing each of the hydrogens with a chlorine atom produces all the other members of the series:



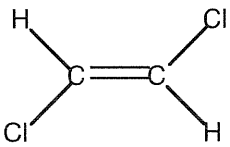
Ethylene



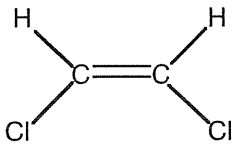
Vinyl chloride



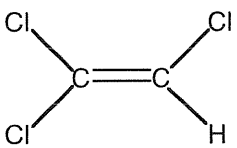
1, 1 - DCE



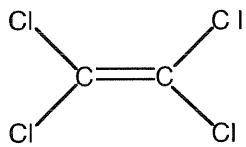
1, 2 - *trans* DCE



1, 2 - *cis* DCE



TCE



PCE

As can be seen, the result of the first substitution (upper right) results in vinyl chloride (VC), a known human carcinogen. There are three ways to make the second substitution (middle row), which produces three forms of the animal carcinogen dichloroethylene (DCE). There is one way to make the third substitution, producing TCE, and one way to make the fourth, which results in PCE. It is easily seen that the various chlorinated ethylenes are chemically closely related. It is often the case that the lower chlorinated chemicals, VC and DCE are degradation products of TCE and PCE from bacterial action in the ground.

All of these chemicals have uses in the chemical industry as feedstocks for other chemicals or plastics, and several, notably PCE and TCE, are commonly used solvents for

cleaning metal parts or in the dry cleaning industry. During natural biodegradation in groundwater PCE and TCE are converted first to DCE and then to VC.

C. Biological consequences of PCE exposure

In the ensuing discussion I will focus primarily on PCE, and my opinions in this case are sufficiently supported by the literature on PCE only. However, I will also make reference to the TCE and VC literature, for two reasons. The first is that there is evidence of TCE and VC contamination in this case. The second is that the additional scientific information on TCE and VC provides corroborating and confirming evidence for the effects of PCE, given the closely related chemical structure and toxicological effects of the compounds. In essence, this information *adds* to the "weight" in the weight-of-the evidence without being determinative in my opinion.

My opinion concerning the public health risks of PCE exposure are related to health outcomes mentioned in the scientific and medical literature which focus is: cancer effects.

D. PCE and the Risks of Cancer

1. What is cancer?

Cancer is a multifactorial process. Thus it is not valid to talk about PCE, or anything else, for that matter, as the *sole* cause of a cancer. Rather, it is an element in the cause of a cancer, and as we shall show shortly, one that can present a substantial risk.

The multifactorial nature of cancer progression is well known to scientists.[75] The multifactorial viewpoint is not merely a generality but has a concrete basis in current science and

---

[75] See, for example, Cullen MR, Rosenstock L, Brooks SM. "Clinical approach and establishing a diagnosis of an environmental medical disorder," Ch. 18 in Brooks SM, Gochfeld M, Herzstein J, Jackson RJ, Schenker MB (eds.)

is the product of scientific analytical principles which are generally accepted. In recent years we have learned an enormous amount about the steps needed to produce cancer and the outlines of the general picture are now clear.[76]

Cancer has it origins in a series of discrete genetic changes (fixed mutations) in formerly normal cells.[77] A normal cell grows or proliferates in a well-regulated manner, sensitive to the needs of the tissue, organ and organism of which it is a part. It must therefore sense and obey signals from outside itself and act accordingly. In this respect it is like a good citizen among the body's cells, seeing to its own needs but doing so in consideration of the needs of the larger community. A cancer cell, by contrast, is like a social deviant, growing where and when it pleases, heedless of the general needs of the body. When such a cell divides, each of the progeny cells behaves similarly. When a sufficient mass of such deviant cells is produced (a "tumor") it can interfere with the structure and function of the body and produce the symptoms and signs of the disease we call cancer. If a vital function is compromised, death ensues.

The system of signals and cues which are part of a cell's essential growth regulatory repertoire has at least four elements. They are: a set of "growth factors" released by nearby cells and tissues; receptors on the surface of the cell that bind with a growth factor and initiate a growth or inhibition signal to the interior of the cell; a system of interacting proteins within the cell that conveys the signal from the surface receptor to the cell's nucleus, where the growth

---

Environmental Medicine Mosby, 1995: "It may be added that the environmental exposure may be only a contributing factor to disease and would not be the sole cause. Indeed, most disease is multifactorial, and the exposure may represent only one factor contributing to disease pathogenesis." (p. 221).

[76] An excellent account of the main discoveries that have made this possible can be found in Weinberg's memoir of his discovery of oncogenes, *Racing to the Beginning of the Road*, Basic Books, 1996.

[77] A superb summary of the current state of knowledge is given by Weinberg, R. "Molecular mechanisms of carcinogenesis," Chapter 12, Part II in Dale DC and Federman DD, *Scientific American Medicine*, WH Freeman, 1978 - 1998. Weinberg is the Director of MIT's Whitehead Institute and one of the world's authorities on oncogenes.

directing elements are located; and a set of "nuclear transcription factors" that translate the conveyed signal into the actions of banks of genes that "orchestrate the growth and proliferation programs of the cell."[78] Each of the four parts of the system is itself under the control of a gene or group of genes.

It is now clear that derangements of various parts of this four-part system are capable of causing malignant change. If a gene controlling a component of this pathway becomes altered so as to cause the cell to undergo uncontrolled proliferation the gene is called an activated oncogene. Activated oncogenes are usually variants of normal genes ("proto-oncogenes"). One thing that can cause a proto-oncogene to be converted to an activated oncogene (the cancer causing variant) is a chemical or physical agent that alters the genetic material (the DNA) that makes up the gene. This is called a genotoxic effect of a chemical. I will discuss below the evidence that PCE is genotoxic.

One activated oncogene (called *ras* from its original discovery in rat sarcomas, but present in humans as well) is found in mutant form in about one quarter of all human cancers. It lies centrally in the pathway between the cell surface receptor and the nuclear transcription factor. When a signal passes the pathway the normal *ras* protein passes the signal along to the next component by turning "on" briefly, and then shutting off. Mutant *ras* proteins, however, get stuck in the "on" position and provide a constant stimulus to the cell to proliferate. Interestingly there is an association between certain kinds of *ras* mutations and exposure to PCE, TCE and one of its metabolites (DCA).[79]

---

[78] Ibid., p. 1.

[79] See Anna CH, Maronpot RR, Pereira MA, Foley JF, Malarkey DE, Anderson MW, "*ras* proto-oncogene activation in dichloroacetic acid-, trichloroethylene- and tetrachloroethylene-induced liver tumors in B6C3F1 mice," *Carcinogenesis* 15:2255-2261, 1994. The number of genes looked so far is small, but the data are suggestive. There seems to be a difference in the kinds of mutation at codon 61 in the H-*ras* oncogene of TCE and DCA treated mice

If a cell has an unusually large amount of one of the growth cell surface receptors it can result in repetitive firing in the absence of an external signal and cause uncontrolled proliferation. This seems to be the mechanism of cancer in tissues as diverse as malignant brain tumors (glioblastomas), breast cancers and stomach cancers, all of which have abnormal amounts of a specific cell surface receptor that signals the cell to proliferate. The same situation is found in breast, ovarian and stomach cancers that each have abnormal amounts of another specific receptor. [80] Thus while cancer is often said to be not one disease but a hundred different diseases, in another sense cancer has *common underlying mechanisms* whose diverse types of cancer depend only upon the tissues in which they happen to occur.

As we have begun to uncover the details of this beautiful system, we are seeing important refinements in the picture. Cells not only have "accelerators" (activated oncogenes) but also "brakes" ("tumor suppressor genes" or anti-oncogenes). Knocking out the brakes can be even more damaging than getting the accelerator stuck in the down position. Even with a stuck accelerator, intact brakes can save you. But the combination of defective brakes and a racing engine is a recipe for a cancer, and chlorinated ethylenes can accomplish both.

For example, there is now evidence that TCE plays a major role in causation of kidney cancer in exposed individuals by its association with mutations of an important tumor suppressor gene in the kidney, the von Hippel Lindau (VHL) gene. Bruning and coworkers[81] have found that VHL mutations, a genetic marker of common kidney cancer, are also present in abnormal

---

compared to this codon in the tumors of untreated mice. Since H-*ras* does not seem to be necessary for tumorigenesis here, the meaning of this effect is still not clear.

[80] The receptor in the first instance is called Epidermal Growth Factor receptor (EGF Receptor), in the second called *erb*-B2. Ibid., p. 1. See also DiFiore PP, Pierce J, Kraus MH, et al., "*erb*-B2 is a potent oncogene when over expressed in NIH/3T3 cells," *Science* 237:178, 1987.

[81] Bruning T, Weirich G, Hornauer MA, Hofler H, Brauch, "Renal cell carcinomas in trichloroethene (TRI) exposed persons are associated with somatic mutations in the von Hippel-Linday (VHL) tumor suppressor gene," *Arch Toxicol* 71:332-335, 1997.

prevalence among TCE exposed workers with kidney cancer. While the VHL gene is mutated in about 30% - 50% of common kidney cancer cases, an examination of 23 kidney cancer patients with TCE exposure found mutations in the VHL gene in 100% of the patients. Follow-up work confirmed this result. Thus Brauch et al.[82] have now shown that in 44 patients with kidney cancer and known industrial exposure to TCE there is a much higher frequency of VHL gene mutation, singly and multiply, than in 107 kidney cancer patients without TCE exposure.[83] In this series 75% of the TCE exposed workers had VHL mutations compared to 58% of the non-exposed workers, but of more interest was the presence of a specific mutation at nucleotide 454, present in 29% of the exposed workers but *none* of the 107 unexposed workers.[84] This is a very striking and important finding, representing, according to the authors, "the first molecular evidence for a relationship between exposure to a defined carcinogen, gene damage, and kidney cancer." (p. 859). There was also an evident dose-response relationship between degree of exposure and codon 454 mutation and number of mutations in the VHL gene (table 4 of paper).

We are just now beginning to unravel the details of the ways various chemicals can cause cancer. Even without knowledge of the finest details, however, it is clear along all the lines of evidence that TCE is capable of causing cancer in intact animals, including humans. Even in these early stages, however, we see evidence that TCE can affect both the accelerator (cell proliferation via the *ras* activated oncogene) and the brakes (the tumor suppressor VHL gene in

---

[82] Brauch H, Weirich G, Hornauer MA, Storkel S, Wohl T, Bruning T, "Trichloroethylene exposure and specific somatic mutations in patients with renal cell carcinoma," *JNCI* 91:854-861, 1999.

[83] The authors of this study interpreted the results to mean the VHL gene is a specific target in TCE induced cancer: "In addition to the available epidemiological studies the results are now further proof for human renal carcinogenicity induced by high occupational exposures to TRI [i.e., TCE]." See also Bruning T, Weirich G, Hornauer MA, Hofler H, Brauch H. "Renal cell carcinomas in trichloroethene (TRI) exposed persons are associated with somatic mutations in the von Hippel-Lindau (*VHL*) tumor suppressor gene," *Arch Toxicol* 71:332-335, 1997.

[84] The paper reports this percentage as 39%, but an examination of the tables shows the true number is 29% (13/44).

the kidney). Thus there is good evidence that this chlorinated ethylene affects the genetic machinery to cause cancer. This brings us to the realm of toxicologic mechanisms, which I discuss below in the section, *"Between the Bookends: Mechanisms."*

I discuss the general considerations for identifying carcinogens first, then give an overview of the evidence specifically for PCE, TCE and VC.

2. Cancer and animal bioassays

It should not be presumed that any of these chemicals is biologically capable of causing cancer, no matter how large or small the dose. Most chemicals cannot do this under any circumstances. This is because transforming a normal cell to a malignant cell is a very special kind of biological effect. A cancer cell is more like a "Super cell" than a damaged cell: it grows where it wants to grow and divides when it wants to divide without any heed or concern for other properly behaving cells in the same organ or tissue. Cancer cells do their damage by "out competing" normal cells for nutrients, blood supply, and space. They are social deviants in the cellular society. Although any chemical, even water, oxygen or common table salt, can make a cell run worse (i.e., have a toxic effect, just as opening up the back of a Swiss-watch and poking around with an ice-pick generally will make a watch run worse), most chemicals, no matter what the dose given, cannot cause a cell to run "better," that is, to become cancerous.[85]

Thus we do not (and should not) presume that any of the chlorinated ethylenes fall into this special category without evidence. The evidence that PCE, TCE and vinyl chloride *are* such chemicals comes from several sources.

---

[85] Here "better" is meant only relative to the cell itself. Clearly from the standpoint of the *organism* this is not a "better running" cell.

The "bookends" of the evidence, so to speak, are, on the one hand clear demonstration that animals given PCE (as well as DCE, VC and TCE, but not ethylene; the chlorines are needed) suffer cancer as a result, and on the other, human evidence from epidemiological studies that this is true for the human species as well (where naturally we cannot do a "formal" experiment). Between our two bookends comes a substantial amount of toxicological evidence to show how this happens on the molecular level. We have mentioned some of this evidence in our discussions of the VHL gene, oncogenes and anti-oncogenes where there is a large body of evidence from the laboratory. Although the picture here is continually being filled in, gaps remain. Below, we summarize the animal bioassays and epidemiological evidence, our two bookends.

3. Left Bookend: Animal bioassays

The "gold standard" for identification of chemical carcinogens is the chronic, long-term animal bioassay.[86] The usual animals are rodents (rats and mice). The laboratory species used are extremely well understood and characterized and can be bred to be genetically homogeneous, thus removing an important source of variation in species like human beings or non-laboratory rodents. The closeness in biochemistry and metabolism (but obviously not outer form) to humans is the basis for their widespread use in human health research, from cancer to Alzheimer's Disease to cystic fibrosis. Dr. David Rall, former Director of the National Institute of Environmental Health Sciences, put it this way:

> Both theoretical consideration and experience indicate that it is possible to test in laboratory animals chemicals to which humans are or will be exposed and to use these test results to predict

---

[86] "...if there is strong evidence that a chemical is carcinogenic in appropriate laboratory animal test systems, it must be treated as if it were carcinogenic in humans." From Rall D, "Relevance of Results from Laboratory Animal Toxicology Studies," in Chapter 13, Toxicology of Environmental Disease, *Public Health and Preventive Medicine, 12th Edition*, ed. by J. Last, 1986.

> in general terms what is likely to occur in the human population. Essential to this premise is the knowledge, derived from considerable basic research, that biological processes of molecular, cellular, tissue, and organ functions that control life are strikingly similar from one mammalian species to another. Processes such as sodium and potassium transport and ion regulation, energy metabolism, and DNA replication vary little in the aggregate as one moves along the phylogenetic ladder [i.e., from species to species]. The classic work on the transmission of neural impulses in the squid axon is directly relevant to humans. Extensive renal function studies in fish, rodents, and dogs set the basis for our current understanding of renal function and the treatment of hypertension in humans. *Also, the processes of cell replication and development of cancer are analogous in all mammalian species.* [emphasis added]

The bioassay is a classical experiment, in Mill's sense. One group of animals (usually a few hundred) is dosed with the chemical and another ("control") group is not. At various intervals the animals are examined for cancer and a comparison made of tumor formation in the two groups.[87] Results of such bioassays are an important part of the basis for regulatory decisions about the carcinogenicity of chemicals like the chlorinated ethylenes.

a. Reliance on animal studies is a standard feature of science and policy

Because of the close resemblance between humans and certain test animals, state, federal, and international agencies all rely on animal studies to establish carcinogenicity.

Typically, the states follow the many Federal regulatory agencies that rely on animal studies to establish carcinogenicity. Thus animal studies are utilized by the Food and Drug Administration (FDA), the Environmental Protection Agency (EPA), the Consumer Product

[87] Of course there are a variety of important technical details in a well-conducted bioassay, and interpreting the results requires making a judgment by applying the factors noted earlier in this report. In the case of the bioassays of PCE and TCE, however, it is generally accepted that these chemicals do cause cancer in the animals. Disagreements, insofar as they pertain to the bioassays, surround the meaning of the results for human exposures.

Safety Commission (CPSC), the Occupational Safety and Health Administration (OSHA), and other federal agencies with responsibilities of determining permissible exposure levels for the safe usage of prescription drugs and over-the-counter medicines, implanted medical devices, cosmetics, fungicides, rodenticides, insecticides, pesticides, and disinfectants, and other industrial, pharmaceutical, and household products.[88]

Indeed, Federal agencies have explained their reliance on animal tests to establish carcinogenicity in humans many times over the last 25 years. For example:

> All policies accept the use of animal data as predictive for human beings. Explicitly or implicitly, all the policies acknowledge that substances shown to be carcinogenic in animals should be presumed to present a carcinogenic hazard to humans. (U.S. Congress, Office of Technology Assessment, 1987)[89]

An often-quoted statement on the value of animal data in assessing human risk is that of the **International Agency for Research on Cancer (IARC)**, a research agency of the **World Health Organization of the United Nations**. Their principle is based on two points. First, that a number of chemicals first identified as animal carcinogens were later confirmed to cause cancer in humans. Second, all chemicals accepted as human carcinogens that have been adequately studied in animals are positive in at least one species of animal. Thus the relationship goes in both directions. IARC concluded:

> 'Although this association cannot establish that all animal carcinogens also cause cancer in humans, nevertheless, in the absence of adequate data on humans, it is biologically plausible and prudent to regard agents for which there is sufficient evidence of carcinogenicity in experimental animals as if they presented a carcinogenic risk to humans.' (IARC, 1987)."

---

[88] P. Chan & A. Wallace Hayes, "Principles and Methods for Acute Toxicity and Eye Irritancy," in *Principles and Methods of Toxicology* 169, 206-12 (A.W. Hayes ed., 1989); A.T. Mosberg & A. W. Hayes, "Subchronic Toxicity Testing", in Hayes, ibid. , pp. 221, 226-31.

[89] Office of Technology Assessment, *Identifying and Regulating Carcinogens: A Background Paper*, Congress U.S., OTA, Nov.1987

Similarly, the **National Research Council**, an arm of the **National Academy of Science** has stated:

> Positive results in properly conducted animal bioassays are considered to be predictors of qualitative response in humans (IARC, 1980; NRC, 1977, 1983; NTP, 1984; OSTP, 1985; OTA, 1981). The scientific rationale for this approach is simply that animals are the closest models to the human for cancer studies. In addition, many carcinogens produce cancer in several species, and all known human carcinogens have been shown to produce tumors in at least one animal model (NTP, 1984). Benzene and arsenic trioxide, the two former holdouts from this general rule, have now been shown to be carcinogenic in animals (Goldstein, et al.., 1982; Maltoni and Scarnato, 1979; Pershagen, et al.., 1984). For some chemicals (e.g., aflatoxin B1, DES, vinyl chloride, mustard gas, melphalan, and 4-aminobiphenyl), the positive results in experimental animals preceded the epidemiological evidence. The overall patterns of chemical metabolisms are generally similar in humans and laboratory animals (Rall, 1979), although the rates of metabolism and the type and site of cancer may differ (IRLG, 1979; OTA, 1981). For example, the metabolism of B(a)P is qualitatively the same in all species and systems studied. (Sims, 1976)."[90]

In a similar vein the **Office of Technology Assessment (OTA)** has stated:

> "Chemicals cannot be tested for carcinogenicity in humans because of ethical considerations. A substantial body of experimentally derived knowledge and the preponderance of expert opinion support the conclusion that testing of chemicals in laboratory animals provides reliable information about carcinogenicity. Animal tests employ whole mammal systems, and although they differ one from another, all mammals, including humans, share many biological features (NRC, 1977). ¶ Effects in animals, properly qualified, are applicable to man. This premise underlies all of experimental biology and medicine, but because it is continually questioned with regard to human cancer, it is desirable to point out that cancer in men and animals is strikingly similar. Virtually every form of human cancer has an experimental counterpart, and every form of multicellular organism is subject to cancer, including insects, fish, and plants. Although there are differences in susceptibility between different animal species, and between individuals of the same strain, carcinogenic chemicals will affect most test species, and there are large

---

[90] Safe Drinking Water Committee, NAS/NRC, *Drinking Water and Health*, National Academy Press, Washington, D.C. 1986. Cites within the NRC Report are given to show that their statement is based on an extensive review of the literature. The full citations can be found in the original Report.

> bodies of experimental data that indicate that exposures that are carcinogenic to animals are likely to be carcinogenic to man, and vice versa.") [91] (My italics).

As a final example, the **Occupational Safety and Health Administration (OSHA)** concluded, on the basis of detailed testimony from a wide range of recognized scientific authorities, and an extensively documented record that:

> "The validity of qualitatively extrapolating animal test results to humans is firmly based upon substantial and empirical evidence in the Record. Not only have experiments in test mammalian animals given positive carcinogenic test results for every compound known to cause cancer in humans, except arsenic and perhaps benzene, but although there may be wide variations in the susceptibility of various species to cancer, evidence indicates that a substance that causes cancer in one mammalian animal species is likely to do so in most other mammalian species tested. Substantial evidence and scientific data in the Record indicate, in sum, that laboratory animals are suitable test models for determining the cancer-causing potential of a toxic substance to humans.
>
> "OSHA concludes that the general principle that substances shown to be carcinogenic in test animals should be presumed to pose a qualitative carcinogenic hazard to exposed humans was overwhelmingly supported, except as so qualified below (in relation to the adequacy of the carcinogenicity test); indeed, the specific scientific documentation for the principle is steadily being enlarged."[92]

In fact, some federal agencies *insist* on animal tests. As a result of this well- grounded and generally accepted appreciation of animal tests, federal agencies not only accept animal tests as reliable predictors of cancer in humans, these agencies often insist upon animal tests before authorizing production, distribution, or use of regulated substances.[93]

---

[91] Office of Technology Assessment, *Assessment of Technologies for Determining Cancer Risks from the Environment*, Congress, U.S., OTA, June 1981

[92] *OSHA, Identification, Classification and Regulation of Potential Occupational Carcinogens*, Fed. Reg. 45:5001-5296, 1980.

[93] See, for example, Chan & Hayes, *op. cit.*, in *Principles and Methods of Toxicology*, pp. 206, 211-12; J. J. Cohrssen & V. T. Covello, U.S. Council on Environmental Quality, *Risk Analysis: A Guide to Principles and Methods for Analyzing Health and Environmental Risks* , CEQ, US Government Printing Office, 1989, p. 38.

Finally, the Federal Courts have recognized that the use of animal tests to predict carcinogenicity in humans is soundly based on the multiple similarities between animals and humans. See, for example, the opinion of Judge Edward Becker in the *Paoli Railroad Yard PCB Litigation*,[94] in which he remarked that humans and monkeys are likely to show similar sensitivity to polychlorinated biphenyls (PCBs). Although the observed rates of metabolism may differ between humans and animals, the biochemical and metabolic processes carried out in most organs are similar.[95]

b. TCE, PCE and VC bioassays for determining cancer risk in humans

By way of summary, the following table gives the results of various animal bioassays of the chlorinated ethylenes:

---

[94] 35 F. 3d 717 at p. 779 (3rd Cir. 1994)

[95] See also Rall DP et al., *Alternatives to* "Using Human Experience in Assessing Health Risks," 8 *Annual Review of Public Health* 8:355ff, 1987

POSITIVE ANIMAL CARCINOGENICITY STUDIES OF 1,1,2-TRICHOLORETHYENE (CAS#79-01-6)[96]

Species	Route	Organs	Reference
Rat	Oral	Kidney (males)	NTP, 1990
Rat	Oral	Kidney (male), testes	NTP, 1988
Rat	Oral	Leukemia (males)	Maltoni et al., 1986
Mouse	Oral	Liver	IARC, 1976: NTP, 1990
Rat	Inhalation	Testes, kidney (males)	Maltoni et al., 1986, 1988
Mouse	Inhalation	Lung, liver	Maltoni et al., 1988
Mouse	Inhalation	Lymphoma (females)	Henschler et al., 1980
Mouse	Inhalation	Lung	Fukuda et al., 1983
Mouse	Inhalation	Lung	Maltoni et al., 1986, 1988

[96] From International Agency for Research on Cancer, *IARC Monographs on the evaluation of the carcinogenic risks of chemicals to humans, vol. 11* (review of carcinogenicity study conducted by National Cancer Institute in 1976); also IARC Monographs, v. 19; National Toxicology Program, *Toxicology and Carcinogenesis Studies of Trichloroethylene (CAS No. 79-01-6) in F344/N Rats and B6C3F1 Mice (Gavage Studies)* Tech. Rept. Ser. 243, Research Triangle Park, NC 1990; National Toxicology Program, *Toxicology and Carcinogenesis Studies of Trichloroethylene (CXAS No. 79-01-6) in Four Strains of Rats (ACI, August, Marshall, Osborne-Mendel) (Gavage Studies) Tech. Rept. Series 273*, Research Triangle Park, NC 1988; Maltoni C, Lefemine G, Cotti G, "Experimental research on trichloroethylene carcinogenesis. In: C. Maltoni and M. Mehlman, *Archives of Research on Industrial Carcinogenesis* V.5, Princeton Scientific Publishing, 1986; Maltoni C et al, *Ann NY Acad Sci* 53:316-342, 1988; Henschler D, Romen W, Elsasser H, Reichert D, Eder E, Radwan Z, "Carcinogenicity study of trichloroethylene by long term inhalation in three animal species," *Arch Toxicol* 43:237-248, 1980: Fukuda K, Takemoto K, Tsuruta H, "Inhalation carcinogenicity of trichloroethylene in mice and rats," Ind Health 21:243-254, 1983.

POSITIVE ANIMAL CARCINOGENICITY STUDIES OF TETRACHLOROETHYLENE (PCE (CASE# 127-18-4)[97]

Species	Route	Organs	Reference
Mouse	Oral	Liver	NCI, 1977
Rat	Inhalation	Kidney (males), leukemia	Mennear et al., 1986
Mouse	Inhalation	Liver	Mennear et al., 1986

Both TCE and PCE have been adjudged to have sufficient evidence in experimental animals for carcinogenicity by the International Agency for Research on Cancer (IARC) of the World Health Organization.[98]

[97] National Cancer Institute (NCI) *Bioassay of Tetrachloroethylene for Possible Carcinogenicity (CAS No. 127-18-4)* Tech. Rept. Ser. 13, Bethesda, MD, 1977; Mennear J, Maronpot R, Boorman G, Eustis S, Huff J, Haseman J, McConnell E, Ragan H, Miller R, "Toxicologic and carcinogenic effects of inhaled tetrachloroethylene in rats and mice," In P Chambers P Gehring and F Sakai, eds., *New Concepts and Developments in Toxicology* Elsevier, pp.l 201-210, 1986.

[98] International Agency for Research on Cancer *Dry cleaning, some chlorinated solvents and other industrial chemicals*, v. 63 of IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, IARC, 1995, See pp. 136, 204.

Vinyl chloride has also been extensively tested in animals for its ability to produce cancer. As noted earlier, animal testing revealed VC's carcinogenicity more than three years before a cluster of a rare cancer in a VC workplace was reported in 1974. The EPA's Integrated Risk Information System (IRIS) database lists VC as a known human carcinogen.[99] The IRIS document makes explicit mention of the weight-of-evidence review process:

Vinyl chloride; CASRN 75-01-4 (08/08/2000)

> Health assessment information on a chemical substance is included in IRIS only after a comprehensive review of chronic toxicity data by U.S. EPA health scientists from several Program Offices and the Office of Research and Development. The summaries presented in Sections 1 and 11 represent a consensus reached in the review process.

The IRIS document gives detailed information on the animal bioassays which went into the weight-of-evidence review, and concluded the evidence for carcinogenicity in animals was:

> **Sufficient:** VC is carcinogenic in rodents by both oral and inhalation routes, and some data indicate that it produces tumors when given i.p. [intraperitoneally], s.c. [subcutaneously], and transplacentally.

Numerous studies detailed descriptions of eleven animal bioassays published in the scientific literature demonstrating the ability of VC to cause cancer in animals.[100]

---

[99] Online entry to the IRIS database here: http://www.epa.gov/IRIS/subst/1001.htm

[100] Feron, V; Hendrikson, CFM; Speek, AJ, et al. Lifespan oral toxicity study of vinyl chloride in rats. Food Cosmet Toxicol 19:317-333, 1981; Til, HP; Immel, HR; Feron, VJ. Lifespan oral carcinogenicity study of vinyl chloride in rats. Final report The Netherlands: Civo Institutes, TNO, Report No. V 83.285/291099, TSCATS Document FYI-AX-0184-0353; Til, HP, 1983; Feron, VJ; Immel, HR. Lifetime (149-week) oral carcinogenicity study of vinyl chloride in rats. Food Chem Toxicol 29:713-718, 1991; Maltoni, C; Lefemine, G; Ciliberti, A; et al. Carcinogenicity bioassays of vinyl chloride monomer: a model or risk assessment on an experimental basis. Environ Health Perspect41:3-29, 1981; Suzuki, Y. Pulmonary tumors induced in mice by vinyl chloride monomer. Environ Res 16:285-301,1978; Maltoni, C; Lefemine, G; Ciliberti, A; et al. Experimental research on vinyl chloride carcinogenesis, Vol. 1 and 2. In: Archives of research on industrial carcinogenesis. Princeton, HJ: Princeton Scientific Publishers, Inc., 1984; Suzuki, Y. Neoplastic effect of vinyl chloride in mouse lung – lower doses and short term exposure. Environ Res 32:91-103, 1983: Elehir, RM; McNamara, BP; McLaughlin, J; et al. Cancer induction following single and multiple exposures to a constant amount of vinyl chloride monomer. Environ Health Perspect 41:63-72; 1981: Hong, CB; Winston, JM; Thornburg, IP; et al. Follow-up study on the carcinogenicity of vinyl chloride and vinylidene chloride in rats and mice: tumor incidence and mortality subsequent to exposure. J Toxicol Environ Health 7:909-924, 1981; Bi, W; Wang, Y; Huang, M; et al. Effect of vinyl chloride on tests in rats. Ecotoxicol Environ Safety 10:281-289, 1985; Dre, RT.

I have carefully considered the weight of the evidence in the case of PCE, TCE and VC and have concluded, in agreement with standard practice and federal agencies like the EPA, as noted above, that the bioassays are relevant for human cancer risks at concentrations found in the Class Area in Madison, Wisconsin.

c. The Right Bookend: Epidemiological Studies of TCE, PCE, VC and Cancer

I note at the outset that an exhaustive review of the scientific evidence on carcinogenicity of TCE and PCE was conducted in 1995 by the International Agency for Research on Cancer (IARC).[101] It concluded:[102]

"Trichloroethylene *is probably carcinogenic to humans (Group 2A).*

"In making the overall evaluation, the Working Group considered the following evidence:

- **Although the hypothesis linking the formation of mouse liver tumors with peroxisome proliferation is plausible, trichloroethylene also induced tumors at other sites in mice and rats.**[103]
- **Several epidemiological studies showed elevated risks for cancer of the liver and biliary tract and for non-Hodgkin's lymphoma." (italics in original).**

"Tetrachloroethylene [PCE] *is probably carcinogenic to humans (Group 2A). "In making the overall evaluation, the Working Group considered the following evidence:*

---

[101] IARC is one of the constituent agencies of the United Nations. My experience with IARC is that it takes relatively conservative stands resulting from the representation of a number of countries whose environmental and occupational protection records are less than optimal, and who are frequently reluctant to accept the hazards of industrial chemicals of economic importance.

[102] International Agency for Research on Cancer, *Dry Cleaning, Some Chlorinated Solvents and Other Industrial Chemicals*, IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 63, Lyon 1995, pp. 136-137.

[103] The question of toxicological mechanism involving peroxisome proliferation is discussed below in section II.C.4. My evaluation is roughly the same as IARC's.

- **Although tetrachloroethylene is known to induce peroxisome proliferation in mouse liver, a poor quantitative correlation was seen between perioxisome proliferation and tumor formation in the liver after administration of tetrachloroethylene by inhalation. The spectrum of mutations in proto-oncogenes in liver tumors from mice treated with tetrachloroethylene is different from that in liver tumors from mice treated with trichloroethylene.**
- **The compound induced leukemia in rats.**
- **Several epidemiological studies showed elevated risks for esophageal cancer, non-Hodgkin's lymphoma and cervical cancer." (italics in original).**[104]

It is important to emphasize that IARC is not a public health regulatory agency and they make clear in their documentation that their evaluation is on the basis of the underlying science. It is not intended to be a "health protective" judgment but just a scientific judgment.

The evolution of IARC's treatment of TCE and PCE over the years is instructive. In a review of the status of IARC animal carcinogens, Karstadt[105] made the following observations about the two chemicals:

"Trichloroethylene and tetrachloroethylene have been reviewed by IARC panels several times: three times (volumes 11, 20, supplement 7, and volume 63) for trichloroethylene. Until the consensus meeting that resulted in volume 63 (published 1995) animal evidence for the two chemicals was evaluated as limited and human evidence as inadequate; both evaluations were raised in volume 63, to sufficient in animals and limited in humans. The IARC reviews of those two chemicals clearly show the gradual accretion of human evidence over the years as well as the development of definitive animal data."

---

[104] Ibid.

[105] Karstadt M, "Availability of epidemiologic data for chemicals known to cause cancer in animals: an update," *AM J Ind Med* 34:519-525, 1998.

Similarly, both TCE and PCE were listed, since the 9th Edition of the Report on Carcinogens, as "reasonably anticipated to be a human carcinogen." The Report is from the National Toxicology Program, National Institute of Environmental Health Sciences of the NIH.

Once again, the judgments are based on scientific, not regulatory considerations.

At the very least, it is clear there is independent, informed scientific opinion that accepts the proposition that TCE and PCE are *probable* human carcinogens.

Because of the scientific evidence of carcinogenicity both PCE and TCE are regulated as carcinogens by federal and state governments. The current public health based Maximum Contaminant Level Goal (MCLG) for both PCE and TCE in community drinking water is 0 parts per billion. An MCLG is defined by USEPA as the level of a contaminant in drinking water below which there is no known or expected risk to health. Furthermore, California has listed both TCE and PCE as substances that cause cancer under Proposition 65. This statute requires the Governor to revise and republish at least once per year a list of chemicals known to the State to cause cancer.

These chemicals are not carcinogens because they are regulated as such. They are regulated as such because they are considered to be carcinogens.

i. Evaluating the epidemiological literature

Given the carcinogenic potential of PCE and TCE in animals and the demonstrated ability of VC to cause cancer in animals, it is a natural question to ask if humans exposed to these chemicals are similarly affected. The available epidemiologic evidence is entirely consistent

with a PCE and TCE cancer risk in humans and abundantly confirms that VC is a cancer risk in humans.

PCE and TCE are used in a variety of workplaces, many of them very difficult to study epidemiologically. Many of these work settings are in small shops with transient workforces and consequent difficulty in following them for the periods of time and in sufficient numbers to obtain informative results. Another major problem is the presence in most of these workplaces of other substances, often closely related solvents. This is an unavoidable complication, but one commonly found in occupational studies, such as aniline dye workers, chemical workers, roofers, and many others. In those cases the presence of possible other agents did not prevent a judgment that particular substances in the workplace were among the causes of cancer, nor should this "chemical shell game" be used in this instance.

One exception to the multiple exposure problem in the occupational setting has been studies of laundry and dry-cleaning workers. Dry-cleaners have used a variety of solvents over the years, but principally PCE, petroleum solvents, and in earlier years, TCE. They thus constitute a working group exposed to two of the main chlorinated ethylenes, PCE and TCE. Taken together, these studies have showed associations of dry cleaning work with blood cancers, cancer of the urinary tract (bladder and kidney), cervix, lung, colon, pancreas, and liver.

Likewise, environmental investigations of drinking water contamination with PCE and TCE constitute additional opportunities to verify that the increases in cancer seen in PCE/TCE-workplaces are at least in part due to exposure to these chlorinated ethylenes.

A summary of the epidemiological literature follows:

SUMMARY OF EPIDEMIOLOGICAL STUDIES
PERTAINING TO TCE, PCE AND CANCER*

REFERENCE	STUDY POPULATIONS	STUDY DESIGN	RESULTS
McMichael et al., 1975	Rubber workers	Retrospective cohort mortality study	Leukemia in rubber workers exposed to several solvents, including TCE
Blair et al., 1979	330 deceased members of laundry and dry cleaning union, St. Louis	Proportional mortality study US population	Relative excess total cancers, lung, cervix; slight excesses in liver & leukemia
Kaplan, 1980	1597 drycleaners whose principal exposure was PCE	Retrospective cohort mortality study	Colon cancer (SMR 182) plus elevations involving small numbers for cancers of the pancreas and urinary tract and diseases of the blood forming organs
Olsson and Brandt, 1980	25 men admitted with HD, 50 matched controls	Case-control study	OR solvents 6.6, 3 cases, no controls exposed to TCE
Hardell, 1981	169 lymphoma cases, 338 controls in Sweden	Case-control study	OR solvents 4.6; OR TCE 4.8
Katz and Jowett, 1981	671 white female drycleaning workers who died in Wisconsin	Proportional mortality study	Relative excess in deaths from cancer of the kidney (unspecified site), with smaller excesses in bladder cancer, skin cancer and lymphosarcoma
Peters, 1981	Parents 92 children with brain CA $\leq$ 10 yrs., 92 matched controls in LA	Case-control study	OR = 9 for employment in aircraft industry, 2 fathers exposed to TCE
Duh and Asal, 1984	Deaths from 1975-1981 among laundry and drycleaning workers in Oklahoma	Proportional mortality study	Elevated standardized proportional mortality odds ratios for respiratory cancer and cancer of the kidney
Barret, 1984	235 deaths in TCE and cutting oil exposed workers	Cohort study	Naso- and oropharynx CA (SMR 2.5)

Hardell, 1984	102 liver cancers and 204 matched controls in Sweden	Case-control study	OR solvents and primary liver CA 1.8, hepatocellular CA 2.1
Hernberg et al., 1984	126 primary liver CA from Finnish registry 1979-80; 324 hospitalized controls with dx MI	Case-control study	OR solvents 2.3
Shindell and Urich, 1985	2646 workers (production and office) employed ≥ 3 months 1957-1983 in plant where TCE used as degreaser; some water contamination at plant; f/u to 1983	Cohort study	Decreased mortality from all causes in workers compared to national rates
Brown and Kaplan, 1985	1597 drycleaners whose principal exposure was PCE, exposed > 1 year before 1960	Extension of cohort study reported in Kaplan, 1980	Excess deaths from malignant neoplasms, cervical cancer, urinary tract cancer (both bladder and kidney)
Barret et al., 1985	Workers exposed to TCE and cutting oils	Cohort study	Excess cancer of the naso- and oropharynx
Axelson et al., 1986	Workers exposed to TCE and PCE	Cohort study	Slight excess incidence of bladder cancer and lymphoma
Lagakos et al., 1986	Population of Woburn, MA	Nested case-control	Leukemia OR = 2.2 (1.5-2.9)
Lowengart, et al., 1987	Parents of children with leukemia	Case-control study	Excess risk if parents occupationally exposed to TCE
Garabrant et al., 1988	14,067 workers employed ≥ 4yr in aircraft plant 1958-1982	Cohort study	No excesses
Hernberg, 1988	344 primary liver CA in Finnish registry, 1976-1978, 1981; 385 controls with dx MI and 476 deceased stomach CA controls	Case-control study	OR solvents .6 men, 3.4 women
Silverman, et al., 1989	Bladder cancer patients in national study	Case-control study	Excess risk in solvent-exposed workers
Sharpe, 1989	164 kidney CA, 161 non-CA kidney disease controls	Case-control study	OR solvents 3.4 (TCE, PCE, TCA and DCM most commonly used)
Fredricksson, et al., 1989	Colon cancer patients	Case-control study	Excess cancer in solvent-exposed workers

Olsen, 1989	2610 white males employed ≥ 1yr in chemical company in LA between 1956 and 1980; f/u to 1981; plant made PCE and other solvents	Cohort study	Leukemia/aleukemia SMR 4.9 (various types, differing employment histories)
Blair, 1990	5365 members of a drycleaning union employed ≥ 1yr between 1945 and 1977, followed through 1978	Retrospective cohort mortality study	Increased mortality from esophagus (SMR 2.1, black men 3.5), larynx (SMR 1.6), lung (SMR 1.3), cervix (SMR 1.7), bladder (SMR 1.7), NHL (SMR 1.7), HD (SMR 2.1), thyroid (SMR 3.3), high exposure to drycleaning solvents, blood CA (SMR 4.0)
Bond, 1990	44 liver CA from 6259 deaths in hourly workers Dow Chemical 1940-1982; random sample of other deaths (1888) as controls	Nested case-control study	OR PCE 1.8
Lynge and Thygesen, 1990	10,600 Danish laundry and dry cleaning workers, followed for 10 years from 1970; 1/4 worked in dry cleaning but could not be individually identified; PCE, TCE and CFC exposure	Cohort study	Excesses in lung (1.2), liver (2.2) and pancreatic CA (1.7)
Mallin, 1990	Town in NW Illinois	Cross-sectional	Bladder, RR = 1.7 male, RR = 2.6 female
Fagliano et al., 1990	Residence in one of 42 towns in NJ	Cross-sectional, TCE in town water	Leukemia RR = 1.4 (1.1-1.9) females, RR = 1.0 (.7 – 1.5) for males
Vartiainen et al., 1993	Residence in two villages	Cross sectional, TCE exposure in town water; comparison nat'l rates	Leukemia Town A, 1.2 (.8-1.7), Town B .7 (.4-1.1); HD Town A, .8 (.3-1.7), Town B, 1.4 (.7-2.5); Liver Town A, .7 (.3-1.4), Town B, .6 (.2-1.3) Multiple myeloma Town A, .7 (.3-1.3),Town B, .6 (.2-1.3) NHL Town A, .6 (.3-1.1), Town B, 1.4 (1.0-2.0)

Cohn et al., 1994	Residence in one of 75 towns in NJ	Cross-sectional, TCE exposure in town water	Leukemia RR = 1.4 (1.1, 1.9) females, RR = 1.1 (.8, 1.4); NHL RR = 1.4 (1.1, 1.7) females, RR = 1.2 (.9, 1.5) males
Spirtas et al., 1991	6929 employees exposed to solvents (include. PCE) and TCE in aircraft maintenance 1952-56	Cohort mortality	SMRs men Buccal/pharynx .9 (.3, 2.1) Esoph. 1.1 (.4, 2.3) Stomach .9 (.5, 1.5) Colon 1.1 (.7, 1.6) Rectum .6 (.2, 1.6) Bil./liver 2.0 (.9, 3.9) Pancreas .8 (.5, 1.4) Larynx .3 (.0, 1.9) Lung 1.0 (.8, 1.3) Prostate .8 (.5, 1.2) Kidney 1.2 (.5, 2.4) Bladder 1.4 (.7, 2.5) Melanoma1.0 (.3, 2.2) Brain .9 (.4, 1.7) Bone 2.6 (.5, 7.7) NHL 1.0 (.5, 1.9) HD.9 (.3, 2.4) Leukemia, .7 (.3, 1.3) Mult. Myeloma 1.1 (.4, 2.6) SMRs women: Colon .4 (.0, 1.3) Pancreas .8 (.1, 2.9) Breast .8 (.4, 1.5) Uterus 1.0 (.3, 2.5) Cervix 2.2 (.6, 5.7) NHL 2.9 (.8, 7.3) SMRs PCE exposure, women Mult. myeloma 17.1 (2.1, 61.6) NHL 9.7 (1.2, 35.0)
Aschengrau et al., 1993	Permanent residents of 5 towns on Cape Cod, MA, exposed to PCE in water	Case-control	Leukemia OR = 8.3 (1.5-45.3) Bladder OR = 4.0 (.7-25.1)

Axelson et al., 1994	1421 workers exposed to TCE, 1958-1987, biomonitored for U-TCA	Cohort mortality	SIRs Stomach .7 (.2, 1.6) Colon 1.0 (.4, 2.0) Liver 1.4 (.4, 3.6) Pancreas .3 (.0, 1.4) Larynx 1.4 (.2, 5.0) Lung .7 (.3, 1.3) Prostate 1.3 (.8, 1.8) Testis 2.0, (.3, 2.5) Kidney 1.2 (.4, 2.5) Bladder 1.0 (.4, 2.0) Skin 2.4 (1.0, 4.7) NHL 1.6 (.5, 3.6) HD 1.1 (.0, 6.0) Multiple myeloma .6 (0.0, 3.2)
Ruder et al., 1994	1109 women, 592 men drycleaners, employed at least 1 yr. before 1960 at shop using PCE followed through 1990 (update Brown/Kaplan)	Cohort mortality	SMRs PCE-only sub-cohort (CI) All 1.01 (.76, 1.32) Buccal 2.5 (.52, 7.33) Tongue 7.25 (.88, 26.2) Esoph. 2.64 (.72, 6.76) Stomach 0 Colon 1 (.32, 2.33) Rectum 0 Pancreas .73 (.09, 2.62) Lung 1.12 (.61, 1.88) Breast 1 (.36, 2.17) Female genital 1.57 (.68, 3.1) Male genital .89 (.11, 3.21) Kidney 1.16 (.03, 6.45) Bladder 0 Lymph/hem .49 (.06, 1.77) SMRs PCE-plus sub-cohort (CI) All 1.33 (1.13, 1.56) Buccal 1.2 (.3, 3.6) Tongue 1.8 (.0, 9.7) Esoph. 1.9 (.7, 4.1) Stomach.9 (.3, 2.0) Colon 1.8 (1.1, 2.7) Rectum 1.8 (.6, 4.2) Pancreas 2.1 (1.1, 3.6) Lung/resp. 1.2 (.8, 1.7) Breast 1.1 (.6, 1.9) Female genital 1.2 (.6, 2.0) Male genital .9 (.3, 2.0) Kidney 1.6 (.3, 4.7) Bladder 3.5 (1.6, 6.7) Lymph/hem. .8 (.3, 1.6)

Anttila et al., 1995	2050 men and1924 women who were biomonitored and followed up between 1962 and 1992	Cohort study	SIRs PCE Cervix 3.2 (.4, 12) Kidney 1.8 (.2, 6.6) Brain 1.2 (.1, 4.2) NHL 3.8 (.8, 11) SIRs TCE Stomach 1.3 (.8, 2.0) Colon .8 (.4, 1.7) Liver/bil. [1.9. IARC] (.9, 3.6) Cervix 2.4 (1.1, 4.8) Prostate 1.4 (.7, 2.4) Kidney .9 (.3, 1.9) Bladder .8 (.3, 1.9) Brain 1.1 (.5, 2.1) Lymp./hem. 1.5 (.9, 2.3) NHL 1.8 (.8, 3.6) HD 1.7 (.4, 5.0) Leukemia 1.1 (.3, 2.5)
Henschler et al., 1995	169 men exp. to TCE at factory working at least 1 yr. between 1956 and 1975, followed to 1992; control of 190 men at same factory w/o exposure to TCE (no office workers)	Cohort study	Kidney SIRs with 3 comparisons 11.2 (4.5, 23.00)Den. reg. 13.5 (5.4, 27.9) GDR reg. 7.2, internal comparison SMRs Lung 1.4 (.6, 2.9) Kidney 3.3 (.4, 11.8) Brain 3.7 (.1, 20.6) Lymph./hem. 1.1 (.0, 6.1)
Mass DPH, 1997	19 leukemia cases/37 controls, Woburn Mal, 1969-89	Case control	Leukemia OR with TCE contam. water 2.4 (.54, 10.6), overall OR w/ exp. in preg. 8.3 (.7, 95) OR 2 y before concep. 2.6 (.5, 14) OR p/birth 1.2 (.3, 5)
Aschengrau et al., 1998	258 breast cancer cases and 686 controls, permanent residents of 5 towns on Cape Cod, MA , exposed to PCE in water	Case control	Breast CA OR 7.8 (.9,16.7), 9 yrs latency and 90th%

Blair et al., 1998	Cohort of 7204 aircraft maintenance workers(1952-1990) exposed to TCE	Cohort mortality study	Esophagus SMR 5.6 (.7, 44.5) Stomach SMR .9 (.4, 1.9) Colon SMR 1.4 (.8, 2.4) Rectum SMR.4 (.1, 1.5) Biliary/liver SMR 1.3 (.5, 3.4) Prim. liver SMR 1.7 (.2, 16.2) Pancreas SMR 1.2 (.6, 2.3) Lung SMR .9 (.6, 1.3) Breast SMR 1.8 (.9, 3.3) Cervix SMR 1.8 (.5, 6.5) Prostate SMR 1.1 (.6, 1.8) Kidney SMR 1.6 (.5, 5.1) Bladder SMR (.5, 2.9) Melanoma SMR 1.0 (.3, 3.1) Brain SMR .8 (.2, 2.9) Endocrine SMR .7 (.1, 5.4) Bone SMR 2.1 (.2, 18.8) Lymph./Hem. SMR 1.1 (.7, 1.8) NHL SMR 2.0 ((.9, 4.6) Leukemia SMR .6 ((.3, 1.2) HD SMR 1.4 (.2, 12) Mult. myel. SMR 1.3 (.5, 3.4)
Morgan et al., 1998	4733 aerospace workers exposed to TCE	Cohort mortality study	RR from internal analysis w/ Cox prop. hazards, cum. high Lymph/hemat. 1.03 (.59, 1.79) Lymphoma .81 (.1, 6.49) Liver 1.19 (.34, 4.16) Kidney 1.59 (.68, 3.71) Bladder 2.71 (1.1, 6.65) Prostate 1.35 (.75, 2.44) Ovarian 7.09 (2.14, 23.54)
Vamvakas et al., 1998	58 kidney cancer cases and 84 controls (accident wards)	Hospital-based case-control study	Adj. ORs TCE/PCE 10.8 (3.36, 34.75)
Paulu et al., 1999	326 colorectal CA, 252 lung CA, 37 brain CA, 37 pancreas CA, and controls, permanent residents of 5 towns on Cape Cod, MA , exposed to PCE in water	Case control	Adj. ORs Lung CA, 90th% 3.7 (1.0,11.7) Colorectal CA 1.7 (.8,3.8) ever exp. and 9 yrs latency Crude OR Brain .7 (0,3.4), ever exp., 9 yrs latency Pancreas 0, ever exp. 9 yrs latency

Boice et al, 1999	77,965 aircraft manufacturing workers potentially exposed to TCE, PCE and CrVI, 1960-1997	Historical cohort	SMRs (C.I.)/TCE, tbl 8 All .86 (.76, .97) Buccal .93 (.2, 1.4) Esophagus 83 (.34, 1.72) Stomach 1.32 (.77, 2.12) Colon 1.07 (.72 1.52) Rectum 1.29 (.59, 2.45) Liver .54 (.15, 1.38) Pancreas ..41 (.17, .85) Larynx 1.1 (.3, 2.82) Lung .76 (.6, .95) Bone 1.44 (.04, 8.02) Connec. Tissue .1.94 (.4, 5.67) Melanoma ..46 (.06, 1.67) Breast 1.31 (.53, 2.69) Uterus .74, (.02, 3.57) Cervix 0 (0, 5.42) Ovary .58 (.01, 3.22) Prostate 1.03 (.7, 1.45) Testis/genital 0 (0, 5.42) Kidney .99 (.4, 2.04) Bladder .55 (.18, 1.28) CNS (.54 (.15, 1.37) NHL 1.19 (.65, 1.99) HD (2.77 (.76, 7.1) MM .91 (.3, 1.99) Leukemia 1.05 (.54, 1.84) SMRs (C.I.)/PCE, tbl 8 All .89 (.82, .86) Buccal .43 (.2, .82) Esophagus .83 (.49, 1.31) Stomach .76 (.48, 1.13) Colon 1.05 (.81, 1.33) Liver .92 (.54, 1.47) Pancreas .77 (.52, 1.09) Larynx .55 (.18, 1.29) Lung .88 (.77, 1.01) Bone .57 (.01, 3.18) Connec. Tissue 1.21 (.39, 2.82) Melanoma .87 (.42, 1.6) Breast 1.26 (.7, 2.07) Uterus .31 (.01, 1.71) Cervix 0 (0, 2.37) Ovary .57 (.07, 2.07) Prostate 1 (.78, 1.26) Testis/genital 3.04 (1.12, 6.63) Kidney .81 (.44, 1.36) Bladder .85 (.49, 1.35) CNS .68 (.36, 1.16) NHL 1.02 (.68, 1.47) HD 1.61 (.59, 3.51) MM .98 (.55, 1.61) Leukemia 1.02 (.68, 1.48)

Dosemici et al., 1999	796 cases, 707 controls (both population based, controls stratified for age and gender	population-based case-control for kidney cancer	TCE OR men OR 1.04 (.6, 1.7) women OR 1.96 (1.0, 4.0) PCE OR men OR 1.12 (.7, 1.7) women OR .82 (.3, 2.1)
Ruder et al., 2001	1708 drycleaners exposed to PCE and possibly other solvents for at least 1 yr. prior to 1960 followed to 1996 (update of Brown and Kaplan and Ruder 994)	Retrospective cohort mortality study	SMRs PCE-only all sites 1.08 (.85, 1.36) tongue 9.03 (1.86, 26.39) esophageal 2.65 (.85,6.20) inatestinal 1.18 (.51, 2.33) rectal 0 pancreatic .8 (.17, 2.35) trachea, bronchus lung 1.17 (.71, 1.83) breast .78 (.28, 1.69) female genital 1.6 (.77, 2.95) cervix 1.89 (.52, 4.84) male genital .65 (.08, 2.35) kidney 1.73 (.21, 6.25) bladder 0 hemato 1.08 (.39, 2.36) SMRs, PCE+ all sites 1.35 (1.16, 1.55) tongue 3.04 (.37, 10.99) esophageal 2.40 (1.10, 4.56) inatestinal 1.63 (1.004, 2.42) rectal 2.16 (.86, 4.45) pancreatic 1.89 (1.06, 3.11) trachea, bronchus lung 1.46 (1.07, 1.95) breast .98 (.54, 1.65) female genital 1.24 (1.69, 2.04) cervix 1.98 (.85, 3.91) male genital 1.02 (.46, 1.93) kidney 1.27 (.26, 3.72) bladder 3.15 (1.51, 5.79) hemato .61 (.25, 1.26)
Ojavari et al., 2001	Pancreatic cancer literature 1969-1998 where chlorinated HC solvents are studied	Meta-analysis	Pancreatic cancer meta-relative risk 1.24 (.79, 1.97)
Hansen et al., 2001	803 Danish workers exposed to TCE, files of individual air and urine measures	Retrospective cohort	SIRs: NHL 3.5, esophagus 4.2, cervical 3.8 (all based on small numbers; no CI reported)

Costas et al. 2002	Published version of Mass DPH study; water model for exposure; 19 cases 37 matched controls	Case-control	Childhood leukemia OR 8.33 (.73, 94.67) for contaminated; statistically stable for exposure prior to birth
Morgan, Cassady 2002	Incident cases of 16 cancer types, 1988 – 1998, in California community with water contaminated with perchlorate and TCE	Cohort	SIRs, 99% CIs: all sites, .97 (.93, 1.02) thyroid, 1.0 (.63, 1.47) lung/bronchus .71 (.61, .81) colon/rectum, .86 (.74, .99) uterine, 1.35 (1.06, 1.7) melanoma, 1.42 (1.13, 1.77)
Bruning et al., 2003	Hospital based 134 renal cell cancers and 401 hospital controls w/o dementia or cancer, frequency matched by sex and age	Hospital based case-control study	Longest held job in TCE industry, OR 1.8 (1.01, 3.2) "metal degreasing" OR 5.6 (2.3, 13.2) narcotic sx OR 3.7 (1.8, 7.5)

*citations[106]

---

[106] McMichael A, Spirtas R, et al, "Solvent exposure and leukemia among rubber workers: an epidemiological study," *Journal of Occupational Medicine,* 17: 234-39,1975; Blair, A, Decoufle, P, Grauman, D. "Causes of death among laundry and drycleaning workers." *Am J Public Health* 69:508-511, 1979; Kaplan S, *Dry Cleaners Workers Exposed to Perchloroethylene – A Retrospective Cohort Mortality Study* NTIS PB81-231-367, 1980; Olsson H, Brandt L, "Occupational exposure to organic solvents and Hodgkin's disease in men. A case-referent study," *Scand J Work Environ Health* 6:302-305, 1980; Hardell L, Eriksson M, Lennert P, Lundgren E, "Malignant lymphoma and exposure to chemicals, especially organic solvents, chlorophenols and phenoxy acids: a case-control study," *Br J Cancer* :169-176, 1981; Katz R, Jowett D, "Female laundry and dry cleaning workers in Wisconsin: a mortality analysis," *Am J Pub Health,* 71: 305-307, 1981; Peters J, Preston-Martin S, Yu M, "Brain tumors in children and occupational exposure of parents," *Science* 213:235-2371981; Duh R, Asal N, "Mortality among laundry and drycleaning workers in Oklahoma," *Am J Pub Health* 74:1278-1280, 1984; Barret L, Faure J, Danel V, "Epidemiological study of cancer in a community of workers occupationally exposed to trichloroethylene and cutting oils (Abstract 34), Stockholm, Association Europeenne des Centres Anti-poison, 1984 [cited in IARC, vol. 63]; Hardell L, Bengtsson N, Jonsson U, Eriksson S, Larsson L, "Aetiological aspects on primary liver cancer with special regard to alcohol, organic solvents and acute intermittent prophyria—an epidemiological investigation," *Br J Cancer* 50:389-397,1984; Shindell S, Ulrich S, "A cohort study of employees of a manufacturing plant using trichloroethylene," *J Occup Med* 27:577-579, 1985; Brown D, Kaplan S *Retrospective cohort mortality study of dry cleaning workers using tetrachloroethylene,* NIOSH. Industry wide Studies Branch. Cincinnati, OH, 1985; Barret L, Faure J, Danel V, "Epidemiological study of cancer in a community of workers occupationally exposed to trichloroethylene and cutting oils," *J Toxicol Clin Toxicol* 23:438, 1985; Axelson, O, Epidemiological studies of workers with exposure to tri- and tetrachloroethylene. In: *New Concepts and Developments in Toxicology,* Chambers, PL, Gehring, P and Sakai, P, Eds., Elsevier, 1986, pp. 223-230; Lowengart R, Peters J Cicioni C, et al, "Childhood leukemia and parents' occupational and home exposures," *JNCI,* 79: 39-46,1987; Garabrant D, Held J, Langholz B, Bernsn L, "Mortality of aircraft manufacturing workers in southern California," *Am J Ind Med* 13:683-693, 1988; Blair A, Stevens Tolbert P, Grauman D, Moran F, Vaught J, Rayner J. "Cancer and other causes of death among a cohort of dry cleaners." *Br J Ind Med.* 47: 162-168,1990; Hernberg S, Kauppinen T, Riala R, Korkala M, Asikainen U, "Increased risk for primary liver cancer among women exposed to solvents," *Scand J Work Environment Health* 14:356-365, 1988; Silverman D, Levin L, et al. "Occupational risks of bladder cancer in the United States: I. White Men," *JNCI,* 81:1472-1483, 1989; Sharpe C, Rochon J, Adam J, Suissa S, "Case-control study of hydrocarbon exposures in patients with renal cell carcinoma," *Can Assoc J* 140: 1309-1318, 1989; Fredricksson M, Bengtsson N, et al., "Colon cancer, physical activity, and occupational exposures," *Cancer* 9:1838-1842, 1989; Olsen G, Hearn S, Cook R, Currier M, Allen S, "Mortality experience of a cohort of Louisiana chemical workers," *J Occup Med* 31:32-34, 1989; Blair A, Stewart P, Tolbert P, Grauman D, Moran F, Vaught J, Rayner J, "Cancer and other causes of death among a cohort of dry cleaners," *Br J Ind Med* 47:162-168, 1990; Bond G, McLaren E, Sabel F, Bodner K, Lipps T, Cook R, "Liver and biliary tract cancer among chemical workers," *Am J Ind Med* 18:19-24, 1990; Lynge E, Thygesen L, "Primary liver cancer among women in laundry and dry-cleaning work in Denmark," *Scand J Work Environment Health* 16:108-112, 1990; Fagliano J, Berry M, Bove F, Burke T, "Drinking water contamination and the incidence of leukemia: an ecologic study," Am J Public Health 80:1209-1212, 1990; Spirtas R, Stewart, Lee J, Marano D, Forbes C, Grauman D, Pettigrew H, Blair A, Hoover R, Cohen J, "Retrospective cohort mortality study of workers at an aircraft maintenance facility. I. Epidemiological results," *Br J Indust. Med.* 48:515-530, 1991; Aschengrau, A., Ozonoff, D. Paulu, C., Coogan, P. Vezina, R., Heeren, T., Zhang, Y., "Cancer risk and

There are *also* numerous studies of cancer risk among workers who have a history of exposure to solvents, among which PCE and TCE are extremely common. Because such work usually involves mixtures, it does not unequivocally implicate PCE or TCE as the *sole* causes of the elevated risks, but it is clearly pertinent in that these are PCE and TCE exposed workers.[107] As already noted, it is not valid to consider any one factor as a *sole* cause in a multifactorial disease, but in any event, my opinion is not determined by the more non-specific studies of solvent-exposed workers. These studies, however, add further to the weight of the evidence.

ii. Evaluating the strengths and weaknesses of the epidemiological studies

I have already discussed the importance of evaluating internal and external validity in epidemiological studies. The results of evaluating these qualities of a given study influence the weight given to a study when a scientist "assembles the picture" integral to a judgment of causality. I have discussed, too, how my Boston University colleagues and I carefully train

---

tetrachloroethylene (PCE) contaminated drinking water in Massachusetts," *Archives of Environmental Health*, 48:284-292, 1993; Varianinen T, Pukkala E, Rienoja T, Strandman T, Kaksonen D, "Population exposure to tri- and tetrachloroethene and cancer risk: two cases of drinking water pollution," *Chemosphere* 27:1171-1181, 1993; Cohn P, Klotz J, Bove F, Berkowitz M, Fagliano J, "Drinking water contamination and the incidence of leukemia and non-Hodgkin's lymphoma," *Environ Health Perspectives* 102:556-561, 1994; Anttila A, Pukkala E, Sallmen M, Hernberg S, Hemminki K, "Cancer incidence among Finnish workers exposed to halogenated hydrocarbons," *J Occup Med* 37:797-806, 1995; Axelson O, Andersson K, Selden A, Hogstedt C, "Updated and expanded Swedish cohort study on trichloroethylene and cancer risk," *J Occup Med* 36:556-562, 1994; Ruder A, Ward E, Brown D, "Cancer mortality in female and male dry-cleaning workers," *J Occup Med* 36:867-874, 1994; Henschler D, Vamvakas S, Lammert M, Dekant W, Kraus B, Thomas B, Ulm K, "Increased incidence of renal cell tumors in a cohort of cardboard workers exposed to trichloroethene. " *Arch Toxicol.* 69:291-299, 1995;; Morgan R, Kelsh M, Zhao K, Heringer S "Mortality of aerospace workers exposed to trichloroethylene," *Epidemiology* 9:424-431, 1998; Blair A Hartge P, Stewart PA, McAdams M, Lubin J, "Mortality and cancer incidence of aircraft maintenance workers exposed to trichloroethylene and other organic solvents and chemicals: extended follow up," Occup Environ Med 55:161-171, 1998; Aschengrau, A, Paulu C, Ozonoff D, "Tetracholoroethylene-contaminated drinking water and the risk of breast cancer, " Environ Health Perspect, 106(suppl4):947-953, 1998; Vamvakas S, Bruning T, Thomasson B, Lammert M, Baumuller A, Bolt HM, Dekant W, Birner G, Henschler D, Ulm K, "Renal cell cancer correlated with occupational exposure to trichloroethene," J Cancer Res Clin Oncol 124:374-382, 1998; Paulu C, Aschengrau A, Ozonoff D, "Tetrachloroethylene-contaminated drinking water in Massachusetts and the risk of colon-rectum, long, and other cancers," Environmental Health Perspectives, 107:265-271, 1999; Boice JD, Marano DE, Fryzek JP, Sadler CJ, McLaughlin JK. "Mortality among aircraft manufacturing workers," Occup Environ Med 56:581-597, 1999; Dosemici M, Cocco P, Chow W-H, "Gender differences in risk of renal cell carcinoma and occupational exposures to chlorinated aliphatic hydrocarbons," Am J Ind Med 36:54-59, 1999.

[107] See Zoloth S, Michaels D, Villalbi J, Lacher M, "Patterns of mortality among commercial pressmen," *JNCI* 76:1047-1051, 1986; Spiegelman D, Wegman D, "Occupation-related risks for colorectal cancer," *JNCI* 75:813-821, 1985; Hernberg S, Kauppinen T, Riala R, Korkala M, Asikainen U, "Increased risk for primary liver cancer among women exposed to solvents," *Scand J Work Environ Health* 14:356-365, 1988.

student epidemiologists to do this evaluation of internal and external validity. I present here my evaluations, doing so first exhaustively, to illustrate the full method, then in the more usual way of a practicing professional epidemiologist.

I will begin with a study by Boice et al.,[108] performed by defendants to defend themselves in toxic tort litigation, but published in the scientific literature. It purports to show that exposure to TCE in the workplace and the community is without risk to those exposed.

a. My evaluation of the internal and external validity of the Boice study.

The Boice study is a study of some 78,000 workers engaged in the manufacture of aircraft at the Burbank facility of the Lockheed-Martin Corporation. In addition to carefully reviewing the Boice study itself, I also have examined two other documents that underlie the Boice study: (a) the exposure assessment, by Marano, et al.., and (b) a portion of the backup material prepared prior to onset of the study (the so-called "feasibility study"), which Boice and his colleagues used to demonstrate to Lockheed's lawyers their ability to undertake the study Lockheed desired. While this study and its ancillary materials was performed specifically with litigation in mind, it has been published in the open scientific literature and therefore forms part of the material upon which I relied for this case.

As noted above, evaluation of scientific evidence for causation involves, among other things, an assessment of existing publications for validity, both internal and external.

Evaluation of the internal validity of the Boice study

Study type: Retrospective cohort mortality study.

Study objectives: This study appears to be litigation-driven, since the main exposures of interest were just those involved in the state and federal lawsuits against Lockheed and it was the

---

[108] Boice JD, Marano DE, Fryzek JP, Sadler CJ, McLaughlin JK. "Mortality among aircraft manufacturing workers," Occup Environ Med 56:581-597, 1999

company, through its legal counsel, that arranged to have it performed. The authors did not state if Lockheed's lawyers sought and were allowed prior "review" of the study, and if so, how the final paper differed from drafts submitted before review. Oddly, although the stated study objective was to evaluate the ***risk of contracting cancer*** and other diseases among aircraft manufacturing workers potentially exposed to chromate, TCE, PCE, and mixed solvents, the study did not evaluate the risk of contracting cancer but rather ***the risk of dying from cancer***, which, as discussed below, is a related but not identical matter.

Outcome of interest: The outcome of interest was stated to be risk of cancer, but the outcome studied was death from cancer. Because death from cancer is composed of two different components (the risk of cancer plus the risk of dying from that cancer), the study failed to accurately measure the authors' stated outcome of interest. Documents from the proposal and feasibility study suggest the authors were interested in pursuing cancer incidence (a better measure of risk), but for some reason that was not explained, did not do so. It should be noted, too, that SMRs (measures of risk of death) are usually lower than SIRs (measures of risk of contracting cancers: adjusted incidence rates) for the same diseases.

Even within the outcome that the authors actually studied (mortality), there are some unanswered questions about the accuracy of the outcome (cause-specific mortality). Cause (and cancer site) specific mortality was obtained from death registration files or death certificates. Scientists recognize that these sources are not completely accurate, and the accuracy varies with cause of death. Thus scientists recognize that although the accuracy for cancer is better than for some other causes, there is considerable variation between different cancer sites, and, moreover, for certain treatable cancers (e.g., some leukemias, kidney and bladder cancers, thyroid cancer) death certificates may only poorly reflect a history of cancer. Although this also affects the

comparison group, *i.e.*, persons who did not die of cancer in the general comparison population the result would tend to blunt, understate, and underestimate any true effects of exposure (non-differential outcome misclassification).

In addition, the authors failed to supply any information regarding the source and quality of medical care for the Lockheed employees. If their care was in some sense comparable to the general population this might make little difference. But if it were different in some way (e.g., through a private insurer paid by the company as part of a collective bargaining agreement with the worker's union) the documentation and interpretation of medical results might be skewed in one direction or another. With fewer than 11% of all deaths nationwide now being autopsied, the judgment of medical care providers becomes even more important in classifying a cause of death.

Primary exposure of interest: The authors of the Boice study were primarily interested in TCE, PCE, chromates, and mixed solvents, although there was exposure to many other chemicals at the plant. The authors' principal difficulty resided in how to determine who was exposed to what and when. To fulfill this requirement they devised an elaborate protocol to classify which workers had been exposed to one or another chemical – routinely, intermittently, or not at all.

The question here is not whether a credible effort was made to devise an elaborate protocol (it appears that it was), but whether in the circumstances of the Burbank plant, where workers changed jobs frequently, where jobs themselves were altered frequently, and where workers frequently suffered from multiple exposures to many different agents, it is possible to devise an exposure assessment scheme that would not result in a crippling – and disqualifying – degree of misclassification, misclassifications that could completely mask any true associations.

My concerns over the adequacy of the exposure assessment stem from two sources.

First, the Boice strategy of trying to determine homogeneous exposure groupings, which underlies their scheme, is known to result in a great deal of potential misclassification (*i.e.*,, to result in non-uniform exposures to members of the supposedly homogeneous group).[109] Substantial within- and between-worker variability were manifest in Kromhout's dataset (much of which was derived from the more homogeneous work setting of the chemical industry), despite the fact that Kromhout's classifications were based on considerably more information than available to Boice. Boice's claim that they had reproducible and accurate exposure classifications (Report on Feasibility, p. 10), seems to be rather exaggerated since they have no method to verify the exposures. Significantly, the authors failed to document this claim.

Second, despite Boice's repeated descriptions of how he and his colleagues clarified exposure classifications, it is still not clear how "estimated potential for exposure" to TCE/PCE was determined for job titles. Lockheed apparently made little or no exposure data available to Boice, and the plant was not operating and empty of workers at the time he conducted his study. I believe that it is simply insufficient to give the sources of information without a description of how those sources were used, particularly when the study to be used by Lockheed was solicited and funded by Lockheed, and would not have been undertaken (at least not by Boice) "but for" Lockheed's payment of Boice's fee. Although job "families" were determined *a priori* by Lockheed (which clearly had a stake in the study and thus a stake on how data was arranged and classified), Boice failed to document the basis and accuracy of these "families" for the purpose of making even gross estimates of exposure. Although members of Boice's own team assigned

---

[109] *See, for example,* Kromhout H, Symanski E, Rappaport S, "A comprehensive evaluation of within- and between- worker components of occupational exposure to chemical agents," Ann. Occup. Hyg. 37:253-270, 1993: "Unfortunately, it seems impossible to predict which groups, based on job title and factory, are more-or-less homogeneously exposed.").

historical job titles to job families, Boice did not say whether these assignments were tested for reproducibility between or within team member judgments.

Boice reported that Lockheed conducted industrial hygiene walk-through inspections of the Burbank facility with "knowledgeable former Burbank employees" to indicate locations of departments and process equipment lines for each location. Boice failed to report, however, who these employees were (were they hourly employees or management, for example) and the extent of their true knowledge of what went on at various locations; indeed, Boice failed to state whether he asked these most elementary questions. To the extent that this knowledge varied, there is additional potential for exposure misclassification.

Boice also failed to report factory "floor plans" and "chemical usage patterns" were translated into job exposure potentials. Boice reported that he conducted interviews with selected long-term employees, but he failed to say who these employees were, who made the selection and how, who arranged for access to these former employees (and for the former employees' access to Burbank facilities), and what their true knowledge might have been.

Marano stated that the exposure metric was the "length of time spent in jobs with potential exposure to the chemical" (Marano ms, p. 2), but it is not clear if this means that Marano counted (1) only time spent in jobs when the chemical was used is counted (e.g., prior to 1966 for TCE) or (2) all time spent in the job. For example, for a Process Operator/Plater who started employment after 1966, presumably *no* person years of follow-up would be counted under "TCE worker" in table 8, but Marano (and Boice) failed to explain whether this was done, and, if so, how this was done. I would assume dermal and inhalation exposures were both included, although it is known (see Kromhout) that dermal exposures show a great deal of

between-worker variability. Marano and Boice also failed to explain how person-years of follow-up were calculated for various occupational classes thought to involve TCE exposure.

Boice reported that information of job histories came from Lockheed's collection of employee "Kardex cards" and that these cards were complete. But Boice also revealed that computerized retirement information was used "to supplement" and confirm information found on the Kardex cards (Marano ms., p. 5). This leads to a question that Lockheed and Boice have never answered: if the cards are the "gold standard" of employee information, why was supplementation used? And how were payroll data listings used to inform the job history database?

Although the number of workers studied and followed-up by Boice is comparatively large (relative to other studies), the number of workers even putatively exposed to TCE and PCE is not large. Only 12% (TCE) and 13% (PCE) of the *factory* workers were judged to have been exposed to those chemicals "routinely" or even "intermittently," and the numbers for "routine" exposure to these chemicals were only 5% and 6%, respectively. The latter figures are roughly the same as the proportion of factory workers (4.3%) for whom duration of exposure to solvents could not be estimated from the records. It is not clear if the 4.3% is part of the 5% (or 12%) or in addition to them.

Only 31,000 of the total cohort of 78,000 (and factory subcohort of 45,000) were in the eight job families that Boice and Marano judged were most likely to entail highest potential for significant exposures to TCE and PCE. (This assumes that the breakdown given in table III of the Marano ms. has no overlaps, *i.e.*,, workers in more than one category; with overlap, fewer than 31,000 would be in the group with potential exposure.) Thus even if Boice confined himself to the factory subcohort described in table 3 of his report (for which he failed to provide any

confidence intervals), he was looking at a population which even he was compelled to admit had at least one third of its members not likely to have even *potential* for exposure. Thus the relevance of any part of table 3 is rather questionable.

Table III of the Marano ms. also purports to give the distribution of exposed workers by job family and exposure to TCE and PCE ("70%…were exposed to PCE"), but all that was determined was whether this was a job with potential for exposure to PCE, not whether any particular worker *was* exposed to PCE.

Imprecise exposure assessment, which is almost certainly present here, is well-recognized as the Achilles heel of most environmental and occupational epidemiological studies. In addition, Boice failed to supply sufficient information to allow an independent evaluation of the adequacy of exposure determination. Because even moderate amounts of non-differential misclassification of exposure will dramatically reduce estimates of true effects, failure to find effects in the one comparison of relevance (Table 8), does not support the opinion that TCE and PCE are without effect.

Study base and comparison group: The study base and comparison group consisted of workers with at least one year tenure at the LM Burbank plant on or after 1/1/60. Thus there were two possible sources of both selection and "late entry" bias. Boice stated that he analyzed data to account for the better mortality experience of newly hired workers and to account for latency (a 10 year lag), but he failed to present that data. Boice stated that the data made "little difference" in SMRs for the full dataset. If there were no difference, I would think the methodologically less impeachable data would have been preferred. The reason Boice gave for using this sort of data (including the entire cohort allowed comparison with the aircraft studies in the literature) are neither relevant nor persuasive, because the comparisons given in table 10 are

between the total cohorts of each of the studies, not the factory sub-cohorts. Furthermore, Boice's habit of diluting the cohorts with office and administrative workers – *i.e.*, with employees who had no potential for exposure – is not an informative procedure for the claimed objective of evaluating the effects of factory exposures. As a result, Boice's comparisons in the table must be judged to be completely worthless.

Most importantly, there is substantial evidence for a general downward bias in the SMRs in the Boice study. This can be seen in two ways. Both the "all cause" and "all cancer" SMRs are significantly below 1.00 for both the total cohort and the factory subcohort. Because no one has ventured to claim a "chemoprotective effect" for any of the chemicals in the plant, *i.e.*,, that industrial pollutants miraculously improve the health of the people who ingest them, this is evidence of bias (chance being unlikely). Whatever the source of the bias, it seems to operate on cancer and non-cancer effects alike, and is not confined to only a few causes of death. Thus for the 25 cancer SMRs reported in table 3 among factory workers (the only subcohort of any relevance here), only 7 have SMRs at one or above. If there is no effect of exposure, one would expect the probability of being above or below one to be 0.5. A calculation of the likelihood of there being "no bias" when 18 of 25 results are below 1.00 does not support the contention of "no bias." The same is true of the Morgan and Garabrant studies, both studies of aircraft manufacturing workers (see below). Thus it is necessary to look for a source of downward bias common to these studies.

One finds it most plausibly in the Healthy Worker Effect (HWE). The HWE is clearly evident in the non-cancer effects. One might think it less likely to affect cancer mortality, but as noted above, a bias is clearly evident in the results. One explanation is the special nature of being selected into manufacturing work on military aircraft, where it is possible (and for many

jobs likely) that additional screens for lifestyles that might also be correlated with unfavorable mortality outcome would be applied (e.g., drinking, drug use, or other lifestyle choices that might pose a security risk). In any event, it is clear that something pushed all the effect levels downward, to an extent not determinable from the data as presented.

Confounding: Boice controlled for only four confounders or effect modifiers: age, race, sex and calendar year. Although these are important to control, Boice made no attempt to control for other important confounders, confounders that might mask (or produce) associations, for example smoking. Although Boice acknowledged the importance of controlling confounders in his feasibility proposal, he made no attempt to do so in the main study.

Stability of measures of association: The confidence intervals for various sites in table 8 are wide, which substantially reduces the informativeness and relevance of the results that Boice reports. This is due primarily to the small number of workers exposed. It should also be noted that calculation of confidence intervals assumes there is no bias. In as much as we know some downward bias exists, the confidence intervals themselves are biased. The most apt comparisons (internal comparisons) had too few observations to be informative (according to the authors, and I agree) for routine exposures to TCE or PCE.

Interpretation and external validity (generalizedability) of the Boice study

At best the Boice study is non-positive with respect to **cancer mortality**, not **cancer risk**. As it is, even the study's conclusions about mortality are not warranted.

Although it is true that in this study both the "all cause" and cancer mortality rates are mostly below national population norms, this is readily explained as the result of bias, most likely selection bias. The only alternative is to posit a chemoprotective effect for the exposures that occur at the Burbank facilities, something which neither Boice nor anyone else has had the

temerity to broach. Thus the results provide evidence that impeach the credibility of the study, not the credibility of propositions about the relationship of exposures to TCE and PCE and cancer risk.

The same caveats concerning bias apply with respect to total and cancer mortality for Burbank employees resulting from exposure to TCE and PCE. Additional caveats apply to the almost certain non-differential misclassification of exposure, misclassifications that would dilute or mask any real effect. The relatively small proportion of workers exposed to TCE and PCE in this workforce make any estimate of effect, even in the absence of bias, tenuous. Indeed, the fact that only a portion of each worker group in table 7 was exposed to TCE or PCE, essentially turns this study into an "ecologic design." Thus questions of between group confounding or effect modification become even more pressing and the results even more difficult to interpret. It is significant, however, that one of the cancer sites identified by IARC, non- Hodgkin's lymphoma, shows up as elevated in the job title with the highest proportion of TCE and PCE exposure (Fabrication, table 7), even in the face of almost certain downward bias from selection and misclassification.

The problems that are endemic to the aircraft industry in the Boice study are also evident in other studies of aircraft manufacturing/maintenance (Spirtas/Blair, Garabrant and Morgan), where selection and exposure misclassification bias the results (downwards or to the null).

The insensitivity and downward bias (not to mention lack of generalizability) of the Boice results is shown by its inability to detect an increased rate of lung cancer in the asbestos-exposed subcohort. If we were to depend upon the Boice study to signal the carcinogenicity of asbestos (a strong carcinogen) we would still be allowing indiscriminate exposure at work and in the environment.

Finally, the generalization that a lack of showing of cancer at LM Burbank logically implies a lack of risk to populations in the neighborhood of the LM facilities is problematic. The workforce at LM Burbank is not representative of that population, nor can risks that would be of significant public health impact on the surrounding large population of all segments of the community be remotely detectable in the Boice study. The workers that lived in Burbank were subject to the same selection biases as the ones remaining at work.

My review of other, independent, non-litigation driven epidemiological studies on TCE and PCE

There is a rather large epidemiological literature on TCE and PCE. A survey of this literature is in Wartenberg et al.[110] In addition to the Boice study reviewed above, a number of the older studies have been updated and some new populations of workers exposed to PCE and TCE have been studied, specifically, the update of the Axelson study, published in 1994, the study by Spirtas et al. of 1991, the study of Henschler of 1995, the Anttila et al. study of 1995, the Blair update of 1998, and the Morgan study of 1998. IARC took special note of the studies by Axelson, Spirtas and Anttila, while the studies of Blair and Morgan were published after the IARC Monograph was published. In addition, Ruder et al. published two updates of the Kaplan (1980) and Brown and Kaplan (1985) studies of dry cleaners, and Blair et al. reported in 1990 on a follow-up of an earlier union cohort of drycleaning workers.

I briefly discuss each of these studies. I organize the discussion under several broad headings: studies of aircraft manufacturing and maintenance workers; studies of workers

---

110 Wartenberg D, Reyner D, Siegel C, "Trichloroethylene and Cancer: Epidemiologic Evidence," *Environ Health Perspect* May;108 Suppl 2:161-176, 2000: "Evidence of excess cancer incidence among occupational cohorts with the most rigorous exposure assessment is found for kidney cancer (relative risk [RR] = 1.7, 95% confidence interval [CI] 1.1-2.7), liver cancer (RR = 1.9, 95% CI (1.0-3.4), and non Hodgkin's lymphoma (RR = 1.5, 95% CI 0.9-2.3) as well as for cervical cancer, Hodgkin's disease, and multiple myeloma."

measured with biomarkers of exposure; studies of dry cleaning workers; German studies of kidney cancer; and environmental studies of drinking water contaminated with TCE and/or PCE.

Other epidemiological studies of aircraft manufacturing and maintenance workers

Lockheed's Boice study is one of four occupational cohort studies in the literature of aircraft manufacturing or maintenance workers, although it is the only one I know of that was solicited and completely financed by a party in the midst of ongoing litigation. All the studies show evidence of significant selection bias (healthy worker effect and selection for unusually low risk workers). All are mortality studies except for some portions of the Spirtas/Blair cohort, described below. Use of mortality also lowers observed risks. All of the studies also suffer from significant exposure misclassification, further blunting, diluting, or masking any risks that may be actually present. One of the studies, by Garabrant et al.,[111] is uninformative and will not be considered further.

Spirtas, et al.[112] – This cohort study consisted of workers who maintained and overhauled aircraft and missiles. Some 7200+ civilian workers working in a 5-year window (1952 – 1956) at Hill Air Force Base in Utah were classified as having been exposed to TCE, although exposure levels were not quantified[113]. Utah State rates were used for comparison and follow-up was through 1982 (i.e., maximum 26 year latency). The kinds of selection bias noted for the Boice study (a workforce with less than usual lifestyle risk factors because of the nature

---

[111] Garabrant D, Held J, Langholz B, Bernsn L, "Mortality of aircraft manufacturing workers in southern California," Am J Ind Med 13:683-693, 1988;

[112] Spirtas R, Stewart PA, Lee JS, Marano DE, Forbes CD, Grauman DF, Pettigrew HM, Blair A, Hoover RN, Cohen JL, "Retrospective cohort mortality study of workers at an aircraft maintenance facility. I. Epidemiological results," *Br J Indust Med* 48:515-530, 1991.

[113] An attempt was made to estimate relative exposure on the basis of job title. See articles by Kromhout, et al. and Rappaport et al., op. cit., note 84, for the misclassifications likely to result from this method.

of the employment) would also be expected to bias the results downward for Spirtas's study. A relatively stable excess for biliary tract cancer deaths was seen, and less stable excesses for bone cancer in men; cervical cancer, and non-Hodgkin's lymphoma in women also were seen. This study differs from the Axelson study and the Henschler study (discussed below) in studying mortality (deaths) rather than cancer incidence. This makes a quantitative comparison of its results with those studies less commensurable and less interpretable.

There was evidence in Spirtas' tables 6 and 7 of an increased risk for Multiple Myeloma (MM) and NHL to both men and women exposed to solvents as a class, including TCE:

Spirtas tbls 6-7	hi	low	SMR	obs	chemical
MM	2.8	0.78	1.56	11	any solvent, men
MM	2.8	0.78	1.57	11	mixed solvent, men
MM	2.6	0.36	1.11	5	TCE, men
MM	6.4	0.45	2.2	3	any solvent, women
MM	6.5	0.46	2.23	3	mixed solvent, women
MM	7.2	0.03	1.3	1	TCE, women
NHL	1.9	0.72	1.21	18	any solvent, men
NHL	1.9	0.72	1.22	18	mixed solvent, men

NHL	1.9	0.49	1.03	10	TCE, men
NHL	5.8	1.13	2.82	7	any solvent, women
NHL	5.6	1.15	2.86	7	mixed solvent, women
NHL	7.3	0.78	2.86	4	TCE, women

Although the exposures here are generally to mixed solvents or unspecified solvents, Spirtas did remark that elevated SMRs for MM and NHL were seen for PCE exposures in women (SMR 17.05, CI 2.06-61.59 for MM, SMR 9.68, CI 1.17-34.96 for NHL). Although these findings could have other explanations, Spirtas et al.. comment that "the associations between these tumors and chemicals such as carbon tetrachloride and perchloroethylene that cause cancer in laboratory animals, plus similarities to other epidemiological investigations that have noted associations between various solvent exposures and risks of lymphatic and hematopoietic neoplasms, provide a biological plausibility which, we believe, does not allow these findings to be clearly dismissed as chance occurrences."

A paper by **Blair** [114] et al. reports an update of the Spirtas et al. cohort study. The definition used by Spirtas and Blair of "TCE exposure" is of an individual who ever held a job "in which exposure to TCE *may* have occurred." (Emphasis added). Thus, some of the jobs so classified may **not** have had any TCE exposure, and even for those that did, holders of those jobs may never have been exposed to TCE (indeed, TCE use ceased in 1966). Thus considerable

---

[114] Blair A, Hartge P, Stewart PA, McAdams M, Lubin J. "Mortality and cancer incidence of aircraft maintenance workers exposed to trichloroethylene and other organic solvents and chemicals: extended follow up," *Occup Environ Med* 55:161-171, 1998. This is a follow up to Spirtas et al., discussed above.

exposure misclassification is almost certain, which would tend to reduce any estimate of a real risk that might have been present.

The estimates of NHL are increased (RR = 2) in the TCE sub-cohort (as are estimates for cancer of the esophagus, colon, liver breast, kidney, cervix and bone), but the estimates of the increases are not statistically stable, affect men and women differently, and are not clearly related to levels of exposure. In particular, kidney cancer risk was increased for women, with the two reported cases both being in the high exposure category (RR = 3.6, C.I. .5 – 25.6), or alternatively in the group experiencing frequent peak exposures (RR = 5.7, C.I. .5 – 63.3). The considerable width of the confidence intervals indicates the relatively unstable nature of these estimates. No similar risk patterns were evident among men, but again with such unstable estimates any existing monotonic pattern could be obscured.

On their face, the Spirtas/Blair data suggest an increased risk of kidney cancer for women but not for men. A similar result has been reported by **Dosemici et al.**[115], where the gender difference in kidney cancer risk from TCE was highlighted. In this study 796 newly diagnosed kidney cancer patients (all white) were identified in the population-based Minnesota cancer registry in the period 1988 – 1990, and compared via interview methods with 707 population-based controls stratified by age and gender. Response rates (86% and 87%) were good and comparable in each group. Occupational histories were obtained (along with a complete demographic profile and set of confounders) and those thought exposed to TCE or PCE were identified by means of a National Cancer Institute Job Exposure Matrix previously used by the investigators, who are from the National Institutes of Occupational Safety and Health (NIOSH). Analysis was via multivariate logistic regression, with the ORs used as an estimate for the RR,

---

[115] Dosemici M, Cocco P, Chow W-H, "Gender differences in risk of renal cell carcinoma and occupational exposures to chlorinated aliphatic hydrocarbons," *Am J Ind Med* 36:54-59, 1999.

controlling for age, smoking, hypertension status and/or use of diuretics and/or anti-hypertension drugs, and body mass index (gender specific).

As expected, males were much more frequently exposed to organic solvents (34% versus 11%), but the difference between the cases and controls was much greater for females than males, with an OR = 1.96 (1.0, 4.0) for women compared to an OR = 1.04 (.6, 1.7) for men. No similar risk difference by gender was seen for those giving a history of exposure to PCE, although differences were seen for some other solvents. The authors point out that previous occupational studies of kidney cancer tended to focus on men, although a few have included women, including the Ruder et al.. study of dry-cleaners, discussed below, and two other studies that mention dry-cleaners (Asal et al., 1988, and Mellemgaard et al., 1994).[116] I quote the authors' comments on this matter:

> Although in some studies, risks were not significant due to the small number of women in the study population, women consistently showed higher RCC [kidney cancer] risk than men for the same exposures. Asal et al. (1988) and Ruder et al. (1994) reported almost 3-fold risk differences between men and women who worked in the dry-cleaning industry, in which various chlorinated aliphatic hydrocarbons, e.g., carbon tetrachloride, TCE, PCE, and 1,1,1 trichloroethane) have been used since the 1930s (ref. to IARC). Mellemgaard et al. (1994) reported more than 4-fold significant risk differences between men and women exposed to solvents in general.[117]

Dosemici et al. also discuss the biological mechanisms that might produce higher risk among women than men, which has been fairly consistently observed. I note that any

---

[116] Asal NR, Geyer JR, Risser DR, Lee ET, Kadamani S, Cheng N, "Risk factors in renal cell carcinoma. II. Medical history, occupation, multivariate analysis, and conclusions," *Cancer Detect Prev* 13:263-279, 1988; Mellengaard A, Olsen JH, McLaughlin JK, Engholm G, "Occupational risk factors for renal-cell carcinoma in Denmark," *Scand J Work Environ Health* 20:160-165, 1994.

[117] Dosemici et al. also note drily that Dr. McLaughlin's published review of kidney cancer and TCE exposure does not note that "[i]n some of the articles reported in this review, women employed in the dry-cleaning industry showed excess risk of [kidney cancer]," and citing four studies.

discounting of the TCE or PCE kidney cancer relationship on the basis that it is not the same for men and women is shown here to be invalid reasoning.

Morgan et al.[118] – The study by Morgan et al. examined mortality in a cohort of 20,508 workers at a Hughes Aircraft manufacturing plant in Arizona. Any worker employed more than 6 months between 1950 and 1985 was identified from company records and the vital status of each worker (whether they were alive or dead) was determined using Social Security Administration (SSA) files and, after 1979, the National Death Index (NDI). The published paper does not indicate use of any other means of follow-up. However, the unpublished earlier report by Wong and Morgan indicates that the usual follow up means were used. It is unclear why the published report does not take note of this, but I assume that the earlier (unpublished) method was followed in the updated study as well.[119] The study found an excess of ovarian cancer and elevated relative risks for cancers of the kidney, bladder and prostate, but did not find an association between TCE and lung or liver cancers or blood cancers, such as NHL.

Morgan's study is a study of aircraft manufacturing workers, like the Garabrant and Boice studies. Like those studies, Morgan shows evidence of severe selection bias. Indeed if we

---

[118] Morgan R, Kelsh M, Zhao K, Heringer S, "Mortality of aerospace workers exposed to trichloroethylene," *Epidemiology* 9:424-431, 1998.

[119] The method of using only SSA files to ascertain vital status as given in the published version of the Morgan study, can lead to serious undercounting. The National Death Index (NDI) provides a fairly complete record of deaths after 1979, but did not exist before then. In earlier years SSA files were routinely and successfully used for these types of studies, but in 1988 the Social Security Administration replaced its old system with a new Death Master File (DMF) for use in searching prior to 1979. Schnorr and Steenland compared the DMF results with a list of known decedents from seven previously established cohorts. Only 75% of the known deaths in the cohort were found successfully by the DMF, with the best results in the 1975-1979 period (89% – 95%). Reanalyzing two cohorts by excluding deaths not found by the DMF resulted in a 20% - 35% decrease in the SMRs and dose-response trends. Their conclusion was that the DMF is inadequate for vital status determination in cohort mortality studies for any cohort with a substantial number of deaths prior to 1979. Roughly two-thirds of the follow-up time for the Morgan et al. study was before this date. The undercounting found by Schnorr and Steenland during some of this period is severe: For the period 1950-54, only 1.7% of deaths were found in the DMF; for 1955-59 only 3.7%; from 1960-64 only 20.6%, from 1965-69 only 53.1%, and even as late as 1970-74, only 72.7%. Over the entire follow-up period, to 1991 the undercount by DMF is about 38%.

look at the three cohorts that examined cancer mortality in aircraft manufacturing workers we find that for 54 independent cancer mortality estimates only 10 in all three studies were above 1.0. If this were an unbiased study in which work exposures had no effect on mortality we would expect about half to be above 1.0 and half below 1.0. The observed split (10 above and 44 below) is highly unlikely. It would be as if a fair coin were tossed 54 times and came up tails 44 times and heads 10 times. The observed results suggest either that the coin was not fair (there was a bias) or indeed something at work is chemoprotective, again, a claim no one has ever made. Most epidemiologists would think selection bias was the most plausible explanation (the so-called Healthy Worker Effect or HWE).

Because of this selection bias, the most appropriate comparisons in the Morgan study is an "internal comparison," which was done using a Cox proportional hazards model. Here are the results for the comparison of peak low and no TCE exposures versus peak medium and high exposures:

Morgan, Cox tbls 3-5	upper limit	lower limit	RR	exposed	not exposed	contrast
all CA	1.24	0.9	1.06	177	923	peak hi vs low
bladder	3.81	0.52	1.41	5	18	peak hi vs low
hemat	1.82	0.64	1.08	17	90	peak hi vs low
kidney	4.23	0.85	1.89	8	24	peak hi vs

						low
leukemia	2.49	0.49	1.1	7	35	peak hi vs low
liver	3.35	0.29	0.98	3	17	peak hi vs low
lung	1.4	0.82	1.07	64	324	peak hi vs low
NHL	6.08	0.28	1.31	2	9	peak hi vs low
ovarian	8.99	0.84	2.74	4	9	peak hi vs low
prostate	2.55	0.85	1.47	16	60	peak hi vs low

In every case except for liver cancer there is an increase in risk with high TCE exposure in this cohort, *i.e.*,, evidence of a dose-response relationship. In particular, there is a RR of 1.89 (.85, 4.23) for kidney cancer in the internal comparison. Increased RR in the blood cancers are slight (with the highest being for NHL, RR = 1.31, C.I. .28 – 6.1) and not impressive. The downward bias that affects the other estimates also affects these, suggesting the actual risks may be considerably higher (and well above 2).

Epidemiological studies of workers with biological monitoring for exposure to TCE and PCE

Axelson et al.[120] – The Axelson ("Swedish") study published in 1994 is a second update of his original effort from 1978 (the first update was published in 1984). It uses methodology similar to the study of Finnish workers by Anttila et al., establishing a cohort of workers exposed to TCE by using existing records of urine tests given to monitor occupational exposures to chlorinated hydrocarbons (urinary trichloroacetic acid measurements, U-TCA). However, unlike the Anttila study (discussed below), the workers who were offered free U-TCA monitoring in Axelson were employees of *customers* of the TCE producer. There is no evidence to suggest that all workers who took advantage of this service were exposed to TCE, since it appears that any worker of such a customer could be monitored, regardless of exposure. U-TCA is also used for exposures to PCE and 1,1,1-trichloroethane, of which only PCE is sufficiently similar to TCE to put it in the same category of potential carcinogenicity. In fact, detectable U-TCA levels do not ensure that a worker was exposed to *any* of these chemicals. Studies cited in the IARC monograph on TCE show median U-TCA levels of 6 mg/L, with a range of 0.6 - 261 mg/L in *unexposed* individuals, comparable to the levels found in 80% of the study subjects in Axelson.[121]

---

[120] Axelson O, Selden A, Andersson K, Hogstedt C, "Updated and expanded Swedish cohort study on trichloroethylene and cancer risk," J Occup Med 36:556-562, 1994.

[121] Note that Axelson uses arithmetic means (always higher than the geometric mean often used here), while the 6 mg/L figure just cited in the IARC Monograph is a median, much less subject to highly skewed distributions such as those typical of these measurements. Thus the difference in published values for unexposed populations and the bulk of Axelson's study subjects is even less than apparent from these figures.

Moreover, there is almost certainly substantial exposure misclassification involved in the U-TCA exposure assessment.[122] This misclassification tends to dilute, blunt, or obscure existing differences between exposed and unexposed groups, obscuring any real relationships that might exist. Because of the lack of information about how the samples were taken and from whom, any exposure information (and certainly dose information) should be considered unreliable.

As noted elsewhere, this kind of problem almost always tends to reduce or mask any true effects that might be present. It should also be noted that Axelson's measure of "duration of exposure" is not actually exposure duration, but length of time from first *sample submission.* Some of these samples could have been submitted many years after beginning work, so that the measure of cumulative exposure is also in error, further biasing the risk results downward. The Anttila study (discussed below) does not have these defects, because monitoring was mandatory and only TCE exposed workers had samples submitted for U-TCA exposure. On the other hand, Anttila acknowledges that some employers even in this setting may have confused trichloroethane with trichloroethene and sent in the incorrect sample (urine instead of blood or vice versa), and even here misclassification is a possibility.

There is some evidence of a dose-response relationship in Axelson's tables (the data are for all malignant tumors):

---

[122] See recent work on this subject: Rappaport S, Kromhout H, Symanski E, "Variation of exposure between workers in homogeneous exposure groups," *Am Ind Hyg Assoc J* 54:654-662, 1993; Rappaport SM, Symanski E, Yager JW, Kupper L, "The relationship between environmental monitoring and biological markers in exposure assessment," *Environ Health Perspect* 103(Suppl 3):49-54, 1995; Kromhout H, Symanski E, Rappaport S, "A comprehensive evaluation of within- and between-worker components of occupational exposure to chemical agents," *Ann Occup Hyg* 17:253-270, 1993.

monitored cohort by mean U-TCA and exposure time, all CA, tbl 2	upper confidence limit	lower confidence limit	SMR	obs	time from 1st monitored sample
0-49	2.04	0.45	1.04	8	<2 yrs
100+	4.37	0.15	1.21	2	<2 yrs
0-49	0.9	0.37	0.56	20	>2 yrs
100+	2.8	0.2	0.96	3	>2 yrs

Here the contrast is between those with 0 – 49 mg/l U-TCA and 100+ mg/l. The middle category has been eliminated so as to reduce as much as possible the kind of misclassification the Axelson study suffers from. In this rendition, within each duration category (< 2 years and >2 years) there is an apparent dose-response gradient. The depressed SMRs are explainable from a healthy worker effect.

If we look at Axelson's table 5 which gives incidence data for selected sites in the monitored cohort, with the same contrast and men with >2 years exposure and at least 10 years latency, we see a dose response for all cancers combined and for NHL and prostate cancer. There is no apparent gradient for liver cancer and skin cancer. In all these instances the numbers are small and the stability of the estimate (as measured by the confidence intervals) is small (wide intervals), which could lead one to say the study itself is not very informative (*i.e.*,, it is compatible with a range of interpretations), but the data are certainly compatible (and show evidence of) a dose response effect.

Axelson table 5: monitored cohort SIRs >2 yrs from 1st sample and 10 yr latency	upper confidence limit	lower confidence limit	SIR	obs	exposure group
all CA	1.38	0.75	1.02	41	b - 0-49 mg/l
all CA	3.35	0.56	1.54	6	d - 100+ mg/l
liver	6.83	0.23	1.89	2	b - 0-49 mg/l
liver	35.52	0	0	0	d - 100+ mg/l
NHL	5.92	0.2	1.64	2	b - 0-49 mg/l
NHL	46.43	0.22	8.33	1	d - 100+ mg/l
prostate	2.33	0.65	1.3	11	b - 0-49 mg/l
prostate	8.67	0.29	2.4	2	d - 100+ mg/l
skin	8.52	1.19	3.65	5	b - 0-49 mg/l
skin	28.69	0.09	0	0	d - 100+ mg/l

Anttila et al.[123] – The Anttila study was an update and expansion of a 1980 study by Tola et al. [124] This study (the "Finnish" study) was roughly twice the size of Axelson's[125]. An "in press" version was made available to the IARC Working Group at the time of the latter's deliberations and formed part of their assessment. Like Axelson, Anttila studied workers who were monitored via U-TCA for exposure to three halogenated chemicals, one of which was

---

[123] Anttila A, Pukkala E, Sallmen M, Herberg S, Hemminki K, "Cancer incidence among Finnish workers exposed to halogenated hydrocarbons," *J Occup Med* 37:797-806,1995

[124] Tola S, Vilhunen R, Jarvinen E, Korkala M-L, "A cohort study on workers exposed to trichloroethylene," *J Occup Med* 22:737-740, 1980.

[125] 3089 men and women in Anttila versus 1421 men in Axelson.

TCE.[126] 208 cancers occurred in the TCE exposed group. Stable excesses were seen for cervical cancer, an excess that increased with exposure. There was also a stable six-fold excess in liver cancer, a stable three-fold excess for cancers of the blood system, a stable three fold excess for stomach cancer, and a stable three-fold excess for prostate cancer, all in the high exposure groups, and all after latency was taken into account.

The Anttila study investigated associations between TCE exposure and cancer incidence. The mandatory reporting of urine measures and its analysis by a single government laboratory made the exposure misclassification less severe than the Axelson study. However, as Anttila et al.. note, even here some misclassification was inevitable and would bias risk estimates downward if such risks existed. The use of cancer incidence rather than mortality is also an advantage over the aircraft manufacturing studies, all of which are mortality studies.

For a follow-up period of >20 years, Anttila notes that for the TCE –exposed cohort risks were:

> significantly increased for overall cancer as well as for cancer of the stomach, liver, prostate, and lymphatic and hematopoietic tissues combined (Table 3). The increase in the overall cancer incidence for the follow-up of >20 years was the same both in women and men and similar for the specific primary sites other than liver. Anttila, p. 800.

There was also evidence of a dose-response relationship for cervical cancer and cancers of the blood system, although the latter was not as strong. Interestingly, analysis of the same data for mortality did not show the same relationships, indicating that effects on cancer incidence may be obscured when mortality is the endpoint as it is in the aircraft manufacturing/maintenance studies. Attempts to analyze the data for the PCE exposed cohort were hampered by the small numbers of exposed subjects. The explanation that cervical cancer

---

[126]The cautions with regard to U-TCA noted earlier in Axelson's study apply here as well. The resulting effects of exposure misclassification would tend to reduce the TCE association, not produce a false one.

is a result of confounding by low socioeconomic status is not supported by the Anttila study, since there is a relationship of cervical cancer with measured U-TCA levels.

Anttila's conclusion was that the study provides "support to the hypothesis that trichloroethylene and other halogenated hydrocarbons are carcinogenic for the liver and lymphohematopoietic tissues [blood cancers], especially non-Hodgkin's lymphoma. The results also suggest that exposure to these solvents may increase the risk of pancreatic cancer…[and] cancers of the stomach, cervix uteri, prostate, and nervous system … "

Epidemiological studies of drycleaning workers

The non-aqueous cleaning of clothes ("drycleaning") is performed with organic solvents and detergents. Since the 1930s, the three most common solvents have been petroleum solvents, TCE, and PCE, with PCE being the predominant agent since the 1960s.[127] Unlike many industrial uses of organic solvents, where several solvents may be used simultaneously, drycleaning workers were usually exposed to one or at most two solvents, either PCE only, TCE only (in earlier years in the US and in Europe and Japan) or petroleum solvents and PCE. Drycleaning workers have thus attracted interest as a way to understand the effects of these solvents. A number of studies of drycleaning workers have been performed, and updates of earlier studies have appeared. I review primarily the latest of the studies in the case of follow-ups.

In the 1970s Blair et al. began to study a union of drycleaning workers from Missouri. Their latest paper on this cohort appeared in 1990 (**Blair et al.**)[128] In this study data from union records on 5365 members enrolled before 1978 and employed for one year or more were

---

[127] 1,1,1-trichloroethane (TCA) and chlorofluorocarbons were used to a lesser extent for special purposes like furs as well.

[128] Blair A, Stewart PA, Tolbert PE, Grauman D, Moran FX, Vaught J, Rayner J, "Cancer and other causes of death among a cohort of dry cleaners," *Br J Indust Med* 47:162-168, 1990.

abstracted for sex, age, race, job title at time of entry, and most recent firm of employment. Their vital status (alive or dead) as of the end of 1978 was determined, with a fairly large 12% loss to follow-up (*i.e.*, only 88% of the cohort could be successfully traced). In addition, 425 workers were excluded because there was not sufficient information to trace their vital status. As a general rule, losses to follow-up and incomplete information tend to lower observed risks of death. Mean duration of follow-up was only about 20 years for each race/sex group, which, again, as a general rule is fairly short to establish risk of death from a solid tumor.[129]

Selection bias in terms of an evident healthy worker effect for all causes of death was evident (SMR = .9), but interestingly, risk of death from cancer was increased (SMR = 1.2). Significant increases were seen for esophageal and cervical cancer, and statistically unstable increases were seen for several other sites, including non-Hodgkin's Lymphoma (NHL) and Hodgkin's Disease. Risk of death from kidney cancer was not increased, which was not unexpected given the relatively short latency and use of mortality as an endpoint. Risk of death from lymphoma and other blood cancers was highest in the group with highest estimated levels of exposure to dry cleaning solvents, with a dose response trend that was statistically significant (all race/sex groups combined: SMR = 4.0, rate ratio comparing high to low exposure = 3.7). The authors comment that the increase in lymphomas but not leukemia seen in their study contrasts with the increase in leukemias reported by others. There was no control for confounding in this study, although it is unlikely confounding alone could produce such a large risk, although it could lower an even higher one.

---

[129] The IARC Monographs include this guideline in their preamble about evaluating studies of human cancer: "Experience with human cancer indicates that, in some cases, the period from first exposure to the development of clinical cancer is seldom less than 20 years; latent periods substantially shorter than 30 years cannot provide evidence for lack of carcinogenicity." (p. 17 of IARC Monograph 63, cited above).

Given the significant limitations of this study that would tend to lower risks (use of mortality instead of incidence, likely exposure misclassification according to the authors, and incomplete follow-up, relatively short-term follow-up, and significant exclusions), the increases in lymphoma risk are fairly impressive.

In 1994 and again in 2001, **Ruder,** et al.[130] published updates to Brown and Kaplan's 1985 study of drycleaning workers. The Ruder studies were a follow-up study of 1701 union members in four states employed for at least a year before 1960 on premises where PCE and no carbon tetrachloride was used. In the most recent follow-up mortality was observed in this cohort through 1996. The means for ascertainment given in the Brown and Kaplan study seemed appropriate. Two sub-cohorts were established, one consisting of workers in premises where PCE only was known to be used, another "PCE-plus" cohort, consisting of workers in premises where PCE and other solvents were used. The authors have continued to observe an excess of all cancer deaths (SMR 1.25, CI 1.11 – 1.41), and specific excesses in tongue, esophageal, kidney, bladder, cervical and intestinal cancer.

More specifically, elevations in mortality from esophageal, intestinal and bladder cancer observed in the previous update (deaths through 1990) continue to be seen through 1996. Esophageal cancer was in excess in all four gender – race categories, although the estimates were not very stable due to small numbers. Intestinal cancer was substantially elevated (SMR 2.3, CI 1.0 – 4.53, white men; 2.06, 1.153.0, white women). Bladder cancer deaths were increased four-fold in non-white men (4.19, 1.14 – 10.7). There were also reported elevations that were stably above 1.0 for cancer of the tongue, long and cervix. The lung cancer SMRs were 1.88 (1.07, 3.05) for women, 1.52 (1.05, 2.39) for men. Cervical cancer was also elevated for both white and

---

[130] Ruder A, Ward E, Brown D, "Cancer mortality in female and male dry-cleaning workers," *J Occup Med* 36:867- 874, 1994; Ruder A, Ward E, Brown D, "Mortality in dry-cleaning workers: an update," *Am J. Ind. Med.* 39:212 – 132, 2001.

non-white women. Kidney cancer was increased for the entire cohort (SMR = 1.41, C.I. .4 – 3.3). The SMR for the PCE-only cohort was 1.73 (.21, 625) and for the PCE-plus cohort 1.27 (.26, 3.72). The use of mortality as an endpoint is problematic for kidney cancer, as this disease can be successfully treated with nephrectomy (removal of the kidney). In addition, the usual selection problems of healthy worker effect and exposure misclassification may have biased results lower, thus underestimating the risk. There are thus sound scientific reasons to believe the risks for kidney cancer might be higher than indicated here.

The German kidney cancer epidemiological studies

For several decades investigators in Germany have been studying the cancer effects of TCE, PCE and their metabolites, both in the laboratory and in epidemiological studies.

Henschler had been studying TCE since 1977. The epidemiology of kidney cancer began in 1995 with a study by Henschler and his colleagues.

Henschler, et al.[131] – In 1995 Henschler performed a retrospective cohort study of cardboard factory workers: 169 workers who had been exposed to TCE and 190 unexposed workers (classified on the basis of job title) over an average period of 34 years. By all accounts, exposures to TCE (the predominant solvent used, with little exposure to other agents) were very heavy, but no air monitoring data were available. Controls were matched for age, physical activity (no office workers were used as controls), and information on weight, height, blood pressure, use of diuretics, smoking habits, and alcoholic beverages recorded. Individuals were traced in both groups for determination of cause of death or appearance of kidney cancer.

---

[131] Henschler D, Vamvakas S, Lammert M, Dekant W, Kraus B, Thomas B, Ulm K, "Increased incidence of renal cell tumors in a cohort of cardboard workers exposed to trichloroethylene," *Arch Toxicol* 70:131-133, 1995.

Comparisons were made by appropriate statistical methods. In addition, separate comparisons were made with the German and Danish cancer registries.

In addition to the five kidney cancers originally observed, and presumably the impetus for the study, two additional tumors were found as a result of the study and follow-up. All seven of these tumors were in the exposed group, none in the unexposed group. External comparisons (cancer registry) produced SIRs (standardized incidence ratios) of 10 or above (a ten-fold increase in risk). The internal comparison (with the factory unexposed group) showed increased risks of five- to seven-fold. All results were statistically stable.

Compared to other studies that had presented results on TCE exposure and kidney cancer to this point, the Henschler study was unusually strong for the following reasons: (1) the long latency of 34 years allowed enough time for the cancer to develop; (2) exposures were heavy, thus increasing the chance of seeing an effect in a small population of workers; (3) there was little in the way of other exposures to confuse the picture; (4) known confounders for kidney cancer were taken into account; (5) there was biological plausibility and a demonstrated mechanism available for this particular tissue, reviewed by the authors at the end of their paper.

On the other hand, some scientists perceive a weakness in the study, in that its origin was apparently the detection of a "cluster" of cases among TCE exposed workers in the factory. Thus epidemiologic confirmation in this study was seen by some to be just a repetition of this same observation.[132]

As a consequence of this criticism (which was rebutted by the authors), a series of follow-up investigations were conducted, using different designs and different patients. Each study was designed either to confirm or disconfirm the initial findings. The results have been striking and

---

[132] This view was taken by some on the IARC panel, who reviewed a manuscript of the study.

startling. In particular, a relationship has been found between TCE exposure and a particular "cancer gene" (a tumor suppressor gene), mutations of which were previously known to be characteristic of kidney cancer. I briefly review this set of important studies.

In 1998, **Vamvakas et al.**[133] published an important case-control study of kidney cancer in Germany. None of the cases recorded in the original Henschler study were used. Thus the Vamvakas study provides independent confirmation of the results of the original Henschler study. All cases of kidney cancer seen at a large hospital urology department were eligible for the study. Of 73 such cases, 11 could not be contacted and 4 had died, with no occupational history obtainable. This left 58 cases, each of which was reviewed histologically by a specialist in kidney pathology, who confirmed that each cancer was cancer of the kidney proper (excluding renal pelvis). The region where the hospital is located is heavily industrialized, with many small premises involved in metal work or electrical device manufacture, and solvents like PCE and TCE were often used. A set of 84 controls were enrolled from accident victims who had been seen at three local hospitals, all within a radius of 20 km of both each other and the study hospital, so as to be from the same geographic area as the cases.

Occupational, medical and personal histories were taken from cases and controls (or from relatives or former colleagues in the case of deceased cases), and demographic information as well as occupational exposure to TCE, PCE, cadmium, lead, nickel, chromium, gasoline, benzene, asbestos, pesticides and PCBs were recorded, along with information about body mass index, blood pressure, smoking habits, alcohol consumption, use of diuretics and history of kidney disease and family history of kidney disease and cancer. Comparisons between the two groups for occupational exposure to TCE and PCE were made and analyzed using appropriate

---

[133] Vamvakas S, Bruning T, Thomasson B, Lammert M, Baumuller A, Bolt HM, Dekant W, Birner G, Henschler D, Ulm K, "Renal cell cancer correlated with occupational exposure to trichloroethene," *J Cancer Res Clin Oncol* 124:374-382, 1998.

statistical methods (multivariate logistic regression) and Mantel-Haenszel methods for combining stratified tables.

Cases and controls were similar in body mass index and smoking history (combining current and former smokers) as well as all other factors except for blood pressure and diuretic use. Because hypertension and diuretic use are associated with having kidney disease (not causing it), this was an expected and not relevant difference. TCE and PCE exposures were combined into one variable "because of their identical toxicological mechanisms…" (p. 380). Controlling for confounders, logistic regression resulted in an OR of 10.80 (3.36, 34.75) for exposure to TCE and PCE. When stratified by exposure intensity the results were stronger: low-level, OR 6.61 (.5, 85.76), medium level OR 11.92 (2.55, 55.6), high level 11.42 (1.96, 66.79). When stratified by age and analyzed by Mantel-Haenszel methods, the estimate of the OR was 8.96 (2.9, 27.75).

This study shows a strong association between exposure to TCE/PCE and kidney cancer, confirming earlier independent data. There is a potential for information bias in this study, and the usual problem of exposure misclassification. The two biases would likely work in opposite directions, *i.e.*, they tend to cancel out each other. The relatively large case size, good control of confounding, and careful confirmation of diagnosis make this study of particular importance.

These studies take on added significance when seen in the context of the emerging data about the molecular epidemiology of kidney cancer and individual susceptibilities. **Bruning et al.**[134] examined 45 kidney cancer patients from a group with long-term occupational exposure to TCE only, and compared them to exposed workers who did not develop kidney cancer. In

---

[134] Bruning T, Lammert M, Kempkes M, Their R, Golka K, Bolt H, "Influence of polymorphisms of GSTM1 and GSTT1 for risk of renal cell cancer in workers with long-term high occupational exposure to trichloroethene," *Arch Toxicol* 71:596-599, 1997.

particular the authors were looking for genetic variations in two specific enzymes important in detoxifying TCE, and in the process converting it to a form thought to be either the cause of or a contributing factor in producing kidney cancer. It was hypothesized that individuals with greater ability to perform this transformation would be at higher risk, and indeed, individuals without this ability (GSTT1 null and GSTM1 null phenotypes) were under represented among the cancer cases compared to the non-cases.

Of even more interest was the discovery, also by Bruning and coworkers[135] that a genetic marker of common kidney cancer is also present in abnormal prevalence among TCE exposed workers with kidney cancer. The gene, a tumor suppressor gene called the von Hippel-Lindau (VHL) gene, is mutated in about 30-50% of common kidney cancer cases. An examination of 23 kidney cancer patients with TCE exposure, however, found mutations in the VHL gene in 100% of the patients. Follow-up work confirmed this result. Thus, **Brauch et al.** [136] have now shown that 44 patients with kidney cancer and known industrial exposure to TCE had a much higher frequency of VHL gene mutation, singly and multiply, than 107 kidney cancer patients who had not been exposed to TCE. In this series 75% of the TCE exposed workers had VHL mutations compared to 58% of the non-exposed workers, but of more interest was the presence of a specific mutation at nucleotide 454 that was present in 29% of the exposed workers but *none* of the 107 unexposed workers.[137]

---

[135] Bruning T, Weirich G, Hornauer MA, Hofler H, Brauch, "Renal cell carcinomas in trichloroethene (TRI) exposed persons are associated with somatic mutations in the von Hippel-Linday (VHL) tumor suppressor gene," *Arch Toxicol* 71:332-335, 1997.

[136] Brauch H, Weirich G, Hornauer MA, Storkel S, Wohl T, Bruning T, "Trichloroethylene exposure and specific somatic mutations in patients with renal cell carcinoma," *JNCI* 91:854-861, 1999.

[137] The paper reports this percentage as 39%, but an examination of the tables shows the true number is 29% (13/44).

This is a very striking and important finding, representing, according to the authors, "the first molecular evidence for a relationship between exposure to a defined carcinogen, gene damage, and kidney cancer." (p. 859). There was also an evident dose-response relationship between degree of exposure and codon 454 mutation and number of mutations in the VHL gene (table 4 of paper).

Environmental epidemiological studies of blood cancer caused by PCE and TCE in drinking water

The Lagakos et al. study ("The Harvard Study")

Non-occupational studies have also revealed cancer risks in PCE and TCE exposed populations. One of the most well-known is the study from the Harvard School of Public Health of the Woburn[138] leukemia cluster by Lagakos, et al. (1986), linking access to PCE and TCE contaminated well-water to increased risk of childhood leukemia. The Woburn cluster is the subject of the famous *"Civil Action"* case that was the featured in the book *Civil Action* by Jonathan Harr and the movie of the same title with John Travolta.

The design was case-control (not, as sometimes erroneously reported, ecologic). This study has engendered much discussion,[139] but the results are entirely consistent with both animal studies and other epidemiological work, as noted. With respect to the supposed criticisms in the literature concerning this study, some clarification is in order. The Harvard-Lagakos study appeared in the *Journal of the American Statistical Association (JASA)*, a highly regarded specialty journal. It was followed by several critical comments by various scientists, with a

---

[138] Woburn is an industrial town of 35,000 people 13 miles northwest of Boston.

[139] See comments by MacMahon, Prentice, Rogan, Swan and Robins, and Whittemore following Lagakos S, Wessen B, Zelen M, "An analysis of contaminated well water and health effects in Woburn, Massachusetts," *J Am Stat Assoc* 81:583-596, 1986.

rebuttal by the authors. The existence of these criticisms (but not the rebuttal) has been used by some to argue that this paper is unusually controversial and unreliable. On the contrary, the authors first presented the paper at the American Statistical Association's annual meeting and took the unusual step of first circulating the paper to colleagues and *inviting* criticism at that same meeting, with eventual simultaneous publication of the critiques and the paper. Rather than evidence of unreliability, then, this is evidence of unusual scientific probity and candor.

One of the authors, Marvin Zelen, former Chair of the Department of Biostatistics at the Harvard School of Public Health and now Director of Biostatistics at the famed Dana Farber Cancer Center at Harvard, has persuasively addressed the issues raised by the critiques of his study. I have evaluated the claims and counterclaims of bias in the study and find his defense convincing. In particular, the question of regional confounding is also well discussed in the paper. There is no evidence that such confounding was acting here, and there is good evidence that it was not (Tables 8 and 9 and text associated with them).

The criticism that the Woburn study as a data-driven "cluster investigation" is also incorrect. Clusters of cases can sometimes arise purely by chance, and when this is so no matter how sophisticated the methods, all one achieves is verification of the starting point, that a cluster exists. However this is not what was done in the Woburn study. True, a cluster of cases had already been revealed in east Woburn, so any verification that there were more cases there would be redundant and not especially informative. But the Woburn study involved testing a true *a priori* hypothesis (which could easily have been false), that cumulative exposure to contaminated water was higher in the leukemia cases then a matched risk set of Woburn residents. The exposure was not to *east Woburn* but to contaminated *well water*. Thus no rates were compared between the census tracts. The study population was all of Woburn, not just east Woburn.

As for the role of chance, it is generally acknowledged that one of the strengths of the study was the innovative use of new statistical techniques (Cox Proportional Hazards Modeling in a case-control study). It has been asked if ignoring interactions might have resulted in some bias. Since the authors of the Woburn study were only interested in the existence or not of an association, and interaction effects would only diminish such an association if present, this is not a serious issue. The question of confounding is addressed on page 587 of the paper where none of the risk factors were found to be correlated with wells G/H exposure. As regards the so-called "multiple comparison" problem, it should be noted that there was an *a priori* hypothesis for water exposure and leukemia.

The case for "real effect" in this paper is complemented by the other information available about PCE and TCE, which is being discussed here.

The Massachusetts Department of Public Health follow-up study

An important follow-up to this paper was performed by the Massachusetts Department of Public Health (MDPH).[140] The MDPH study was a matched case-control design study with two controls per case. Children 19 years old or younger who had been diagnosed with leukemia between 1969 and 1989 while living in Woburn comprised the case group, while the control group was randomly selected from Woburn school records and matched for date of birth, sex and race. Twenty-one (21) cases and 42 controls met the definitions used. Investigators recorded residential history, occupational information of parents, and medical histories on cases and controls for the appropriate time periods (from two months before conception to the date of diagnosis). For the purpose of analysis the time periods were divided into the two years before conception to conception, during pregnancy, and from birth to diagnosis. A refined

---

[140] Massachusetts Department of Public Health, *Woburn Childhood Leukemia Follow-up Study: Final Report, Volumes I (Analyses) and Volume II (Appendices)* MDPH Bureau of Environmental Health Assessment, July 1997.

computerized water distribution model[141] was used to estimate exposure in a more precise geographic and temporal resolution. It estimated the proportion of water from wells G and H to reach households on monthly intervals. The model was calibrated and validated.

Odds ratios (ORs) were calculated after controlling for socioeconomic status, maternal smoking during pregnancy, maternal age at birth of the child, and maternal alcohol consumption during pregnancy. Adjusted ORs for the effect of water from the contaminated wells (wells G and H) were overall OR = 2.39 (.54, 10.59), and for the sub-periods: 2-years prior to conception, OR = 2.61 (.47, 14.37); during pregnancy OR = 8.33 (.73, 94.67); birth to diagnosis, OR = 1.18 (.28, 5.05). The wide confidence intervals are the result of the relatively small sample sizes (obviously new cases of leukemia could not be created). There was a statistically significant dose-response trend for exposures during pregnancy (using a trichotomous exposure metric of Never, Least, Most, the latter categories obtained with a median cut point from the water model results).

The MDPH investigators concluded that "the risk of developing childhood leukemia was greater for a child whose mother drank water from contaminated wells while pregnant with the child." However risks were also elevated for the preconception period and the period after birth, although these results were not included in the MDPH conclusion. Of special interest is the fact that one of Hill's considerations, that of "The Experiment" was (unusually) fulfilled in this instance. That consideration noted that a valuable indication of causality would be if some intervention that halted exposure caused the disease to disappear. In Woburn, only one additional case of childhood leukemia has developed in a person who could not have been exposed to contaminated wells G and H water in the period since the wells were shut down (and

---

[141] The Lagakos study used the Waldorf-Cleary Model, while the refined model used here is known as the Murphy Model.

hence exposure ceased; the wells shut down in 1979, and cases continued to occur until 1987 and then suddenly stopped; this lag is the latency period for blood cancers). This is a dramatic, if unplanned, confirmation of both the original Harvard-Lagakos study and the MDPH follow-up.

I note, in passing, that the relationship between parental occupational and home exposures had previously been studied by a number of investigators, notably Lowengart, et al. [142] This was a matched case-control study of specific exposures of both parents from one year before conception until the diagnosis of leukemia in the child. A variety of other suspected risk factors were included. In concept this study is similar to the MDPH study just considered. Cases were identified through the population-based cancer registry in Los Angeles County, which covers the same area as the Boice study. Cases of acute leukemia (ALL and AML combined) 10 years of age or less (as compared to the 19 years or less in the MDPH study) were compared to friends of the cases, *i.e.,* individuals of the same age, or if none were suitable or available, a control individual selected from the area by random-digit dial. There were 159 (79%) of eligible cases and 130 age, sex, race and Hispanic origin matched controls. Structured interviews were conducted for occupational exposures of the parents and relevant risk factors and time periods of exposure. There was no attempt to confirm the exposures.

There were elevated ORs for parents who worked in the Transportation Manufacturing industry, most of whom worked in aircraft manufacturing (OR = 2.5, p = .03; OR = 1.8 for aircraft manufacturing, p = .12). The OR for childhood leukemia and chlorinated solvent exposure (noted as TCE, PCE and carbon tetrachloride) was 3.5 (1.1, 14.6), and for specific solvents was OR = 2.7 (.64, 15.6) for TCE exposure after birth, OR = 2.0 (p = .16) for TCE exposure during pregnancy, and OR = 2.0 for TCE exposure in the year before pregnancy (p =

---

142 Lowegart RA, Peters JM, Cicioni C, Buckley J, Bernstein L, Preston-Martin S, Rappaport E, "Childhood leukemia and parents' occupational and home exposures," *JNCI* 79:39-46, 1987.

.16). For PCE exposure only one case and no matched controls were exposed, so no ORs could be calculated (they would be undetermined, or, as indicated in the paper, "infinite").

This study found the same risk factors for ALL and AML, showing that for etiologic purposes, there was no distinction in the leukemia types.

The authors point out that the lack of a general tendency toward elevated ORs among those with chemical exposure (as opposed to specific exposures) argues against recall bias being a significant factor in the results. Taking information bias into account also did not change the risks. Confounding was reduced by the matching. On the other hand, misclassification of exposure, likely in this case, would have tended to bias the risks downward, thus underestimating them. In summary, this study, like other similar studies of occupation and risk of childhood cancer, and the MDPH study in particular, showed a substantial increased risk for leukemia in offspring of parents exposed to chlorinated solvents at work (given as TCE, PCE and carbon tetrachloride). This increase was not readily explainable on the basis of selection bias, information bias or confounding. This was a well-done study by established investigators.

Upper Cape studies of PCE in drinking water

For nearly two decades my Boston University colleagues and I studied an unusual exposure to PCE in drinking water in the Upper Cape region of Massachusetts. These epidemiological studies of cancer and PCE began over twenty years before my involvement in this litigation, and the results have been reported in two different highly reputable peer review journals.[143] In 1993 we published an article in *Archives of Environmental Health* (Aschengrau et

---

[143] Aschengrau, A. and Ozonoff, D. Upper Cape Cancer Incidence Study. Final Report. Massachusetts Department of Public Health, Boston, January 9, 1992, 700 pp; Aschengrau, A., Ozonoff, D. Paulu, C., Coogan, P. Vezina, R., Heeren, T., Zhang, Y., "Cancer risk and tetrachloroethylene (PCE) contaminated drinking water in Massachusetts," Archives of Environmental Health, 48:284-292, 1993; Aschengrau, A, Paulu C, Ozonoff D, "Tetracholoroethylene-contaminated drinking water and the risk of

al.) which showed a marked increase risk of leukemia in people exposed to the highest levels of PCE in drinking water (leukemia OR = 8.3 [1.5 - 45.3]). There were essentially no other contaminants in the water. Increased risks were also seen for bladder cancer, but not for kidney cancer. As my BU colleagues and I noted in our 1993 article, however, the relatively short period between first exposure and ascertainment of diagnosis (maximum of 14 years) is insufficient time to allow development of a solid tumor like kidney cancer. Blood cancers like leukemia have much shorter latencies and can be seen earlier than solid tumors.

This was a population-based case-control study, with careful confounder control and the use of a mathematical model to estimate exposure to individuals (exposure assignments were done blind to case status). Significantly, although the number of cases was relatively small, the demonstrated effect was relatively strong.

The ecologic designs in New Jersey

Fagliano et al.[144] and **Cohn et al.**[145] have investigated New Jersey towns with organic chemical contamination, especially involving TCE, and found higher leukemia and lymphoma (*i.e.*,, blood cancer) rates for women in those areas. The Cohn study is a follow-up and expansion of the Fagliano et al. study, so I will confine my remarks to this study. Seventy-five (75) New Jersey towns were compared for TCE contamination and the incidence of leukemia and lymphoma, with data obtained from the New Jersey Cancer Registry. Stable excesses in

---

breast cancer," Environ Health Perspect, 106(suppl4):947-953, 1998; Paulu C, Aschengrau A, Ozonoff D, "Tetrachloroethylene-contaminated drinking water in Massachusetts and the risk of colon-rectum, lung, and other cancers," Environmental Health Perspectives, 107:265-271, 1999

144 Fagliano J, Berry M, Bove F, Burke T, "Drinking water contamination and the incidence of leukemias: an ecologic study," *Am J Public Health* 80:1209-1212, 1990.

145 Cohn P, Klotz J, Bove F, Berkowitz M, Fagliano J, "Drinking water contamination and the incidence of leukemia and non-Hodgkin's lymphoma," *Environ Health Persepct* 102:556-561, 1994.

lymphoma among women were seen,[146] and stable excesses for total leukemia in women, acute lymphatic leukemia of childhood in girls, chronic myelogenous leukemia in women, and chronic lymphocytic leukemia in men and women were seen when comparing towns with the highest TCE contamination to those without detectable TCE contamination.

The basic design shared by both the Cohn et al. and Fagliano et al. studies involves use of some population measure of exposure, together with individual outcome data. The exposure measure is a weighted average of measured water supply values for TCE and other contaminants, which common value is then assigned to all subjects in the township. The outcomes are cancer incidence rates for the blood cancers in each township. The question investigated was whether there was a relationship between the level of contamination in the township's water supply and the rate of blood cancers in the township's population.

This is not unlike a cohort study, such as the Lockheed-Boice study, where a common occupational exposure is used for a job title. Unraveling the various claims and counterclaims of such study designs is highly technical and depends upon specifics of each study. Across-the-board statements about the reliability and types of bias of this kind of design are inappropriate and misleading.

Specifically, there are two ways to look at this problem, either as an individual level study with mismeasured exposure, or as a "semi-individual" level design[147] with aggregated exposure measure. Depending upon the error model[148] one uses, there are different implications,

---

146 PCE also seemed associated with lymphoma, but because many towns with PCE contamination also had TCE contamination it was more difficult to interpret the results as due only to PCE.

147 The term semi-individual for these studies seems to have been coined in Kunzli N, Tager IB, "The semi- individual study in air pollution epidemiology: a valid design as compared to ecologic studies," *Environ Health Perspect* 105:1078-1083, 1997.

148 I refer specifically to a Classical Error Model or the so-called Berkson Error Model. Good (but highly technical) treatments can be found in Fuller WA, *Measurement Error Models*, Academic Press, 1987; Carroll RJ, Ruppert D, Stefanski LA,

as there are for particular error structures for any model (relationship between the covariance of the individual exposures compared to the between group variance of the aggregated exposures)[149]. The technical question is delicate and complex, but it is certainly not properly treated by rejecting semi-individual designs out of hand[150].

In the case of the Cohn et al. and Fagliano et al. studies, we find there is indeed a relationship (statistically stable) between population rates of blood cancers and contamination of the water supply with TCE. Because of the difficult technical questions involved, this result, standing alone, would be less interpretable than it is in the context of corroborating studies (such as the Upper Cape and Woburn studies of blood cancer and environmental exposure through drinking water, appearance of blood cancers in the animal bioassays, blood cancers in the dry cleaner studies, IARC recognition of blood cancer as related to both PCE and TCE in epidemiological studies). It thus plays a relatively ancillary, but still contributing, role in my conclusion that drinking water contaminated with PCE and TCE is related to blood cancers.

Vinyl Chloride Studies

Creech and Johnson[151]were the first to report on the carcinogenicity of VC in workers when they published a paper detailing three cases of a very rare cancer of the liver

---

Measurement Error I Nonlinear Models (Monographs on Statistics and Applied Probability 63), Chapman & Hall/CRC Publishers, 1998. Applications to the aggregation bias problem can be found in Steenland K, Deddens JA, "Design and analysis of studies in environmental epidemiology," In: Steenland K, Savitz D (eds) *Topics in Environmental Epidemiology,* Oxford University Press, 1997, p. 23.

149 In fact, there is no measurement bias at all in a semi-individual study when the group mean is used for exposure, as might be expected, since least squares estimation regresses the expected value of the outcome on the mean of the exposure. When substituting the mean, one does not bias this result. However there can be a magnification of the confounding bias, whose direction depends upon the error structure (it can go in either direction or even change sign).

150 This might be expected from their conceded value as "hypothesis generators" (and hence a concession they contain actual information).

151 Creech, JL; Johnson, MN. Angiosarcoma of the liver in the manufacture of polyvinyl chloride. J Occup Med 16:150-151, 1974.

(angiosarcoma), all occurring in workers manufacturing polyvinyl chloride (PVC) plastic. All three workers were exposed to the vinyl chloride monomer, the elementary building block of PVC plastic. This was a striking observation, as only 20-30 cases of angiosarcoma of the liver are reported in the US each year.

Since then, numerous studies, using various study designs, have confirmed the carcinogenicity of VC for humans. I do not believe this is a matter of significant disagreement amongst scientists knowledgeable about this chemical. I give some of the citations in this footnote.[152]

d. Between the Bookends: Mechanisms

The toxicological literature on PCE, TCE and VC is large. For the present purposes I touch on only a few points of relevance to considering the carcinogenic risks of these chlorinated ethylenes. These issues are related to whether the mechanisms and modes of action observed in the animal and toxicological studies are relevant to human health risk, including whether any evidence of cancer seen at high doses in animal and workplace studies are of any relevance to risks at the presumably lower exposures seen in the residential environment as a result of the contamination by PCE in the Class Area.

The entire debate is sometimes (mistakenly) framed in terms of whether chlorinated ethylenes like PCE, TCE and VC are "genotoxins" or not. Genotoxicity, the ability of a chemical to alter the genetic make-up of a cell (cause a somatic mutation), is a factor in assessing the ability of a chemical to cause cancer at environmental doses, although it is not completely necessary. The importance of this factor relates to the clear connection to a known mechanism

[152] Waxweiler, RJ; Stringer, W; Wagoner, JK; et al. Neoplastic risk among workers exposed to vinyl chloride, Ann NY Acad Sci 271:40-48, 1976; Monson, RR; Peters, JM; Johnson, MN, Proportional mortality among vinyl chloride workers, Environ Health Perspect 11:75-77, 1975; Tabershaw, IR; Gaffey, WR, Mortality study of workers in the manufacture of vinyl chloride and its polymers, J Occup Med 16:509-518, 1974; Byren, D; Engholm, G; Engund, A; et al. Mortality and cancer morbidity in a group of Swedish VCM and PVC production workers, Environ Health.

of carcinogenesis (mutation in an important gene that regulates growth, positively or negatively); and the realization that this minute change, the mutation, once made no longer requires the presence of a minute amount of chemical that produced it but will constitute a self-reproducing damage as each cancer cell now makes two new ones, and so on with each replication of the cell. Thus the minute chemical change in the DNA of an original single cell is the proverbial spark in the dynamite factory, leading to downstream catastrophe. I give my reasons below for believing that PCE, TCE and VC work via a genotoxic mechanism, although other mechanisms are possible in addition or instead of it.

The opposing view, that PCE, TCE and VC are not genotoxic, is related to the fact that it is believed, that some chemicals can cause cancer without themselves causing mutations. There have been claims that the chlorinated ethylene TCE is such a chemical, and moreover, as an ***added*** claim, that it causes cancer in animal models through a non-genotoxic mechanism that is not possible in humans. These two claims are frequently confused, but they are independent and separate. I discuss genotoxicity first, then the special claim that any non-genotoxic mechanism is irrelevant to humans.

i. The genotoxicity of PCE and TCE

The question of genotoxicity is not determinative here. Whether PCE and TCE are, or are not, genotoxic is a separate question from whether a non-genotoxic effect would allow as much risk at environmental or occupational doses as a genotoxic one. Claiming a specific non-genotoxic mechanism that does not allow cancer effects at doses to which humans are exposed environmentally *or* occupationally, as some have tried to do, is a separate question, although one

sometimes confused with the former one.[153] As the epidemiologic data show, humans *do* experience increased cancer risks at exposures encountered both in the workplace and the environment. The question has been asked about what mechanisms it uses to do this and further, whether those mechanisms operate in humans and at the doses seen in the Class Area.

The specific effects on the VHL gene are one important indication that chlorinated ethylenes are genotoxic but it is not the only one.[154] We also discussed, above, the indication that TCE affects the H-*ras* oncogene in a specific and characteristic way. Two additional studies are also of interest. Both involve means to detect a broader range of genetic damage from foreign chemicals than older methods. The work of **Schiestl et al.**,[155] reported in 1997 in the *Proceedings of the National Academy of Sciences,* is in essence a method that detects large scale genetic changes like deletions of whole genes. Using this method, Schiestl et al. were able to show that TCE, benzene and sodium arsenate caused genetic deletions (a form of genetic damage). These are all carcinogens whose mutagenic activity was not readily detectable by the usual assays.

Of even more importance is a comprehensive report from the Gene-Tox program of the USEPA on the mouse lymphoma specific gene and chromosomal mutation assay (abbreviated

---

[153] For example, the most potent carcinogen known, 2,3,7,8-TCDD ("dioxin") is an alleged non-genotoxic carcinogen. Thus even vanishingly small quantities of a non-genotoxin can apparently cause cancer.

[154] In addition to the oncogene/tumor suppressor work see Rasmussen K, Sabroe S, Wohlert M, Ingerslev HJ, Kappel B, Nielsen J. "A genotoxic study of metal workers exposed to trichloroethylene,: sperm parameters and chromosome aberrations in lymphocytes," *Int Arch Occup Environ Health* 60:419-423, 1988; Walles SAS. "Induction of single-strand breaks in DNA of mice by trichloroethylene and tetrachloroethylene," *Toxicol Letters* 31:31-35, 1986; Konietzko H, Haberlandt W, Heilbronner H, Reill G, Weichardt H. "Cytogenetische untersuchungen an Trichlorathylen-Arbeitern," *Arch Toxicol* 40:201-206, 1978.

[155] Schiestl RH, Aubrecht J, Khogali F, Carls N. "Carcinogens induce reversion of the mouse pink-eyed unstable mutation," *Proc Natl Acad Sci USA* 94:4576-4581, 1997.

MLA).[156] Like the method of Schiestl et al., the MLA can detect large scale genetic alterations, not just point or localized mutations. The results of the MLA on 602 chemicals were reviewed and compared with animal bioassays to see the relationship of the MLA and animal carcinogenicity. The MLA reconfirmed the fact that PCE and TCE were genotoxic and placed it in the category of "definitively positive." The 602 chemicals were grouped into 30 chemical classes, containing chemicals of a similar chemical and biological nature. PCE and TCE fell into class 2 (102 of the 602 chemicals).[157] Of those chemicals tested for carcinogenicity by the National Toxicology Program (NTP), the concordance with the MLA in class 2 chemicals (for both positive and negative results) was 94% (i.e., of all the positive *and* negative carcinogenicity results by NTP of class 2 chemicals, 94% of them agreed with the positive or negative results of the MLA). This was further confirmation that the MLA was identifying genetic changes of significance for the production of cancer. In particular, TCE was judged by the panel to be definitively positive (the highest category) in the MLA.[158] Thus TCE is not only genotoxic in the MLA, but the genetic alterations it causes are relevant for its potential to cause cancer. This further confirms the results obtained from a different line of research (oncogene work).

The genotoxicity of PCE and TCE shows them capable of acting by known mechanisms of cancer causation.

---

[156] Mitchell AD, Auletta AE, Clive D, Kirby PE, Moore MM, Myhr BC. "The L5178Y/tk+/- mouse lymphoma specific gene and chromosomal mutation assay. A phase III report of the U.S. Environmental Protection Agency Gene-Tox Program," *Mut Res* 394:177-303, 1997. A review of the mutagenicity of TCE and its metabolites, Moore M, Harrington-Brock K, "Mutagenicity of trichloroethylene and its metabolites: implications for the risk assessment of trichloroethylene, *Environ health Perspect* 108(suppl 2):215-223, 2000, fails to take note of either Schiestl et al. or the recent MLA paper by Mitchell et al., cited here.

[157] Acyl and aryl halides, halogenated ether, halohydrins, saturated and unsaturated alkyl halides. Ibid. p. 264.

[158] See Table 1, p. 227 in ibid.

ii. Non-genotoxic mechanisms and their relevance

The fact that PCE and TCE are genotoxic under some circumstances does not mean that they are genotoxic under all circumstances, or even that one of or their sole mode of causing cancer is through genotoxicity. Other mechanisms might exist. Indeed, some people consider PCE and TCE "non-genotoxic" carcinogens, i.e., chemicals that cause cancer by some means other than fixing an alteration of the DNA (mutation). Exactly what this other mechanism might be, if indeed there is one, is a matter of some debate, even among those who do not believe PCE and TCE are genotoxic. One popular suggestion was the "peroxisome proliferator" mechanism.[159] But the underlying hypothesis (that peroxisome proliferation resulted in either cell proliferation, oxidative stress or tumor promotion) has not been borne out by subsequent studies.[160] The subject is a matter of some controversy since it also affects other chemicals

---

[159] Despite its popularity in the early nineties, however, IARC rejected the argument. See International Agency for Research on Cancer, *Dry Cleaning, Some Chlorinated Solvents and Other Industrial Chemicals*, IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 63, Lyon 1995, pp. 136-137.

[160] See, for example, Reddy JK, Rao MS. "Oxidative DNA damage caused by persistent peroxisome proliferation: its role in hepatocarcinogenesis," *Mut Res* 214: 63-68, 1989, for a statement of the oxidative stress hypothesis, and recent reviews, Melnick RL, Kohn MC, Portier CJ. "Implications for risk assessment of suggested nongenotoxic mechanism of chemical carcinogenesis," *Environ Health Perspect* 104(Suppl 1):123-134, 1996, p. 123. See also Nelson MA, Lansing AJ, Sanchez IM, Bull RJ, Springer DL. "Dichloroacetic acid and trichloroacetic acid-induced DNA strand breaks are independent of persoxisome proliferation," *Toxicology* 85:239-248, 1989, p. 240; Mukherjee R, Jow L, Noonan D, McDonnell DP. "Human and rat peroxisome proliferator activated receptors (PPARs) demonstrate similar tissue distribution but different responsiveness to PPAR activators," *J Steroid Biochem Molec Biol* 51:157-166, 1994, p. 165; Lake BG. "Peroxisome proliferation: current mechanisms relating to non-genotoxic carcinogenesis," *Toxicology Letters* 82/83: 673-681, 1995, p. 676; Hwang J-J, Hsia MTS, Jirtle RL. "Induction of sister chromatid exchange and micronuclei in primary cultures of rat and human hepatocytes by the peroxisome proliferator, Wy-14,683," *Mutation Research* 286:123-133, 1993, p. 123; Tsutsui T, Watanabe E, Barrett JC. "Ability of peroxisome proliferators to induce cell transformation, chromosome aberrations and peroxisome proliferation in cultured Syrian hamster embryo cells," *Carcinogenesis* 14: 611-618, 1993, p. 611; Rao MS, Subbarao V. "Incidence of pancreatic and testicular tumors in rats treated with ciprofibrate, a peroxisome proliferator," *Cancer Letters* 97: 185-188, 1995, p. 185; Richert L, Price S, Chesne C, Maita K, Carmichael N. "Comparison of the induction of hepatic peroxisome proliferation by the herbicide oxadiazon *in vivo* in rat and human hepatocytes," *Toxicol appl Pharmacol* 141: 35-43, 1996, p. 41; Hofstra AH, King LM, Walker RM. "Early effects of CI-924 on hepatic peroxisome proliferation, microsomal enzyme induction, PCNA, and apoptosis in B6C3F1 mice and Wistar rats," *Arch Tox* 71:250-257, 1997.

prevalent in the environment, like the phthalate plasticizers.[161] There, as in the case of PCE and TCE, it has yet to be demonstrated that peroxisome proliferation is an obligatory step to produce cancer. It certainly is *not* the case that science has shown that PCE and TCE cannot cause cancer in humans at all, or can only do so at low dose.

Even if the mechanism of cancer induction by PCE or TCE proceeds, at least in part, by some non-genotoxic mechanism, this says nothing about their relevance for either human cancer or the probability of low dose effects. Exposures to very small amounts of a non-genotoxic chemical can produce a cancer as easily as a genotoxic one. We have examples to show this (TCDD, for example) as well as arguments that show this could easily happen from currently known mechanisms.[162]

iii. The Genotoxicity of VC

Vinyl chloride has also been shown to be a genotoxic chemical, that is, one that causes changes in the genetic blueprint of the cell found in all of the animal's DNA. The IRIS weight-of-evidence review concludes this about genotoxicity:

> VC carcinogenicity occurs via a genotoxic pathway and is understood in some detail. VC is metabolized to a reactive metabolite, probably chloroethylene oxice (CEO), which is believed to be the ultimate carcinogenic metabolite of VC. The reactive metabolite then binds to DNA, forming DNA adducts that, if not repaired, ultimately lead to mutations and tumor formation.

IRIS later cites the evidence for VC's genotoxicity (full cites from the text are in the footnote):

---

[161] Melnick, R, "Is peroxisome proliferation an obligatory precursor step in the carcinogenicity of Di(2-ethylehxyl)phthalate (DEHP)?", *Environ Health Perspect* 109:437 – 442, 2001.

[162] A discussion of the many factors that might affect the dose-response dynamics can be found, among other places, in Bull R, "Mode of action of liver tumor induction by trichloroethylene and its metabolites, trichloroacetate and dichloroacetate," *Environ Health Perspect* 108(suppl 2):241-259, 2000.

> Several lines of evidence indicate that VC metabolites are genotoxic, interacting directly with DNA. Occupational exposure to VC has resulted in chromosome aberrations, micronuclei, and sister chromatid exchanges (SCEs); response levels were correlated with exposure levels (Hansteen et al., 1978; Purchase et al., 1978; Sinues et al., 1991). VC is mutagenic in the Salmonella typhimurium reverse mutation assay, with the mutagenic activity decreased or eliminated in the absence of exogenous metabolic activation (Bartsch et al., 1975; Rannug et al., 1974). The VC metabolites CEO and CAA are both mutagenic in the Salmonella assay (Bartsch et al., 1975; Rannug et al., 1976). The highly reactive metabolite CEO was much more mutagenic than CAS, suggesting that this is the metabolite responsible for VC carcinogenicity. DNA adducts formed by VC have also been identified (Swenberg et al., 1992, 1999).[163]

The principle that non-threshold low dose extrapolations are valid for chemicals that are "genotoxic" is generally accepted by scientists.[164] There has been much discussion, however, about whether high doses convert a non-genotoxic chemical to a genotoxic one (e.g., by metabolic overload) or whether a particular chemical is genotoxic at all, instead causing cancer by some (as yet unspecified) "non-genotoxic" mechanism.

As has been remarked several times in this Report, while any chemical can be toxic to a cell at very high doses, the result is usually to make the cell falter and die, not become malignant, a very special and unusual biological response. It has been estimated that the proportion of chemicals capable of inducing such a response is less

---

[163] Hansleen, Il; Hillestad, L; Thlis-Evensen, E; et al. Effects of vinyl chloride in man: a cytogenetic follow-up study. Mutat Res 51:271-278, 1978; Purchase, IFH; Richardson, CR; Anderson, D; et al., Chromosomal analysis in vinyl chloride exposed workers, Mutat Res 57:325-334, 1978; Sinues, B; Sanz, A; Bernal, ML. Sister chromatid exchanges, proliferating rate index, and micronuclei in biomonitoring of internal exposure to vinyl chloride monomer in plastic industry workers. Toxicol Appl Pharmacol 108:37-45, 1991; Hartsch, H; Malaveille, C; Montesano, R. Human, rat and mouse liver-mediated mutagenicity of vinyl chloride in S. typhimurium strains. Int J Cancer 15:429-437, 1975; Rannug, U; Gothe, R; Wachtmeister, CA. The mutagenicity of chloroethylene oxide, chloro-acetaldehyde, 2-chloroethanol and chloroacetic acid, conceivable metabolites of vinyl chloride. Chem Biol Interact 12:251-263, 1976; Rannug, U; Johansson, A; Ramel, C; et al. The mutagenicity of vinyl chloride after metabolic activation. Ambio 3:194-197, 1974; Swenberg, JA; Fedtke, N; Ciroussel, F; et al. Etheno adducts formed in DNA of vinyl chloride-exposed rats are highly persistent in liver. Carcinogenesis 13(4):727-729, 1992; Swenberg, JA; Bogdanffy, MS; Ham, A; et al. Formation and repair of DNA adducts in vinyl chloride and vinyl fluoride-induced carcinogenesis. IARC Sci Publ 150:29-43, 1999

[164] Genotoxic chemicals alter the cell's DNA, thus initiating cancer as already described.

than 10%.[165] Among bioassays, those conducted by the NTP are among the most authoritative.[166] The protocol for these bioassays has been worked out over the years and requires the use of maximally tolerated doses (MTDs), i.e., doses that cause no pathology (other than cancer) and no more than a 10% weight loss in the dosed animals. Tests are typically conducted at the MTD and half the MTD in both sexes of rats or mice.

The MTD is almost always a dose much larger than any human would be subjected to in ordinary circumstances. Some argue such large doses make the bioassays showing VC to be an animal carcinogen not pertinent to cases where the dose is so much lower. These arguments are somewhat disingenuous as the scientific reasons for the use of high doses are well known and generally accepted.[167] Large doses are used because the objective of the bioassay is to identify just those special chemicals that can cause tumors at some dose.

Finally, whatever the administered dose, the damage to the original cell that later develops into a tumor is minute and undetectable until many generations of cell division have replicated it. The amount of active chemical that does this damage (in this case a reactive product of VC metabolism) is correspondingly minute, and after causing it need no longer take part, the damage being reproduced by the body's own biological

---

[165] Rosenkranz, HS. "Strategies for the rapid detection and identification of environmental carcinogens," Chapter 12 in Rom, WN, ed., *Environmental and Occupational Medicine*, Second Edition, Little, Brown, 1992, p. 136.

[166] "Most investigators agree that owing to the NTP's quality control criteria the rodent cancer data it generated constitute the most authoritative body of carcinogenicity data." Ibid. p. 136.

[167] This is not to say that *all* scientists accept them, only that the procedure is generally accepted. For a discussion See Fung VA, Barrett JC, Huff J. "The carcinogenesis bioassay in perspective; application in identifying human cancer hazards, *Envir Health Perspect* 103:680-683, 1995, where the authors predict, based on extensive experience in the National Toxicology Program, "that less than 5-10% of the 75,000 chemicals in commercial use might be reasonably anticipated to be carcinogenic to humans." (p. 680).