Journal of Environmental and Sustainability Law

Missouri Environmental Law and Policy Review
Volume 8
Issue 2 *2001-2002*

Article 2

2001

# Junk Science, Environmental Risk, and Very Low Detection Levels in NPDES Permits

Todd B. Adams

Barry Michael Levine

# JUNK SCIENCE, ENVIRONMENTAL RISK, AND VERY LOW DETECTION LEVELS IN NPDES PERMITS

Todd B. Adams * & Barry Michael Levine**

The dangers of "junk science[1]" have dominated recent legal discussion about scientific evidence inside[2] and outside[3] the courtroom. This emphasis on the dangers of poor science extends to dischargers emphasizing the dangers of being unfairly penalized because of improper measurement of pollutants in the air, water and ground.[4] Under the Clean Water Act,[5] the U.S. Environmental Protection ("EPA") has drawn one reasonable compromise between protecting the environment and unfairly penalizing dischargers. Recognizing the policy choices[6] inherent within the measurement process justifies, however, a more environmentally protective approach than currently used by EPA.

## I. Introduction to Trace Pollutants and the Clean Water Act

The Clean Water Act regulates the discharge of pollutants to the navigable waters of the United States. It prohibits the discharge of a pollutant from a "point source" into the "navigable waters of the U.S." except in compliance with a National Pollutant Discharge Elimination System ("NPDES") permit.[7] Depending upon the pollutant, the Clean Water Act requires technology based effluent limits[8] and effluent limits based on water quality standards.[9] Dischargers discharging toxic pollutants like PCBs must meet stringent effluent standards based on water quality standards.[10] The States establish the water quality standards necessary to protect the environment.[11] The States then incorporate a water quality based effluent limit ("WQBEL") in the relevant NPDES permit.[12] Science then must determine whether it can reliably detect or quantify[13] the chemical of interest at the WQBEL.

---

* Academic Specialist, Michigan State University, Eli Broad College of Business. J.D., 1984, University of Michigan. I wish to thank Bob Avery, Ph.D., Kathleen Brewer, Amy Cook, Bill Creal, Hector Galbraith, Ph.D., Carol Smith, Terry Walkington, and Lisa Williams, Ph.D. who have helped me understand as much of environmental science as I do. Any errors, of course, are mine alone. The opinions expressed herein reflect solely those of the authors and not any other institution or person, including Michigan State University, the Michigan Department of Attorney General and the State of Michigan.

** Partner, Braun Kendrick Finkbeiner. Ph.D., University of Michigan, 1993; J.D., Emory University, 1983.

1 "Junk Science" is science that has little scientific credibility and often done by "fringe" scientists with few credentials in the relevant field. Peter Huber, *Galileo's Revenge: Junk Science in the Courtroom* 14. (Basic Books 1991). In contrast, science requires "replication, verification, and peer review, the patient development of consensus, the systematic weeding, pruning, and uprooting of spurious data and erroneous theory." *Id.* at 209. Criticism of "junk science" is not new. "Such is the respect paid to science that the most absurd opinions may become current, provided they are expressed in language, the sound of which recalls some well-known scientific phrase." James Clerk Maxwell, *quoted in* John Ziman, *An Introduction to Science Studies: The Philosophical and Social Aspects of Science and Technology* 1 (1984).

2 *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 119 S.Ct. 1167, 143 L.Ed.2d 238 (1999); *Daubert v. Merrell Dow Pharmaceutical*, 509 U.S. 579, 113 S.Ct. 2786, 125 L.Ed.2d 469 (1993).

3 David L. Faigman, *Legal Alchemy* (1999); Kenneth R. Foster & Peter W. Huber, *Judging Science: Scientific Knowledge and the Federal Courts* (1999); and Carl F. Cranor, *Regulating Toxic Substances: A Philosophy of Science and the Law* (1993).

4 Steven J. Koorse, *False Positives, Detection Limits, and Other Laboratory Imperfections: The Regulatory Implications*, 19 Envtl. L. Rep. 10,212, 10.215 (1989).

5 33 U.S.C. §§ 1251-1387 (1999).

6 Faigman labels the issue as a "legal policy" question. Faigman, *Legal Alchemy* at 195. *and see* Wendy E. Wagner, *The Science Charade in Toxic Risk Regulation*, 95 Colum. L. Rev. 1613, 1617 & 1719; Rob Hoppe & Aat Peterse, *Handling Frozen Fire: Political Culture and Risk Management* 1 (1993).

7 33 U.S.C. § 1311(A).

8 33 U.S.C. §§ 1314(b)(1)(B) & (2)(B).

9 33 U.S.C. § 1313(c)(2). "[W]ater quality standards should, wherever attainable, provide for water quality for the protection and propagation of fish, shellfish and wildlife, and for recreation, and consider and value for public water supplies, propagation of fish and wildlife, recreational purposes, and agricultural, industrial, and other purposes, including navigation." 40 C.F.R. § 131.2.

10 33 U.S.C. § 1317(a)(1).

11 33 U.S.C. § 1313(c)(2) & 40 C.F.R. § 131.2.

12 33 U.S.C. § 1311(b)(1)(c). For an overview of this process, see Patton, Boggs & Blow, *Environmental Law Handbook* 166-179 (1994).

13 "Three fundamental processes characterize chemical measurements: detection, identification, and quantification." Lloyd Currie, *Preface* vii in *Detection in Analytical Chemistry, Importance, Theory and Practice* (Lloyd Currie, ed.)(ACS Symposium Series 361)(1988) ("ACS Analytical Chemistry"). This is analogous to using photographs from spy satellites to identify enemy tanks in a war zone. *See* Ziman, *Reliable Knowledge* at

## II. CHEMICAL ANALYSIS OF TRACE POLLUTANTS

Trace amounts of PCBs can harm fish and wildlife because of their bioaccumulative qualities, persistence in the environment, and toxicity.[14] Over twenty years after the U.S. banned the manufacture of PCBs[15] and strictly regulated their use,[16] PCBs still pose a danger to the environment.[17] The water quality standard and resultant WQBEL for PCBs are orders of magnitude below the ability of science to detect and quantify PCBs.[18] This places a premium on understanding the method of chemical analysis and the terms used.

A "detection level" means that the analyst knows the pollutant is there, but not at what concentration.[19] The analyst sees PCBs.[20] In contrast,[21] a "quantifiable" limit is one where the analyst both knows the pollutant is present and has some degree of confidence it is present at a specific concentration.[22] The analyst not only sees the PCBs, but also is confident at a specified level that PCBs are present in a certain amount. A certain minimal uncertainty exists for chemical analysis[23] of very low levels of PCBs. Noise is unavoidable.[24] The laboratory may focus on determining whether PCBs

---

85-87 (photography used as an analogy). The analyst must first determine whether he or she sees something other artificial in the photograph, then identify it as a tank. The last stage is to count the number of tanks. Obviously clouds, camouflage, the limits of the camera's resolution and mistakes all limit the reliability of the analyst's count of tanks in the photograph. Statistical analysis of the technique and of a particular analyst's past performance can provide an estimate of the accuracy and reliability of any particular photographic analysis.

Science has improved the detection of both tanks and chemicals over the years. "Detection limits are constantly changing. Over the last twenty years, engineers and chemists have worked to design instruments that can detect and quantify smaller concentrations of compounds." David Elias & Robert C. Goodman, *When Nothing Is Something: Understanding Detection Limits*, 13 Natural Resources & Environment 519, 520 (1999).

[14] *Environmental Defense Fund v. EPA*, 636 F.2d 1267, 1271-72 (D.C. Cir. 1980). The evidence that PCBs cause adverse human health effects is more uncertain. Renate D. Kimbrough, *The Human Health Effects of Polychlorinated Biphenyls*, in *Phantom Risk: Scientific Inference and the Law* 211-228 (Kenneth R. Foster, David E. Bernstein, & Peter W. Huber eds.)(1993).

[15] 15 U.S.C. § 2605(e)(3)(A)(i) (1999) (prohibiting the manufacture of PCBs after January 1, 1979).

[16] 40 C.F.R., part 761 (1999).

[17] In 1995, EPA stated "scientists and public leaders have reached a general consensus that the presence of environmentally persistent, bioaccumulative contaminants is a serious environmental threat to the Great Lakes Basin Ecosystem." 60 Fed. Reg. 15,366, 15,367 (1995) (Final Water Quality Guidance for the Great Lakes System)("Final GLI").

[18] The current EPA ambient water criterion for PCBs in navigable waters is .001 parts per billion. 40 C.F.R. § 129.105(a)(4). In Michigan, the PCB water quality standard for PCBs as a class is 0.000026 parts per billion. Mich. Admin. Code R. 323.1057 (Table 8: Human Cancer Values for the Protection of Human Health). This value becomes the WQBEL in the NPDES permit. *Id.* at 1211(1). Science can currently detect PCBs, however, only at much higher levels. Elias & Goodman, *When Nothing Is Something: Understanding Detection Limits*, 13 Nat. Res. & Env. at 519.

[19] EPA defines the "Method Detection Limit" ("MDL") as "the minimum concentration of a substance that can be measured and reported with 99% confidence that the analyte concentration is greater than zero." 40 C.F.R., Part 136, appendix B.

[20] An analyst must recognize the pattern of PCBs spikes on a gas chromatograph by comparing the chromatograph of the sample with established chromatographs of PCBs. 40 C.F.R., Part 136, Appendix A, Method 608; *see*, for example, *id.* at figures 4-10 (PCB chromatographs). If the pattern matches, then the analyst has found PCBs. *Id.* at § 12.5. Comparison of analytical results to a standard is common in science. John Ziman, *Reliable Knowledge: An exploration of the grounds for belief in science*, 43 (1978).

Actually performing the analysis is, of course, much more difficult. A laboratory must prepare a sample for analysis in order to reduce confounding noise from chemicals other than the one for which the sample is being analyzed. 40 C.F.R., Part 136, Appendix A, Method 608, §§ 2.2 & 3. The laboratory then calibrates its gas chromatograph/mass spectrometer ("GS/MC") prior by analyzing samples or "standards" with known levels of PCBs and provided by sources external to the laboratory. *Id.* at § 7.2. In technical language, "[c]alibration is the establishment of a quantitative relationship between the response of the analytical procedure and the concentration of the target analyte." Berger, *Environmental Laboratory Data Evaluation* at 2-4. If the analytical results meet certain statistically based standards, then the laboratory proceeds to analyze the samples with unknown levels of PCBs. After running the sample through the GC/MS, an analyst then examines the chromatographs of the unknown sample to determine whether PCBs were present or "detected" in the sample and then to quantify the level of PCBs. 40 C.F.R., Part 136, Appendix A, Method 608, §§ 12 & 13. If the laboratory analyst recognizes a PCB pattern in the sample without being able to quantify the level of PCBs because the PCB level is below the level for which the particular GC/MS is calibrated, then the laboratory analyst "detects" PCBs. *See id.* at § 13.3 ("Quantitate every individual peak unless interference persists after cleanup.").

[21] The term "detection" has been used in quantifiable limits causing linguistic confusion. Lloyd Currie, the editor of the American Chemical Society's Detection in Analytical Chemistry, has written: "We have noted that detection limits dictated by regulatory concerns have been surrounded by considerable confusion, discrepant statistical and ad hoc formulations, ignorance, and even mild deception."

[22] For example, EPA has followed the American Chemical Society and defined the "Limit of Quantitation" as "the concentration above which quantitative results can be obtained with a specific degree of confidence." 64 Fed. Reg. 30,417, 30,425 (1999). EPA has promulgated quantifiable "Minimum Levels" ("ML") and a method for determining quantifiable "Interim Minimum Levels" ("IML"). Water Quality Guidance for the Great Lakes System, Supplementary Information Document ("SID"), 419-20 (EPA March 1995).

[23] "It is a commonplace of elementary scientific method that every experimental result is subject to some degree of uncertainty. . . . It is a fundamental principle of statistics that a phenomenon that occurs, say, once in a million trials *on the average* will not be observed once *every* million trials. The results of any experiment that runs for a finite time (e.g. to collect and analyze ten million photographs) are thus subject to significant

are present instead of determining the precise level of PCBs.[25] Theoretical uncertainties limit the certainty of any analytical result.[26] Errors in collecting and handling the evidence will increase error rates further."[27] Preparing a sample for analysis may affect the results.[28] Confirmation bias, when the analyst finds PCBs because the analyst expects or wants to find PCBs in a sample, is also a danger.[29] Even the most gifted analyst will make mistakes.[30] The individual analyst should, therefore, "minimize, but not to underrate, the noise in the data."[31] The political and legal communities can then decide what to do with the data.[32]

## III. STATISTICS & ANALYTICAL CHEMISTRY

### *A. Significance Testing and Confidence Intervals*

Mistakes and the unavoidable uncertainties of chemical analysis means that perfect proof does not exist in science any more than it exists in the law. In science, statistics are used to describe errors and to help control them.[33] Scientists and regulators use statistics to define "detection level" and "quantification levels" in analytical chemistry.[34] Both definitions require a certain level of statistical certainty that requires some explanation. There is no one definition of statistical certainty in science.[35] For one thing, there are many ways to describe a group of analytical results or other data.[36] Furthermore, if not understood or incompletely presented, statistics can be misleading.[37]

---

statistical fluctuations. and can never be precisely reproduced from one experimental run to the next." Ziman, *Reliable Knowledge* at 64-65, and *see also* Crummett, *ACS Analytical Chemistry* at 293-4 (comments).

[24] Noise is especially troublesome when scientists seek to observe very small effects. Ziman, *Reliable Knowledge* at 67. Ziman gives the following example about the use of a large aluminum cylinder to detect and measure gravity waves. "The vibrations of this cylinder [designed to detect gravity waves] were then recorded by instruments so sensitive that they could detect length variations of the order of one-millionth of the diameter of a single atom. Naturally, the observed output of this apparatus is fluctuating 'noise' due to innumerable small disturbances that could not be completely screened out. Thus, the question whether any particular wiggle might be due to the passing of a gravitational wave through the antenna cannot be decided by mere inspection of the record." *Id.* Foster and Huber give a less esoteric example. "A simple example of such a question would be the claim that a gasoline additive can improve a car's mileage by a millionth of a mile (about a thirtieth of an inch) per gallon. The claim is scientific in Popper's sense, but it isn't scientific in the world of real engines, real tires, and real roads—in practice, the experimental errors are far larger than the claimed effect." Foster and Huber, *Judging Science* at 55.

[25] "[T]he trade-off between sensitivity and specificity [is] inherent in test systems." I Zwieg, M, "Establishing Clinical Detection Limits of Laboratory Tests," *ACS Analytical Chemistry* at 151. Zweig uses the example of early radar systems to demonstrate this point. A very sensitive radar system will warn the operator of almost all incoming planes but it will also warn the operator of incoming birds. *Id.* A radar system set to avoid almost all misidentifications of birds as planes will, however, also not detect some incoming planes. *Id.* If the incoming planes are enemy bombers, the consequences are severe for the defenders. The tragic downing of an Iraqi airliner by an U.S. naval vessel demonstrates these dangers.

[26] "The raw data must be refined, processed, analyzed and interpreted by each research worker in his own laboratory, before they can be made sufficiently compact and sufficiently interesting, for onward transmission. These processes are, themselves, heavily laden with theory, and deeply embedded in the current scheme of thought." Ziman, *Reliable Knowledge* at 70.

[27] *See* Foster & Huber, *Judging Science* at 95 (DNA).

[28] The techniques used to separate the desired signal from the noise are very important in highly instrumented science, and cannot be ignored in any assessment of the ultimate reliability of scientific knowledge . . . . Ziman, *Reliable Knowledge* at 66. "But these techniques may not be so efficient when we look for a signal that may not be there at all . . . . In very sensitive instruments the experimental uncertainties can take on an active role . . . . In very sensitive instruments the unavoidable and errors of observation are magnified into an apparently autonomous random disturbance, impishly impeding the honest search for truth. *Id.*

[29] "This phenomenon is called *confirmation bias.* Sometimes confirmation bias may lead to egregious miscarriages of justice—when a detective, for example, fails to collect potentially exculpatory evidence about a suspect. Scientists can easily fall into similar traps, collecting easy results that support what is already believed rather than doing hard analysis that might contradict accepted truths." Foster & Huber, *Judging Science* at 45 (emphasis in original).

[30] "[O]ur main point here is to notice the uncertainties and errors that can accompany good experimental technique by highly competent research workers, and the necessity of independent replication and verification if we are to acquire reliable/ empirical knowledge concerning the external world." Ziman, *Reliable Knowledge* at 75-76.

[31] *Id.* at 70.

[32] Ziman, *Introduction to the Study of Science* at 178-180.

[33] *See* Caulcutt & Boddy, *Statistical Analysis for Chemists* at 1-4 (description of meaning and uses of statistics in analytical chemistry).

[34] *See supra* notes 18 & 19.

[35] "Scientists constantly argue over the criteria for rejecting the null hypothesis (i.e., for concluding that there is a real difference between the groups being compared). Foster & Huber, *Judging Science* at 77.

[36] Caulcott & Boddy, *Statistical Analysis for Chemists* at 1-2.

Scientists may use significance tests in evaluating analytical data. "The purpose of a significance test is to draw a conclusion about a population using data from a sample."[38] For example, an analyst determines that six samples have a mean of *x* when the true value of the solution from which the samples were drawn was *y*. If the difference between *x* and *y* is large enough, then a significance test might show that chance would result in such a difference only 1% or 5% of the time.[39] Use of a confidence interval also provides useful information. A confidence interval consists of an interval centered on the sample mean with a confidence level.[40] In less technical language, it is a range interval that covers the true value of the pollutant in the pollution a specified percentage of the time.[41] In the example above, a confidence interval would be that a person is 95% or 99% confident that the analyst's bias lies between *x* plus or minus some value *z*.[42] Political pollsters use this familiar format to report their results. Both statistical significance tests and confidence intervals can provide useful information about the reliability of data. Statistics also provides further useful information for the legal and regulatory system.[43]

### *B. False Positives Versus False Negatives*

"False positives" and "false negatives" are by now familiar scientific jargon. In the context of detecting and quantifying trace pollutants, false positives are finding the trace pollutant where it is not or at least not above the WQBEL.[44] False negatives are not finding the trace pollutant above the WQBEL where it is above the WQBEL.[45] False negatives and false positives have, of course, different consequences for the environment and dischargers. This means that setting a compliance level requires a policy judgment on how to balance the risk of finding the trace pollutant where there is none. False positives entail risk to both dischargers and environmentalists. Dischargers, of course, have a vital interest in the measurement of PCBs and other trace pollutants. Finding PCBs where there are none can trigger the expenditure of millions of dollars in control equipment[46] and possibly require the involved discharger to pay fines.[47]

Environmentalists also have a stake in the money spent in the form of lost opportunity costs. Money spent unnecessarily to stop PCB pollution cannot be spent on stopping other pollution.[48] Even money spent unnecessarily

---

[37] *Id.*
[38] *Id.* at 34.
[39] "Most scientists require that the difference between control and exposed groups in a study be 'statistically significant' at the $p = 0.05$ level. This means that there is a probability of less than 5 percent—according to the statistical tests used—that the investigators would have recorded a difference as large in the sample if the populations from which the groups were drawn were the same with respect to the properties being compared. A difference that is statistically significant has a low probability of being a statistical fluke: the two populations probably are different." Foster & Huber, *Reliable Knowledge* at 77. *See also*, Caulcott & Boddy, *Statistical Analysis for Chemists* at 33-35. "In most scientific work . . . .05 or 5%. The probability that the null hypothesis will be rejected incorrectly, assuming the null hypothesis is true . . . while . . . the 5 % criterion is typical. reporting of more stringent 1% significance tests or less significant 10% tests can also provide useful information." Daniel L. Rubinfeld, *"Reference Guide on Multiple Regression," in Reference Manual on Scientific Evidence* 415, 431 (Federal Judicial Institute, 1990)(multiple regression analysis).
[40] Caulcott & Boddy, *Statistical Analysis for Chemists* at 38.
[41] David H. Kaye & David A. Freedman, *"Reference Guide on Statistics,"* in *Reference Manual*, 331, 396.
[42] Caulcott & Boddy, *Statistical Analysis for Chemists* at 38-39.
[43] *Daubert*, 509 U.S. at 594 ("[T]he court ordinarily should consider the known or potential rate of error.")
[44] *See* Foster & Huber, *Judging Science* at 75-76 (false negatives and false positives in general).
[45] *Id.*
[46] The issue of false positives permeates environmental law.
[47] These dangers may have been overstated. "The apparent deception [resulting from the misuse of the term detection] is related to the lack of general understanding or agreement concerning the appropriate nature and magnitude of the error of the second kind ( . . . false negative). By ignoring its presence, whether intentional or not, those who must meet regulatory demands generate a [false positive/false negative] imbalance where . . . false negatives may exceed false positives by nearly a factor of 400." Currie, *Detection: Overview of Historical, Societal and Technical Issues*, in *ACS Analytical Chemistry* in 38. Emphasizing the danger of petitioners being unjustly fined ("false positive") to the exclusion of the danger of petitioners discharging injurious pollutants without being discovered ("false negative") is incorrect. "False-negatives, a particular concern of public health authorities and regulators, may indicate that a health-based water quality standard has been met when in fact it has been exceeded." William M. Draper, et al., *90 J. Am. Water Works Ass'n* 82, 83 (June 1998).
[48] Foster and Huber provide an example of lost opportunity costs in controlling radon. "If EPA's numbers are right [that radon causes 20,000 excess lung cancer deaths every year], radon is one of the most serious environmental health threats in the country. But studies in counties with high radon levels have consistently failed to find any direct evidence of excess deaths from radon. Billions of dollars of remediation expenditures thus ride on a presumed effect too small to detect with any scientific and statistical tools currently available." Foster & Huber, *Judging Science* at 57 (note omitted).

measuring PCBs causes lost opportunity costs for controlling PCBs or other pollutants.[49] Setting compliance levels for trace pollutants that cause environmental damage below levels of reliable measurement requires a policy decision about the relative dangers of false positives and false negatives. Those dangers cannot be evaluated in the abstract.

### *C. Base Rates*

More than the error rate inherent to the test affects the likelihood of a particular erroneous analytical result being a false positive or a false negative. The true ratio of violations to non-violations, known as the base rate, may significantly affect the ratio of false positives to false negatives in compliance monitoring. Most simply stated, if violations occur very rarely,[50] then the likelihood that any particular erroneous result will be a false positive instead of a false negative dramatically increase.[51] In the abstract, a thousand samples would have 25 false positives and 25 false negatives at the 95% confidence level. If the true rate of compliance was 98%, however, then the ratio changes. Of the 980 times the plant complies, there will be about 49 false positives.[52] Of the 20 times the plant violated the standard, there would be 1 false negative.[53] Of the 58 violations found, over 70% of them would be incorrect.[54] If the test has a 99% confidence level of correctly finding compliance and also of correctly finding non-compliance, however, the ratio is about 10 false positives[55] out of 30 violations.[56] Base rates powerfully affect the ratio of false positives to false negatives. Any regulatory approach should take base rates into account in some manner.

## IV. THE REGULATORY APPROACH

Determining the proper regulatory approach for PCBs and other trace pollutants that are environmentally harmful at levels that cannot be reliably measure is obviously difficult. EPA has established a reasonable regulatory approach that balances the danger to the environment with the rights of dischargers.

### *A. Determining a Reasonable Potential That the Discharge Will Contain the Pollutant*

The previous discussion shows that false positives may adversely impact dischargers and to a lesser extent the environment. EPA regulations minimize the danger of false positives by requiring compliance with a WQBEL only for pollutants that have a "reasonable potential" to be discharged. The discharger must submit a fact sheet concerning the proposed discharge at the beginning of the NPDES permit process.[57] From this fact sheet and other information,[58] the permitting authority must determine whether there is a "reasonable potential to cause, or contribute to an excursion" beyond environmentally protective levels.[59] If there is, then the permitting authority must include a WQBEL.[60] This ensures not only that the discharger will not waste resources testing for substances probably not present but also those

49 *American Iron and Steel Institute v. EPA*, 115 F.3d 979, 1004 (D.C. Cir. 1997) (per curiam). "It is all very well to insist that we should go on taking pictures [measurements] until the case is proved 'beyond peradventure' (whatever that means!). Such perfection may cost enormously in money, and in time. The aim of the research is to produce a publishable scientific result, of adequate plausibility, not complete proof." John Ziman, *Reliable Knowledge: An exploration of the grounds for belief in science* 65 (1978).

50 "Tests, even quite good ones, will yield unreliable results—far less reliable than intuition suggests—if used to screen for rare events." Foster & Huber, *Judging Science* at 119.

51 "Predictive value clearly depends on both the qualities of the test itself (sensitivity and specificity) and the base rate. The predictive value is the ratio of true positives (individuals who test positive and who really are infected, for example) to the total number of people who test positive, whether correctly or incorrectly." Foster & Huber, *Judging Science* at 115.

52 980 x .05 = 49.

53 20 x .05 = 1.

54 49/68 = .73.

55 980 x .01 = 9.8 false positives. There would be a .2 chance of a false negative. 20 x .01 = .2.

56 Assuming that all 20 of the times that the plant violated the compliance limit were detected.

57 40 C.F.R., Part 132, Appendix F, Procedure 5 (for the Great Lakes System). For an overview of the NPDES Process, see Patton, Boggs & Blow, *Environmental Law Handbook* at 166-179.

58 This includes monitoring information but also any "valid, relevant, representative information." 40 C.F.R., Part 132, Appendix F, procedure 8, such as intake pollutants, existing controls and sources of the pollutant. *Id.* 40 C.F.R. §§ 122.44(d)(ii) & 122.45(h)(2).

59 40 C.F.R., Part 132, Appendix F, at Procedure 8 (Great Lakes Basin).

60 40 C.F.R. § 122.44(d)(i).

false positives will not result in fines. After all, if an analyst takes enough samples, there will be false positives even without confirmation bias.[61] Ensuring that a reasonable potential exists for the discharge of the pollutant reduces the possibility of a false positive.

### *B. Protecting the Discharger by Reducing False Positives Under the EPA Regulations.*

Seeing PCBs at the detection level only means that PCBs are present in the sample. It provides no reliable information about the concentration of PCBs in the sample.[62] On the other hand, seeing PCBs at the quantification level provides reliable information about the concentration of PCBs in the sample and, therefore, about whether the WQBEL is violated. Accordingly, EPA has promulgated rules to protect dischargers from false positives associated with the detection level.

Where the WQBEL is too low to reliably measure, EPA makes a violation of the quantification level—and not of the detection level—a violation of the permit. "When a WQBEL is below the [quantifiable] ML, one cannot make a definitive statement as to whether or not the concentration of the pollutant in the effluent is above or ½ below the WQBEL."[63] EPA explicitly rejected use of the detection level in order to protect dischargers from false positives. "EPA rejected the use of the MDL and other non-quantifiable concentration levels because these concentrations, by definition, do not represent concentrations that are both reproducible and quantifiable indicators of the actual concentration of a given sample, and hence are not reliable measures for permit compliance purposes."[64] EPA has also left discretion for the permitting authority to determine the confidence level necessary to establish a violation.[65] This allows the permitting authority to take into consideration other evidence of a violation and issues of base rates.

### *C. Pollutant Minimization Plan*

Compliance with a quantification limit above the WQBEL does not protect the environment. In such circumstances, EPA requires dischargers to institute a pollutant minimization plan ("PMP").[66] The goal of a PMP is "to maintain the effluent at or below the WQBEL."[67] The PMP requires monitoring, submission of a control strategy, and "implementation of appropriate, cost-effective control measures."[68] EPA has fully protected the rights of dischargers in its regulations through use of the quantification level. It has also protected the environment through requiring pollutant minimization plans when the WQBEL cannot be reliably measured. As discussed below, however, EPA can take further regulatory steps to protect the environment without unfairly penalizing dischargers.

## V. SCIENCE AND THE LEGAL BURDENS OF PROOF

EPA has established a reasonable regulatory approach for trace pollutants such as PCBs that cause environmental damage below levels of reliable measurement. Placing analytical results back in context supports, however, a more environmentally protective approach without unjustly penalizing dischargers.

### *A. The Scientific and Legal Processes are More Similar than Often Thought*

Much of the recent literature emphasizes the difference between the scientific process and the legal process to the detriment of the legal process.[69] The literature sometimes underestimates, however, the effectiveness of the adversary

[61] *See supra* section III.B.
[62] *See supra* notes 18 & 19.
[63] 58 Fed. Reg. 20,802, 20,978 (1993) (Proposed GLI).
[64] SID at 419.
[65] *American Iron and Steel Institute*, 115 F.3d at 994.
[66] 40 C.F.R., Part 132, appendix F, procedure 8.D. (for Great Lakes System).
[67] *Id.*
[68] *Id.* at 8.D.1-6.
[69] Jasanoff, *Science at the Bar: Law Science, and Technology in America* at 5-6. For example, the writers of one article stated: "Apparently, at the level at which a scientist might barely entertain an argument, the lawyer will convict a criminal defendant beyond a reasonable doubt." Walter M. Gawlak & Daniel M. Byrd, *Divergent approaches to Uncertainty in Risk Assessment: Mathematical Expression compared to Circumstantial*

system[70] and the similarities between the legal and scientific system. "[T]here are important differences between the quest for truth in the courtroom and the quest for truth in the laboratory."[71] In particular, "[s]cientific conclusions are subject to perpetual revision" while legal disputes must be "resolve[d] . . . finally and quickly."[72] This difference is indisputable when comparing science as a philosophical endeavor to find the truth to law as a dispute resolution mechanism. Others draw a sharp philosophical distinction: "[S]cience seeks truth, while the law does justice; science is descriptive, but the law is prescriptive; science emphasizes progress, whereas the law emphasizes process."[73] Replication, verification, peer review, and the patient development of consensus almost completely separate science from the law for some.[74]

These differences are less acute and there significance for the courtroom less apparent, however, when we compare both science and the law as philosophical endeavors[75] and science and the law as practical endeavors.[76] Philosophically, the law no less than science seeks the truth because individual justice fundamentally depends on finding the correct facts.[77] The law relies on logic and evidence in seeking justice.[78] Furthermore, if the law is not viewed as a single case, but as a process, law is self-correcting to some degree over the long term. Legal rules that do not work and bad decisions are abandoned.[79]

Unlike science, the law cannot usually rely on the precision of mathematics or consensus to resolve issues, but this makes the law more difficult without distinguishing it in kind. Aesthetics, history, and cultural needs shape science[80] as much as they do the law. Scientific authority, just like legal authority, may partially arise from chance[81] and inhibit new ideas.[82] Science as much as the law may be an endless process.[83] Second and more importantly for purposes of this paper, science in practice has the same difficulties as law in practice: It is impossible to "wring the last drops of uncertainty from what scientists call their knowledge. . . . [B]oth as individual scientists and historically, is that we only arrive at partial and incomplete truths."[84] Time may constrain science less than the law,[85] but individual scientists, like any other human being, face time and monetary constraints. Nor are scientists immune from human failings.[86] "Theories . . . shape perception."[87] Emotions can powerfully affect scientific theories.[88] Even the greatest scientist is not a

*Evidence*, in *Uncertainty in Risk Assessment, Risk Management, and Decision Making*, 39, 43 (Vincent T. Covello, Lester B. Lave, Alan Moghissi, and V.R.R. Uppuluri, eds. 1987).

[70] *Daubert*, 509 U.S. at 596.

[71] *Id.* at 596-597.

[72] *Id.* at 597.

[73] Jasanoff, *Science at the Bar*, 7 (characterizing the criticism of the law).

[74] *See supra* note 1.

[75] Jasanoff, *Science at the Bar* at 8.

[76] "The approach I take here assumes the contrary [to science is an autonomous and largely self-regulating field of inquiry] that scientific claims, especially those that are implicated in legal controversies, are highly contested, contingent on particular localized circumstances, and freighted with buried presumptions about the social world in which they are deployed." *Id.* at xiv.

[77] Jasanoff argues that "The law also seeks to establish facts correctly, but only as an adjunct to its transcendent objective of settling disputes fairly and efficiently." *Id.* at 9. While it is true that the courts sometimes ignores probative evidence, especially in criminal proceedings, in order to serve other societal purposes and rely on less stringent standards of proof than science, *id.* at 10, this is an overstatement. Most people would agree that finding the truth is a key to doing justice.

[78] *American Iron and Steel Institute*, 115 F.3d at 1005 (model cannot be arbitrary).

[79] This is true with regard to science. *See* William W. Scwarzer, "Management of Expert Evidence," *Reference Manual*, 15 (comparing "mature torts" with developing torts.). This point has been made more generally as well. Benjamin N. Cardozo, *The Nature of the Judicial Process*, 178-179 (1921).

[80] Ziman, *An Introduction to Science Studies* at 191-2.

[81] *Id.* at 77 (1984) ("the *Matthew* effect" of "lucky" early success substantially contributing to later success).

[82] *Id.* at 79, & *see Daubert*, 509 U.S. at 596 ("[Petitioners] suggest that recognition of a screening role for the judge that allows for the exclusion of 'invalid' evidence will sanction a stifling and repressive scientific orthodoxy and will be inimical to the search for truth.").

[83] Karl R. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge* at 6 (1963).

[84] J.M. Ziman, *Public Knowledge: An Essay Concerning the Social Dimension of Science*, 5 (1968) & Wagner, *The Scientific Charade*, 95 Colum. L. Rev. at 663.

[85] Jasanoff, *Science at the Bar* at 9.

[86] "Sophisticated equipment and training don't guarantee that subjectivity and self-deception won't creep into an experiment. Nor is there any guarantee that a substantial fraction of the scientific community may not, for a time, be taken in as well." Alan Cromer, *Uncommon Sense: The Heretical Nature of Science*, 171 (1993).

[87] Bert Black, *A Unified Theory of Scientific Evidence*, 56 Fordham L. Rev. 595, 619 (1988).

[88] *See, for example*, Stephen Jay Gould, *The Mismeasure of Man, passim* (1996) (how racism affected research into human physiology and intelligence).

completely disinterested observer reasoning from pure facts.[89] Sharp philosophical distinctions between science and science and the law seem, therefore, unjustified from a philosophical and a practical perspective. If "junk science" poses a danger in the measurement of pollutants or in the courtroom generally, the main causes would seem to lie elsewhere.

### B. Science and Law as Historical Endeavors.

Science and the law appear remarkably similar when trying to do the same thing: understand past events. This suggests that the most important problem with "junk science" is the law's perceived inability to properly evaluate scientific evidence and the burden of proof. Science and the law can appear remarkably similar when trying to understand past events. Science has carefully sought out facts and constructed theories about why the Titanic sank over 80 years ago. Access to the shipwrecks and computer modeling have brought the answer closer without resolving the mystery. Nor has science been able to answer conclusively why certain airliners crashed. Absent a change in the laws of nature[90] and a change in morality about experimenting on human subjects where the subjects run an avoidable risk of injury or death,[91] science cannot finally answer eliminate all uncertainty any more than the law can.

The law's perceived inability to properly evaluate scientific evidence rests on two solid grounds. First, juries and courts often do not have enough scientific knowledge to understand and discriminate between competing scientific theories.[92] Modern science and statistics often tend toward the impenetrable. A court need not understand, however, every nuance of a scientific dispute.[93] It need only understand enough to make a reasonable decision.[94] This is, of course, by no means an easy task.[95] Education and the use of special masters are two tried and obvious ways to help courts evaluate scientific evidence.[96] Unless science is to remain a "black box" accessible only to a scientific elite despite its importance and implications,[97] then judges and juries will have to muddle their way through as best they can.[98] Second, the burdens of proof differ between science and the law.[99] Science disproves a "null hypothesis" with more or less certainty.[100] A scientific theory gains acceptance by eliminating competing explanations, providing a convincing explanation, and usually by providing testable hypothesis that distinguish it from the competing explanations.[101] Science

[89] Gerald Holton, *Einstein, History, and Other Passions*, 58-61 (1996) (describing Einstein's apparently erroneous reliance on his "little finger" for rejecting "the fundamentality of probabilism in physics"); & Kurt Hubner, *Critique of Scientific Reason*, 51-71 (1983).
[90] Stephen Jay Gould, *Wonderful Life: The Burgess Shale and the Nature of History*, 287 (1987) (alternate versions of history only possible in the movies)).
[91] *See* Carl G. Hempel, *Philosophy of Natural Science*, 6 (1966) (experiments in 1800s to determine the cause of puerperal fever that would be regarded as criminal today).
[92] Jasanoff, *Science at the Bar* at 5-7. "The courts do not have the institutional resources to be statisticians." D.H. Kaye. *"Ruminations on Jurimetrics: Hypergeometiric Confusion in the Fifth Circuit,"* 26 Jurimetrics J., 215, 222 (1986).
[93] *In Re: TMI Litigation*, 193 F.3d 613, 628 (3d Cir. 1999).
[94] *See id.* at 727, n. 179 (reliance on common sense), *but see*, Wagner, *The Scientific Charade*, 95 Colum. L. Rev., 1717 (most judges admit "scientific incompetence").
[95] The *TMI* Court needed over 25 pages to do so. *In Re: TMI Litigation*, 193 F.3d at 628-655.
[96] *See* William W. Schwarzer, "Managing Expert Evidence," *in Reference Manual*, 16 (how judge should approach scientific evidence): Joe S. Cecil and Thomas E. Willinger, "Court Appointed Experts," *id.* at 525; and Margeret G. Farrell. *"Special Masters," Id.* at 575.
[97] Jasanoff, *Science at the Bar*, 232, & Ziman, *Teaching and Learning about Science and Society*, 53 (1980). Ziman. *An Introduction to Science Studies* at 160-61 & 190.
[98] "[Courts] reasoned that such decisions [about carcinogenicity] are usually made at the 'frontiers of scientific knowledge.' They are, to a significant degree, policy judgments rather than findings of fact, and they are therefore entitled to great judicial deference." Ronald Brickman, Sheila Jasanoff & Thomas Ilgen, *Controlling Chemicals: The Policy of Regulation in Europe and United States*, 121 (1985). For a criticism of how the courts defer too much to administrative agencies, see Wagner, *The Science Charade*, 95 Colum. L. Rev. at 1664-65, n. 186.
[99] *See*, e.g., Faigman, *Legal Alchemy*, x (Difficulty of reconciling "uncertainties of science" with use of science in courts, administrative agencies, legislatures) & Ziman, *An introduction to Science Studies* at 113. "Unfortunately, the scientist's concept of statistical error does not translate directly into the judge's concept of legal error. We cannot say, therefore, that a study that is statistically significant at the .05 level of confidence will lead judges, if they admit the evidence, to make only five errors (of the Type I variety [false positive]) out of one hundred. Hence, there is no true correspondence between statistical confidence and legal burdens of proof." Faigman, *Legal Alchemy*, 68. Science searches for comprehensive understanding . . . A trial seeks to resolve a focused legal dispute in a finite period of time." Foster & Huber, *Judging Science* at 17. Cranor extensively discusses the differing burdens of proof in science and the law in his book. Cranor, *Regulating Toxic Substances* at *passim*.
[100] Foster & Huber, *Judging Science* at 49.
[101] *See supra* note 1.

generally demands very strong evidence in support of a theory.[102] Scientific investigation often results, therefore, in not proving anything and continuing uncertainty.[103]

Neither science nor the law can afford, however, to ignore uncertain information in all circumstances. For example, it was "largely a matter of conjecture what would happen if there were a major reactor accident" during the development of nuclear power.[104] Refusing to provide any scientific information on the dangers of a major reactor accident because of uncertainty is, however, "an antisocial attitude, since it would effectively deny access to whatever relevant information might have been gleaned in the course of research, however uncertain or controversial."[105] The question is whether requiring a higher burden of proof for scientific evidence in the courtroom would better serve society than a lower standard of proof. This requires a careful evaluation of the scientific process involved and the dangers of error.[106] What it does not suggest is that a particular scientific or statistical imprimatur should be required for introduction of scientific evidence in the courtroom or administrative agency.[107] This is true in analytical chemistry.

### C. *Can the Adversarial Process Work with Very Low Detection Levels?*

Scientists often fundamentally distrust the ability of the adversarial process to work with scientific evidence.[108] The inability to form a scientific consensus in a courtroom, the cultural clash between scientists trained to carefully qualify their findings and lawyers rewarded for convincing others there is one truth,[109] the precision of mathematics compared to language[110] all play legitimate roles in this distrust. The adversarial process can serve effectively, however, in identifying and correcting errors in the detection, identification and quantification of very low levels of pollutants.

Errors in analyzing very low levels of pollutants can occur many places in the process.[111] With proper

[102] "Reliable evidence and logical consistency are the two basic requirements for achieving such a consensus [of informed opinion]. These place a tight constraint on new knowledge. and scientists tend to be closed-minded about claims that aren't so constrained. This occasionally causes science to miss real knowledge for a time. as in the case of parity violation. But it helps guard against being swamped by a flood of nonexistent phenomena that result from the egocentric tendency toward self-deception." Cromer, *Uncommon Sense* at 171. The Supreme Court found that these differences might prevent juries from learning of "authentic" new scientific insights in order to avoid relying on "[c]onjectures that are incorrect." *Daubert*, 509 U.S. at 596-97.

[103] The uncertainty is especially high with regard to suspected carcinogens like PCBs.
Alyson C. Flourney. *Legislating Inaction: Asking the Wrong Questions in Protective Environmental Decisionmaking*, 15 Harv. Envtl. L. Rev. 327, 333 (1991).

[104] Ziman. *An Introduction to Science Studies* at 179. and *see also*. Thomas O. McGarity, *Substantive and Procedural Discretion in Administrative Resolution of Science Policy Questions: Regulating Carcinogens in EPA and OSHA*, 67 Geo. L.J. 729, 808-9 (1979).

[105] *Id.*

[106] Cranor argues that "justice requires that priority be given to avoiding false positives and underregulation." Cranor, *Regulating Toxic Substances* at 152-153. "In tort law we should not be so concerned to have evidence of harm from toxic torts that is scientifically defensible in the best journals that the wrongful loss of a person's good health goes uncompensated. In administrative agencies we should not be so concerned to develop biologically correct models for assessing each risk in question that we identify few carcinogens or assess few of the risks of those we have identified.' *Id.* at 178. He is unable, however, to identify a philosophical theory that fully supports his position. *See id.* at 155-178. Utilitarianism, Rawls. Daniels. and distributively sensitive consequentialism all fail because they cannot support Cranor's position. *Id.* Is Cranor, despite admirable intentions and considerable insights. wrong or are utilitarianism, Rawls, Daniels. and the others wrong?

To write that there should "be no material impairment of health, to ensure that in *most circumstances* a person will not have to choose between having a job or having good health," *id.* at 175 (emphasis added). begs the question of what "most circumstances" means. As much as we would like to be certain that false negatives have fewer costs than false positives, spending millions of dollars because of false positives may adversely affect lives as much as false negatives may.

[107] *Kumho*. 526 U.S. at 246.

[108] Jasanoff. *Science at the Bar* at 4.

[109] Scientists are not alone in their distrust of the adversarial system. "[T]he rhetoric of absoluteness increases the likelihood of conflict and inhibits the sort of dialogue that is increasingly important in a pluralistic society. In the common enterprise of ordering our lives together, much depends on communication. reason-giving. and mutual understanding. Even the legal profession is beginning to question the utility and legitimacy of the traditional adoption of extreme positions by lawyers." Mary Ann Glendon, *Rights Talk: The Impoverishment of Political Discourse*, 44-45 (1991). Nor are scientists the first. "Socrates: Now which type of persuasion does oratory produce in law courts . . . ? The one that results in being convinced without knowing or the one that results in knowing? Georgias: It's obvious. surely. it's the one that results in conviction." Plato. *Georgias* (Donald J. Zeyl. trans.). in *Plato: Complete Works*, 800 (1997).

[110] *See*. Ziman. *Reliable Knowledge* at 11-13 (difference between spoken language and mathematics).

[111] *See supra* section II.

precautions, however, dischargers can identify and respond to the possibility of errors.[112] Experts hired by dischargers can challenge the interpretation of gas chromatographs and laboratory quality assurance and quality control procedures done by administrative agencies and plaintiffs.[113] The uncertainties involved in extrapolation from animal studies to human beings or from epidemiological studies to causes are not present. The adversarial system can handle the issues raised by very low detection, identification and quantification levels very well.

## VI. CHANGING THE REGULATORY APPROACH TO BETTER PROTECT THE ENVIRONMENT WITHOUT UNFAIRLY PENALIZING DISCHARGERS.

EPA has adopted a reasonable regulatory approach to controlling discharges of PCBs that cannot be reliably measured by science at very low levels. Placing the issue back into context, however, supports additional environmentally protective measures that will not unfairly penalize dischargers. First and foremost, the regulatory approach should distinguish between violations of the quantification level and of the WQBEL. Compliance with the WQBEL protects the environment whereas compliance with the quantification level does not necessarily protect the environment.[114] If an analyst is certain that the discharge sample violates the WQBEL with a high degree of confidence, then uncertainty about how much it violates the WQBEL and the exact amount of PCBs in the sample become relatively unimportant. The environment is being harmed contrary to the statutory mandate.

Confidence intervals provide this type of information better than do significance tests.[115] Confidence intervals do so by providing a range of values to use for comparison purposes[116] instead of a single value. The range of a confidence interval may be very broad and yet still exceed the WQBEL at the lowest point of the range. This would support a violation. Conversely if the WQBEL is within the confidence interval or no confidence interval can be drawn, then a court should examine other evidence more closely before determining that a violation exists.[117]

Second, the regulatory approach should distinguish between discharges to waters whose uses are already impaired by PCBs and discharges to waters whose uses are not impaired within the meaning of the Clean Water Act. EPA regulations prohibit discharges that "cause or contribute" to the violation of water quality standards.[118] Finding PCBs at the "detection level" means that the discharger is putting PCBs into waters that already have too many PCBs. The ecosystem will have to cleanse itself of that many more PCBs thereby increasing the danger from a false negative increases. The danger from a false positive remains the same. This change in balance supports placing the burden on a discharger to show that the WQBEL is not being violated when PCBs are found at the detection level.[119] The industrial discharger can take precautions when the samples are taken, take additional samples, evaluate and test internal waste or process streams, or show how its PMP prevents a violation to meet its burden.[120]

---

[112] Koorse, *False Positives*, 19 Env. L. Rev. at 10220-22 (detailed advice to dischargers on how to reduce and protect themselves from uncertainty). and David Faigman, Elise Porter, & Michael Saks, *Check Your Cystal Ball at the Courthouse Door, Please: Exploring the past, Understanding the Present, and Worrying about the Future of Scientific Evidence*, 15 Cardozo L. Rev. 1799, 1834 (1994). (provides a list of how to assure accurate data including the credentials and experience of the technician, specifics of data collection, general reputation and track record of the laboratory that produced the data, specifics of any sample collection, the extent to which the technique relies on the interpretation of an 'expert.'")

[113] *See supra* note 109.

[114] *See supra* note 15.

[115] A result at the quantification level should, of course, be a violation.

[116] "They should focus on P-values (or better yet interval estimates) and they should concern themselves with the reasonableness of the underlying probability model that gives rise to these numbers" and not force into a "poorly conceived and ill-defined mold of a standard deviation analysis." D.H. Kaye at 26 *Jurimetrics Journal* at 223 (about employment statistics).

[117] One court has rejected the use of confidence intervals to defeat an allegation of violation because the Clean Water Act requires dischargers to submit accurate information. *U.S. v. Aluminum Company of America*, 824 F.Supp. 640, 649 (E.D. Tex. 1993). "To the extent that a permit holder hires a laboratory that produces inaccurate and unreliable test results, the permit holder has failed to fulfill its monitoring requirement in direct violation of the Act. Erroneous laboratory results yield the same result as if no monitoring had been performed at all: the government, citizen-plaintiffs, and the courts have no way of knowing whether discharge violations have occurred or not." *Public Interest Research Group of New Jersey, Inc v ELF Atochem North America*, 817 F. Supp. 1164, 1179 (D.N.J. 1993). Such a rigid approach might lessen the burdens on courts and administrative agencies in the short run, but in the long run will probably not result in better data. Dischargers may simply report the same data differently in order to avoid foreclosing their options.

[118] 40 C.F.R. § 122.4(i).

[119] This is consistent with the Restatement (Second) of Torts § 433A (1963) (burden of proof is on joint tortfeasors of indistinguishable harm). Allocation of the burden of proof is "often decisive" in these contexts. Flourney, *Legislating Inaction*, 15 Harv. Envtl. L. Rev. at 384.

[120] *See supra* section IV.

Third, the regulatory approach and courts should consider evidence about base rates in determining the level of statistical confidence necessary for finding that a sample violates the WQBEL. If a discharger has a history of numerous discharges violating the WQBEL, then the practical danger of false positives recedes and that of false negatives advances.[121] Conversely, if the discharger has a history of compliance, then more statistical confidence might be required for finding a violation based solely on one sample.

## CONCLUSION

Even in the hard science of measurement there are difficult policy choices that science alone cannot resolve.[122] EPA has made one reasonable choice but other choices are both environmentally protective and fair to dischargers. The threat of these choices being labeled "junk science should not dissuade EPA from adopting them.

[121] *See* Kaye. "Reference Manual on Statistics." *Reference Manual* at 386-7 & Michael O. Finkelstein & Bruce Levin, *Statistics for Lawyers*, 94 (1990) (review of the discussion about using Baye's Theorem in evidence) & Michael O. Finkelstein and William B. Fairley, *A Bayesian Approach to Identification Evidence*, 83 Harv. L. Rev. 489 (1970) (which started the discussion).

[122] "The decision to accept or reject a hypothesis on the basis of statistical considerations requires judgment, which, to a lay observer, may be obscured by the apparent precision of the analysis. . . . And where the line [between 'significant' and 'not significant'] lies is not a scientific matter at all . . . nor does it describe the magnitude of the effect or the uncertainties in the measurement." Foster and Huber, *Judging Science* at 78.