REGULATORY SCIENCE IN RULEMAKING AND TORT:
UNIFYING THE WEIGHT OF THE EVIDENCE
APPROACH

THOMAS O. McGARTY†
SIDNEY A. SHAPIRO††

We live in a highly technocratic era, in which both common
law courts and regulatory agencies must draw heavily upon
science for critical information needed to resolve important legal
issues. State common law courts frequently require scientific
evidence in lawsuits in which plaintiffs claim that they have
suffered adverse health effects from a drug, a product they
encountered in the workplace, emissions of a toxic air pollutant
from a nearby industrial facility, or contamination of their
drinking water.¹ In these “toxic tort” cases, scientific information
in the form of clinical studies and/or epidemiological evidence is
essential to resolving the critical question of causation.² Clinical
studies involve controlled exposures to volunteer human beings,
while scientists in epidemiological studies engage in sophisticated
statistical analyses of human populations that have been exposed
to drugs, food, other products, or environmental contaminants.³
Expert testimony is the vehicle through which nonscientist judges
and jurors become educated about the implications of the
scientific information relevant to the causation issue. Regulatory
agencies rely on the same clinical and epidemiological studies in
reaching conclusions about whether environmental contaminants
are endangering human health, or whether the benefits of a drug

† Joe R. and Teresa Lozano Long Endowed Chair in Administrative Law, University of Texas School of Law.
†† University Chair in Law, Wake Forest University School of Law.
² See id. § 5:4.
³ Id. § 5:25–26; see also id. § 5:39 (explaining clinical trials in more detail).
or consumer product outweigh the risks. Controlled studies with laboratory animals of the toxicity of chemicals—the ways in which they are metabolized and the mechanisms of action—can add to the biological plausibility of apparent cause-effect relationships, and hence can add to the confidence with which agencies draw conclusions about their existence.

In the common law and regulatory contexts, the intellectual process of reaching conclusions about cause and effect relationships between exposure to potentially toxic substances and human disease is similar. In the regulatory context, the initial inquiry is whether an agency has sufficient evidence to meet the “risk trigger” that Congress has established as a legal prerequisite for the agency to regulate. In the common law context, the inquiry is directed to the question of “general causation.” Both situations involve the same question: Does the relevant scientific evidence adequately support a conclusion that the chemical of concern is capable of causing particular diseases in human beings who are exposed to that chemical at levels approximating those encountered by people?

Having resolved that inquiry in the affirmative, agencies and courts move on to the ultimate issue that they must resolve. In the case of an agency, the issue is what level of protection the agency is required by its legislative mandate to provide to the public. That is, the agency must either determine the permissible level of exposure, in the case of a health standard, or the conditions under which a product may be sold despite potential dangers to the public, in the case of a licensing regime. In the common law context, the court and/or jury must then decide the issue of “specific causation”—whether the plaintiff’s exposure to the chemical in fact caused his or her disease for purposes of compensating the plaintiff for past damage he or she incurred because of that disease.

5. See infra Part I.A.2.
7. MANAGING THE PROCESS, supra note 4, at 18–19.
This article explores the first issue that is faced in common by agencies and courts. We are interested in how an agency decides whether the scientific evidence is sufficient to meet its risk trigger, and how a court decides whether there is sufficient evidence to establish general causation. Our analysis proceeds in four steps.

Section I describes the role of science in law, and how both agencies and courts face a problem of scientific uncertainty. We note the difference between scientific and legal standards for resolving uncertainty and discuss why the legal system requires less proof to accept a causal connection between a chemical and harm to individuals than scientists typically demand.

Drawing on the Environmental Protection Agency’s (EPA) 2008 revision of the national ambient air quality standard for photochemical oxidants (ozone), section II describes the “weight of the evidence” approach that most scientific bodies and regulatory agencies employ in drawing conclusions about the strength of the cause-effect association between human exposure to environmental contaminants and adverse effects on human health. We then contrast the “corpuscular” approach that regulated entities and defendants in toxic tort cases prefer.

Section III discusses a developing divergence between agencies and courts. The federal appellate courts have approved the use of the weight of the evidence approach for the evaluation of scientific evidence by regulatory agencies. By comparison, it is not clear yet whether the same appellate courts will sanction the weight of the evidence approach when employed by experts in toxic tort litigation. A number of district courts have rejected the approach in favor of the corpuscular approach. The issue is whether the Supreme Court’s decision in Daubert v. Merrell Dow

10. See, e.g., Am. Trucking Ass’n v. EPA, 283 F.3d 355, 370 (D.C. Cir. 2002) (“EPA’s inability to guarantee the accuracy or increase the precision of the PM2.5 NAAQS in no way undermines the standards’ validity.”); Pub. Citizen Health Research Grp. v. Tyson, 796 F.2d 1479, 1495 (D.C. Cir. 1986) (“OSHA need only gather evidence from which it can reasonably draw the conclusion it has reached.”).
Pharmaceuticals, Inc. 12 bars the use of the weight of the evidence approach.

Section IV contends that both agencies and courts should apply the weight of the evidence approach. Although regulatory agencies and courts in toxic tort cases use different burdens of proof, the weight of the evidence approach to evaluating scientific evidence is unrelated to the burden of proof. It has to do with the quality of the scientific studies, the strength of the cause-effect associations, the overall consistency of the scientific studies, and the biological plausibility of statistical observations. 13 For this evaluative purpose, we see no meaningful distinction between the regulatory and tort contexts. Moreover, we will show that Daubert does not require the use of the corpuscular approach. To the contrary, the Court’s decision points toward the weight of the evidence approach. One reason we are confident that Daubert does not require the corpuscular approach is that such a requirement would be tantamount to precluding many, if not most, tort victims from recovering in toxic tort cases. Just as employing a corpuscular approach to evaluating scientific evidence by regulatory agencies would represent a policy judgment against protecting the public health, 14 using Daubert to preclude the jury from considering relevant testimony based on a weight of the evidence approach would turn a rule of evidence into a policy decision against compensating people who become ill from exposure to toxic chemicals.

I. THE ROLE OF SCIENCE IN LAW

Both regulatory agencies and courts in toxic tort cases must determine whether there is sufficient relevant scientific evidence to support a conclusion that the chemical of concern is capable of causing particular diseases in people who were exposed to that chemical at levels approximating those these persons actually encountered. 15 In addressing this issue, both the courts and regulatory agencies confront the problem of scientific

13. See Milward, 639 F.3d at 17.
This first-level determination is therefore contestable in both institutional contexts, and a conflict between experts usually results. While scientists are generally unwilling to conclude that a cause-effect relationship exists until scientific uncertainty is essentially eliminated as a matter of statistical probability, the legal system has not required this degree of scientific certainty in either regulatory determinations or in civil cases concerning whether a cause-effect relationship exists. This toleration for uncertainty in the legal system is not necessarily wrong because policy plays a role in resolving factual questions in both regulation and civil litigation. It does, however, make it more difficult for the legal system to sort out when the available scientific evidence is sufficient to support legally supportable conclusions regarding general causation and risk triggers.

A. The Need for Science in Law

Common law courts and regulatory agencies must resolve the common issue of cause and effect under very different legal regimes. In the judicial arena, the issue is whether the plaintiff has satisfied the burden of proof on general causation, and the ultimate decision-maker is ordinarily a jury. In the regulatory world, the issue is whether an agency can meet the burden of proof established by its risk trigger, and the ultimate decision-maker is either a single administrator or a multimember board. Despite these differences, the underlying intellectual exercise is remarkably similar in both contexts.

i. Tort Law

A plaintiff in a toxic tort case ordinarily has a two-part burden of proof on the causation question. The first-level issue is one of general causation: does a chemical cause the disease suffered by the plaintiff in some people? This step requires a plaintiff to establish that it is more likely than not that the chemical at issue causes the disease in some people, taking into account the available scientific evidence and its reliability. The second-level issue is one of specific causation: did the plaintiff’s exposure to the chemical cause the plaintiff to suffer a particular disease? At this step, the plaintiff must establish by a preponderance of the evidence (more likely than not) that the plaintiff’s illness would not have occurred but for the exposure to the chemical. Epidemiological evidence showing that exposure to the chemical more than doubles the risk of the plaintiff’s disease in the exposed population can be sufficient to establish specific causation. Other questions that arise in determining specific causation include whether the plaintiff’s exposure was one of comparable magnitude and duration, whether the plaintiff was exposed differentially to other causal agents for the same disease, and whether the plaintiff’s individual characteristics render him or her more or less susceptible to the disease than the exposed populations in the relevant epidemiological studies.

To prove general and specific causation, a plaintiff is obliged to offer expert testimony. Unlike the connection between a broken leg and a slip-and-fall attributable to a negligently maintained walking surface, judges and juries cannot reach reasonable conclusions about the cause-effect relationship between exposure to a toxic substance and a particular disease by

24. Id. § 28 cmt. c(4).
25. Id.
26. Id. § 28 cmt. c(1).
27. Id. § 28 cmt. c(4).
28. See id.
relying on common sense. The complex relationships between exposure and result are simply beyond the ken of lay judges and jurors. Courts therefore require that the parties educate the fact finder, be it a jury or the trial judge, through experts who are hired by parties to the proceeding and who are cross-examined by attorneys on the other side, often with the assistance of their own experts.30

ii. Regulatory Law

Health, safety, and environmental regulation also involves a two-step process.31 At step one, the agency determines whether the available scientific information meets the requirements of a statutorily prescribed risk trigger.32 When Congress specifies a risk trigger, it ordinarily provides some indication of when risk is sufficiently serious to warrant regulation under the applicable grant of regulatory authority.33 In practice, the risk trigger operates as an evidentiary burden that the agency has to meet in order to regulate a hazard.34 Regulators must demonstrate that the risk to the public or the environment exceeds some threshold, but Congress typically authorizes them to act on the basis of anticipated harm.35 The Clean Air Act (CAA), for example, authorizes the EPA to regulate new stationary sources of air pollution that may cause or contribute to “air pollution which may reasonably be anticipated to endanger public health or welfare.”36

At the second step, an agency determines the level of regulation by using the “statutory standard” that Congress has established.37 These standards vary, but most are precautionary in nature. For example, the CAA requires the EPA to set national primary ambient air quality standards at a level requisite to protect public health with “an adequate margin of safety.”38 Congress has

30. See id.; RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL AND EMOTIONAL HARM § 28 cmt. c(3) (2010).
32. See id.
33. See id. at 33.
34. Id.
35. Id.
37. SHAPIRO & GLICKSMAN, supra note 31, at 32.
38. Id. at 37.
delegated to regulatory agencies and the professionals on their
staffs the determination of whether a particular chemical meets
these statutory criteria. The heads of the regulatory agencies are
rarely experts in toxicology or epidemiology, and, as often as not,
even the scientists on the agency’s staff are not well-
versed in the
science relevant to the particular risks of particular chemicals.
The agencies thus make frequent use of advisory committees
consisting of panels of experts in the particular subject matter of a
given regulatory initiative to provide expert input.

iii. The Common Issue

Courts and agencies address a common issue at the first
step of the analysis that each must undertake. Both must
determine whether a chemical causes a disease in some people at
the levels of exposure they face. Both rely on scientific experts to
help them make this judgment. Yet the two institutions ultimately
employ a different burden of proof on this question. While courts
apply a preponderance of the evidence test, regulators have a
lesser burden of adducing sufficient scientific evidence to meet
the risk trigger. An agency’s determination that there is sufficient
evidence to meet the risk trigger must meet the “arbitrary and
capricious” (or in some cases the “substantial evidence”) test for
judicial review of the agency’s determination. This reflects
Congress’s intent that the regulatory system be more protective
than tort law. In fact, it was the failure of the tort system to
protect people sufficiently from toxic chemicals that caused

40. See U.S. GEN. ACCOUNTING OFFICE, EPA’S SCIENCE ADVISORY BOARD PANELS:
IMPROVED POLICIES AND PROCEDURES NEEDED TO ENSURE INDEPENDENCE AND BALANCE 1
(2001); U.S. GEN. ACCOUNTING OFFICE, FED. ADVISORY COMMS., ADDITIONAL GUIDANCE
COULD HELP AGENCIES BETTER ENSURE INDEPENDENCE AND BALANCE 1 (2004).
41. U.S. GEN. ACCOUNTING OFFICE, EPA’S SCIENCE ADVISORY BOARD PANELS:
IMPROVED POLICIES AND PROCEDURES NEEDED TO ENSURE INDEPENDENCE AND BALANCE 1
(2001); U.S. GEN. ACCOUNTING OFFICE, FED. ADVISORY COMMS., ADDITIONAL GUIDANCE
COULD HELP AGENCIES BETTER ENSURE INDEPENDENCE AND BALANCE 1 (2004).
42. RESTATEMENT (THIRD) OF TORTS: LIIAB. FOR PHYSICAL AND EMOTIONAL HARM §
28 cmt. c(1) (2010).
44. See Susan Rose-Ackerman, REGULATION AND THE LAW OF TORTS, AM. ECON. REV., May
1991, at 54.
Congress to establish the CAA, the Occupational Safety and Health Act (OSHA), and other protective legislation.\textsuperscript{45}

The differences in the burden of proof, however, should not affect the issue of whether the courts, like agencies, should employ a weight of the evidence approach to evaluate the scientific evidence upon which they rely. The issue in tort law should be whether the weight of the scientific evidence supports a general causation conclusion, not whether each study relied upon by the expert passes an independent test of reliability. This issue, which is typically resolved by the court in a separate proceeding prior to the trial, is antecedent to the question whether the scientific evidence, under a weight of the evidence approach, satisfies the plaintiff’s burden of proving that his or her exposure to the chemical at issue caused his or her disease. The burden of proof, in theory, could differ in agencies and courts. Nonetheless, both agencies and courts may properly adopt a weight of the evidence approach to resolving the question whether scientific evidence is sufficiently reliable to support a conclusion of general causation.

B. Information Quality and the Problem of “Junk Science”

One frequently encountered aspect of decision making about cause-effect relationships is the uneven quality of the scientific information available to the decision-maker. In the case of regulatory agencies, much of the available information may be in agency files because it was submitted by regulatees in connection with licensing proceedings, or because the agency staff or a contractor has conducted searches of the available literature.\textsuperscript{46} In addition, the agencies typically invite interested parties to submit relevant information at the outset of a regulatory proceeding.\textsuperscript{47} Not surprisingly, some information thereby acquired is of higher quality than other information, and the agency staff, often with the assistance of scientific advisory committees, must evaluate the quality of such information in determining its reliability for purposes of supporting conclusions

\textsuperscript{45} See SHAPIRO & GLICKSMAN, supra note 31, at 1–2.


\textsuperscript{47} Administrative Procedure Act, 5 U.S.C. § 553(c) (2006).
about the extent to which the relevant chemical crosses the risk threshold.

In tort law, by contrast, common law courts and juries consider scientific evidence that has already been filtered by experts hired by the parties to present their views on whether the relevant chemical meets the plaintiff’s burden of proof on the issues of general and specific causation. These experts, often with the assistance of counsel, evaluate the quality of individual studies and decide whether and how to rely on them in their expert reports or oral testimony. On cross-examination, an attorney for the other side can challenge an expert’s choice and use of particular scientific studies.48

Critics of regulatory agencies frequently complain that they too often fail to base their findings on “sound science,” and critics of common law courts claim that they too often allow juries to rely on “junk science” in forming conclusions about whether chemical exposures caused particular diseases.49 Both claims seek to characterize the relevant science as too weak to support legal action. In reality, this often represents an effort to characterize scientific uncertainty as disqualifying agencies and courts from acting,50 even in circumstances in which they are entitled to act despite the uncertainty.

C. Uncertainty and Burden of Proof

Although both courts and regulatory agencies need scientific input, the available science is all too often plagued with large uncertainties.51 Epidemiological studies, in particular, are inherently subject to uncertainty.52 The scientists undertaking

49. Sound Science, supra note 29, at 900, 905.
50. Id. at 901 (“Both the demand for ‘sound science’ in regulation and the call for the elimination of ‘junk science’ in the courtroom are artfully framed appeals to scientific objectivity that . . . are highly contestable and, indeed, surprisingly vacuous assertions.”).
52. Flournoy, supra note 51, at 334.
these investigations must take care to avoid unintended bias in the results due to errors in study design, limitations on the sources of data, or “confounding factors”—factors that can account for the differences observed in the populations under comparison but that are unrelated to the exposure under study. One of the largest sources of uncertainty in epidemiological studies stems from the difficulty that epidemiologists face in determining accurate estimates of human exposures to a chemical in the workplace or the environment. Exposures vary from individual to individual, and over time and geography, but existing exposure monitoring devices are often quite crude.

Uncertainties also plague the statistical analyses of possible associations between exposures and diseases. The statistical “power” of a study depends on several factors, including the size of the study group, the duration of exposure, and the normal incidence of the disease being investigated in the relevant population. A consistent pattern of similar results in several epidemiological studies ranging across different places, circumstances, and times can reduce the uncertainty involved in drawing conclusions about cause-effect relationships. Results demonstrating a high relative risk for the disease in the most heavily exposed populations also reduce the uncertainty. If clinical or animal studies suggest biologically plausible disease-producing mechanisms that are consistent with the results of the epidemiological study, uncertainties are further reduced.

Scientists typically demand a level of statistical significance that rejects the null hypothesis with a confidence of greater than 95 percent. This reflects the scientific community’s reluctance to draw statistical inferences about cause and effect relationships.

54. See Flournoy, supra note 51, at 334.
55. See id.
56. See McGarity, supra note 14, at 159.
57. Id.
58. Id. at 160.
59. Id.
60. Id.
when the apparent association between exposure and disease may be attributable to chance.\textsuperscript{62}

By comparison, the existence of scientific uncertainty per se is not a reason to dismiss evidence of cause and effect in most legal contexts.\textsuperscript{63} Even under the “beyond a reasonable doubt” test for criminal liability, some degree of uncertainty is tolerable.\textsuperscript{64} Epidemiological investigations simply cannot demonstrate cause-effect relationships with “the compelling certainty that is reached through mathematical reasoning.”\textsuperscript{65} Because there is no perfect epidemiological study, there can always be disagreements about the degree to which any result supports conclusions about cause-effect relationships.\textsuperscript{66} If regulatory agencies had to wait until they were absolutely certain that a pollutant caused an increase in mortality or morbidity before concluding that it crossed the relevant statutory risk threshold, they would never be able to regulate anything. When they do rely on imperfect scientific information in regulatory contexts plagued by large uncertainties, they must accompany that reliance with explanations that can survive judicial review under an “arbitrary and capricious” test or, sometimes, the substantial evidence test, as mentioned earlier.\textsuperscript{67}

Likewise, if experts in toxic tort litigation were confined to relying on studies that were free of uncertainty, no plaintiff would ever recover. As we shall see, the Supreme Court in \textit{Daubert} assigned a screening role to the trial courts that extends beyond the filtering that the experts hired by the parties perform in preparing their testimony.\textsuperscript{68} Section 702 of the Federal Rules of Evidence was redrafted in the wake of cases like \textit{Daubert} to require that trial courts screen out aspects of expert testimony that do not

\textsuperscript{62} McGarity, supra note 14, at 159.


\textsuperscript{66} McGarity, supra note 14, at 165–66.


meet the Supreme Court’s tests for scientific reliability. That role has evolved into one in which the trial judge, in determining whether to admit an expert’s testimony, is free to evaluate not only the quality of the studies upon which the expert relies, but also the degree to which uncertainties in the available scientific information are sufficiently debilitating to render the expert’s conclusions scientifically unreliable for the purpose of proving causation. After Daubert, trial judges have struggled with determining not whether the studies underlying expert testimony are completely free of uncertainty, but whether they are so clouded by uncertainty that they are not scientifically reliable.

II. TWO APPROACHES TO EVALUATING SCIENTIFIC INFORMATION

As regulatory agencies struggled with how imperfect scientific studies should inform legal conclusions about whether, and the extent to which, exposure to particular chemicals exceeded statutory risk triggers, they looked to the scientific community for guidance and discovered that scientists typically apply a weight of the evidence approach in drawing conclusions from available scientific studies about cause-effect relationships. In the context of toxic tort litigation, however, defendants have often succeeded in persuading trial judges to employ a corpuscular approach to scientific studies that demands a great deal of scientific certainty from each of the studies upon which the plaintiff’s expert relies. This section describes the weight of the evidence and corpuscular approaches to evaluating scientific evidence and how they differ. We then illustrate the difference between the two approaches in the context of the EPA’s 2006 revision of the national ambient air quality standard for ozone.

69. Fed. R. Evid. 702 advisory committee’s note.
70. Fed. R. Evid. 702(b).
71. See Soldo v. Sandoz Pharm. Corp., 244 F. Supp. 2d 434, 537 (W.D. Pa. 2003) (labeling expert witnesses’ reliance on case reports as per se unscientific where the case reports discuss causation in terms of possibilities and uncertainties); Reynard v. NEC Corp., 887 F. Supp. 1500, 1505 (M.D. Fla. 1995) (rejecting causation testimony where the articles on which the expert relied stated that the hypothesis was uncertain).
73. Id. at 172.
A. The Weight of the Evidence Approach

Under the weight of the evidence approach, experts—whether they are informing regulatory agencies or the fact finder in a trial—make a cause-effect determination based on the totality of the available scientific evidence. The expert considers all available studies and determines the weight to be afforded to each on the basis of the strengths and weaknesses of the individual studies. A scientific study may be so affected by inadequate protocols, poor execution, or questionable statistical analysis that it is entitled to no weight at all. Other studies may be flawed in one or more regards but nevertheless useful to inform the decision. The expert need not ignore such studies, but he or she may place less weight on seriously flawed studies than on studies with only minor flaws. The EPA has explained the weight of the evidence approach in the context of carcinogen risk assessment as follows:

Judgment about the weight of evidence involves considerations of the quality and adequacy of data and consistency of responses induced by the agent in question. The weight of evidence judgment requires combined input of relevant disciplines. Initial views of one kind of evidence may change significantly when other information is brought to the interpretation. Generally, no single weighing factor on either side determines the overall weight. The factors are not scored mechanically by adding
pluses and minuses; they are judged in combination. 79

At bottom, the weight of the evidence approach is based on “the exercise of scientific judgment grounded in scientific expertise.” 80

B. The Corpuscular Approach

Under the corpuscular approach to drawing cause-effect conclusions, the expert or the agency must establish the “scientific reliability” of each individual study relied upon and each step of the analysis of the scientific information must pass a high threshold of confidence. 81 As a practical matter, the expert bears the burden of “validating” each individual study as well as justifying the overall conclusion as to cause and effect. 82 The overall conclusion thus stands or falls not only on the expert’s overall rationale, but also on the expert’s successful validation of each relied-upon study. 83 The corpuscular approach invites the parties to a proceeding to search the studies that the expert cites for possible problems with the protocols, errors in data collection, and flaws in statistical analysis, and to speculate at length about the possibility of confounding factors and other sources of potential bias. 84 In the case of epidemiological studies, which are never perfect, the search for possible error and bias is nearly always fruitful. 85

C. The Two Approaches in EPA’s Ozone NAAQS

Under the CAA, the EPA must promulgate and periodically revise “primary” national ambient air quality standards (NAAQS) for pollutants that may reasonably be anticipated to endanger
public health and that derive from numerous or diverse mobile or stationary sources. This requirement constitutes the risk trigger. The EPA then establishes a primary NAAQS for the pollutant at an ambient level that is “requisite to protect the public health” while “allowing an adequate margin of safety.” This second step constitutes the regulatory standard.

To make these decisions, Congress has required the EPA to prepare a Criteria Document (CD), which must “accurately reflect the latest scientific knowledge” about the adverse health effects of the pollutant at various ambient levels in the air. The statute also creates an independent Clean Air Scientific Advisory Committee (CASAC) to assist the Administrator in assessing the scientific evidence. The statute directs the agency to “complete a thorough review” of the existing CD every five years and, if necessary, revise the document and the existing standards to reflect scientific information that has become available since the last revision.

One of the original NAAQS was for photochemical oxidants, measured as ozone. The ozone standard protects the public from the adverse health effects of the irritating oxidants (referred to in the early years as “smog”) that result from the combination of volatile organic compounds and oxides of nitrogen in the presence of sunlight. The EPA revised the CD for ozone in 1997. The agency relied on that document in revising

87. See supra Part I.A.2.
88. Id. § 7409(b) (1).
89. See supra Part I.A.2.
90. Id. § 7408(a) (2).
91. Id. § 7409(d) (2). The CASAC is composed of seven members and must include a physician, a member of the National Academy of Sciences, and a person representing state air pollution control agencies. Id. § 7409(d) (2) (A).
92. Id. § 7409(d) (1).
the primary and secondary standards to change both the level of the standard from 0.12 parts per million (ppm) to 0.084 ppm and the period during which the pollution level was averaged from one hour to eight hours.96

i. The 2006 Criteria Document

The most recent revision of the CD was published in February 2006 after two rounds of public comment.97 Drawing on a large number of new studies that had become available since 1996, the CD concluded for the first time that short-term exposure to photochemical oxidants increased human mortality risk.98 This meant that ozone in the ambient air was killing people, although the exact mechanism for bringing about this unfortunate result was unclear.99

The number of short-term mortality studies involving ozone exposure had “increased markedly” since the earlier CD,100 and most of the studies showed positive associations between ozone exposure and mortality.101 Nevertheless, the studies did not eliminate the scientific uncertainty. Consider, for example, a 2004 time-series study by Professor Michelle Bell of ninety-five communities.102 The Bell study was a primary focus of the 2006 CD because it was the most representative of the U.S. population, and because it focused exclusively on the warm season when ozone levels are highest.103 The document recognized that time-series epidemiological studies consisting of consecutive observations of disease patterns in a population over a period of time were

96. Id.
100. 2006 Ozone Criteria Document, supra note 98, at 7-84.
101. Id. at 7-85 to -86 (see Figure 7-14).
103. 2006 Ozone Criteria Document, supra note 98, at 7-86 to -87.
susceptible to bias due to “confounding by daily to seasonal temporal factors.” 104 The statistical analysis in such studies had to correct for strong seasonal cycles in both health outcomes and ozone levels, and they had to avoid the confounding influence of weather and other air pollutants on a daily scale. 105 Several of the more recent studies, like the Bell study, had successfully avoided these confounding factors, but others had not. 106

The CD ultimately concluded that “[w]hile uncertainties remain in some areas, it can be concluded that robust associations have been identified between various measures of daily [ozone] concentrations and increased risk of mortality.” 107 Furthermore, confounding factors such as weather, time, and co-pollutants could not account for the reported results. 108 The few studies that focused on warm season exposures demonstrated even greater increases in mortality risk. 109

ii. The 2006 Staff Paper

As the final touches were being put on the CD, the Office of Air Quality Planning and Standards prepared a “staff paper” for the Administrator, the purpose of which was to “bridge the gap” between the scientific analysis in the Criteria Document and the policy judgments that the Administrator had to make in determining whether to revise the ambient air quality standards. 110 In addition to evaluating the “policy implications” of the “key studies and scientific information” contained in the CD, the paper presented and interpreted several quantitative risk assessments prepared by the staff based on the scientific data and various risk assessment models. 111 Although the ultimate decision remained within the Administrator’s discretion, the staff also presented its

104. Id. at 7-137.
105. Id.
106. Id. at 7-146.
107. Id. at 7-110.
108. Id.
109. Id. at 7-111.
111. Id.
views “on a range of policy options” that the Administrator should consider. In reaching its “evidence-based” conclusions, the staff paper adopted a weight of the evidence approach, placing “greater weight on associations with health endpoints that the CD has judged to be likely causal based on an integrative synthesis of the entire body of evidence, including not only all available epidemiological evidence but also evidence from animal toxicological and controlled human exposure studies.”

Like the CD, the staff paper underwent two rounds of comment from CASAC and the public. The second draft recommended that the Administrator consider setting a new eight-hour standard in the range of 0.064 to 0.084 ppm (the current standard) with a primary focus on 0.070 ppm. The CASAC unanimously concluded that there was “no scientific justification for retaining” the current primary standard and that the new standard should “be substantially reduced to protect human health.” It therefore unanimously disagreed with the draft paper’s recommendation that the Administrator consider retaining the existing standard. Instead, it unanimously recommended limiting the range of options from 0.060 to 0.070 ppm.

Industry groups launched a corpuscular attack on many aspects of both the staff paper and the CASAC assessment. They took particular aim at the new time-series mortality studies, arguing that the ambient air quality monitors relied on in those studies did not provide sufficiently reliable approximations of

112. Id.
113. Id. at 6-2 (emphasis added).
114. EPA Staff Narrows Recommendations for Ways to Tighten Ozone Standard, 37 ENVT REP. (BNA) Issue No. 29, at 1502 (July 21, 2006) [hereinafter Recommendations]; Steven D. Cook, EPA Staff Calls for Risk Assessments of Possible New Federal Ozone Standards, 36 ENVT REP. (BNA) Issue No. 45, at 2333 (Nov. 18, 2005).
117. Id. at 4–5.
118. Id. at 2.
personal exposure. In their view, the reported associations between ozone exposure and mortality in those studies did not represent cause-effect relationships, but instead reflected inappropriately specified statistical models, inadequate attention to confounding factors, and/or publication bias. Since those individual studies were fatally flawed, the staff-prepared quantitative risk assessments that relied on those studies were likewise unreliable. Consequently, industry groups concluded that the EPA could not legitimately rely on either the time-series studies or the risk assessments in deciding whether or not to revise the primary NAAQS for ozone.

Accepting the CASAC’s advice and rejecting industry criticisms, the final version of the staff paper deleted the recommendation that the Administrator consider retaining the existing standard. The paper stated that “we now conclude that the overall body of evidence clearly calls into question the adequacy of the current standard in protecting . . . against an array of adverse health effects,” including mortality. The staff paper acknowledged that the scientific evidence supported a standard as low as 0.060 ppm. A staff-prepared quantitative risk assessment predicted that a standard set at 0.074 ppm would reduce ozone-caused mortality by about 10 to 40 percent. A standard set at 0.070 ppm would reduce ozone-caused mortality by 20 to 55 percent, and a standard set at 0.064 ppm would reduce ozone-

120. National Ambient Air Quality Standards for Ozone, 72 Fed. Reg. 37818. Attaching a personal air quality monitor to individual subjects (or a representative number of subjects) would have yielded a more accurate estimate of ambient ozone exposures than the stationary monitors that were located in various positions on most cities, but they add considerably to the cost of the study.


123. JAN. 2007 EPA STAFF PAPER, supra note 121, at 6-45 to -46.


125. JULY 2007 EPA STAFF PAPER, supra note 110, at 6-46 to -47 (emphasis added); see also National Ambient Air Quality Standards for Ozone, 72 Fed. Reg. 37818, 37881.

126. JULY 2007 EPA STAFF PAPER, supra note 110, at 6-77.

127. Id. at 6-78 to -79.

128. Id. at 6-79 to -80.
caused mortality by 25 to 75 percent. The staff paper therefore extended the lower limit of its recommended range for the primary standard to 0.060 ppm and dropped its previous preference for 0.070 ppm.

iii. The Notice of Proposed Rulemaking

The Notice of Proposed Rulemaking (NPR) that the EPA published in July 2007 proposed to lower the primary standard from 0.084 ppm to between 0.070 and 0.075 ppm. Adopting a weight of the evidence approach, the preamble to the proposed rule drew on “an integrative synthesis of the entire body of evidence, published through early 2006, on human health effects associated with the presence of O$_3$ in the ambient air.” That body of evidence included more than one hundred epidemiologic studies conducted in the U.S. and Canada. The agency also drew on several quantitative risk assessments that the staff had prepared and highlighted in the staff paper. In making judgments about the extent to which relationships between various health endpoints (including mortality) and short-term exposures to ambient O$_3$ were causal, the agency relied on several factors, including “the nature of the evidence (i.e., controlled human exposure, epidemiological, and/or toxicological studies) and the weight of evidence, which takes into account such considerations as biological plausibility, coherence of evidence, strength of association, and consistency of evidence.” Epidemiological data provided “evidence of associations between ambient [ozone] levels and more serious acute and chronic health effects (e.g., hospital admissions and mortality)” that could not be assessed in controlled human exposure studies. The degree of uncertainty introduced into those studies by “confounding variables (e.g., other pollutants, temperature) and other factors,” however, affected the level of

129. Id. at 6-80 to -81.
130. Id. at 6-81 to -82.
132. Id. at 37823 (emphasis added).
133. Id.
134. Id.
135. Id. (emphasis added).
136. Id.
confidence that the identified health effects were attributable to ozone exposures.  

The preamble explained that in using a weight of evidence approach in forming judgments about the degree of confidence that various health effects were likely to have been caused by exposure to ozone, the agency’s confidence increased “as the number of studies consistently reporting a particular health endpoint” grew and as “other factors, such as biological plausibility and strength, consistency, and coherence of evidence” increased.  

Its conclusions regarding biological plausibility, consistency, and coherence of evidence in turn were “drawn from the integration of epidemiological studies with mechanistic information from controlled human exposure studies and animal toxicological studies.” The agency explained that strength of association in epidemiological studies referred both to the magnitude of the association and its statistical strength. Consistency referred to “the persistent finding of an association between exposure and outcome in multiple studies of adequate power in different persons, places, circumstances and times.” In the case of ozone, the agency believed that the “magnitude of effect estimates” were “relatively consistent across recent studies showing association between short-term, but not long-term, [ozone] exposure and mortality.”

As noted, the NPR proposed to revise the level of the primary ozone standard to some level within a range from 0.070 to 0.075 ppm. The administrator recognized that there was “no evidence-based bright line that indicate[d] a single appropriate level” for the standard. Instead there was “a combination of scientific evidence and other information that needs to be considered holistically in making this public health policy judgment, and selecting a standard level from a range of reasonable values.” Because there was no “bright line clearly directing the choice of

137. Id.
138. Id.
139. Id.
140. Id.
141. Id.
142. Id.
143. Id. at 37878.
144. Id. at 37879.
145. Id. (emphasis added).
level within this reasonable range,” the choice of the appropriate level—“considering the strengths and limitations of the evidence, and the appropriate inferences to be drawn from the evidence and the exposure and risk assessments”—was a “public health policy judgment.”

The Administrator believed that a standard within the range of 0.070 to 0.075 ppm “would reduce the risk of a variety of health effects associated with exposure to O₃, including...mortality effects indicated in the epidemiological studies.”

The public comments on the rule fell into two broad groups. Public health and environmental groups supported the EPA’s use of the weight of the evidence approach and agreed with the general conclusions that it reached under that approach, but they chastised the agency for failing to follow the CASAC’s advice and set the primary standard at the low end of the 0.060 to 0.070 ppm range. The commenters from the industry launched a comprehensive corpuscular attack on the additional epidemiological studies on which the EPA had relied in deciding on a standard in the range of 0.070 to 0.075 ppm. Like the industry’s previous attack on the agency’s evidence, the comments found fault with the individual studies. The Utility Air Regulatory Group, for example, argued that the inconsistencies and uncertainties inherent in the multicity studies rendered them unreliable for purposes of supporting a lower level for the standard.

iv. The Final Rule

The EPA published a final rule in late March 2008 that increased the stringency of the primary standard to 0.075 ppm, a

146. Id.
147. Id.
150. See McCarthy, supra note 149, at 1881; Tumey, supra note 149, at 1876.
151. See, e.g., National Ambient Air Quality Standards for Ozone, 73 Fed. Reg. 16436, 16456.
152. Id.
level that was near the upper end of the range proposed by the staff. In concluding that the standard should be changed, the Administrator rejected the industries’ corpuscular attacks on the new scientific studies. Applying the “weight of evidence” approach used in the Criteria Document,” the Administrator concluded that the entire “body of scientific evidence across all types of studies” was very robust and included “a large number of various types of studies that provide consistent and coherent evidence of an array of [ozone]-related respiratory morbidity effects . . . as well as total nonaccidental and cardiorespiratory mortality.”

The Administrator also relied on the weight of the evidence approach in rejecting the contention of public health and environmental groups that the standard should be set in the 0.060 to 0.070 range. The Administrator acknowledged that his conclusion ran directly contrary to the CASAC recommendations. He pointed out, however, that the CASAC’s recommendations themselves appeared to represent “a mixture of scientific and policy considerations.” The Administrator generally agreed with CASAC’s “interpretation of the scientific evidence” but noted that there was “no bright line clearly directing the choice of level.” The choice of the appropriate level was therefore “a public health policy judgment entrusted to the Administrator.” The Administrator explained that he weighed the uncertainties inherent in the risk assessments more heavily than CASAC. In his view, those uncertainties precluded heavy reliance on the predictions in the risk assessments as “a primary basis for concluding that levels at or below 0.070 ppm” were required to protect the public health with an adequate margin of safety.

153. Id. at 16436.
154. Id. at 16479 (emphasis added).
155. Id. at 16482.
156. Id.
157. Id.
158. Id. at 16482–83.
159. Id. at 16483.
160. Id.
D. Scientific Uncertainty and the Weight of the Evidence Approach

The EPA rulemaking illustrates both the problem of scientific uncertainty and the relationship of the weight of the evidence test towards addressing that uncertainty. The EPA had to resolve two issues. First, it had to determine whether a new NAAQS was necessary because exposure to ozone at the exposure level permitted by the earlier regulation could reasonably be anticipated to endanger public health. Second, if so, the agency had to determine what level of exposure was necessary to protect the public health while allowing an adequate margin of safety. Despite the enormous evidentiary record compiled by the EPA, the resolution of both of these issues was complicated by the scientific uncertainty. The first staff paper recommended that it was reasonable to keep the existing standard, although after hearing from the scientific advisory committee, the second staff paper dropped that recommendation. Further, as the administrator indicated, the decision concerning the exposure level was a policy issue because the science was uncertain concerning what level of exposure was necessary to protect the public. The administrator rejected the recommendation of the scientific advisory committee for a lower exposure limit, noting that their recommendation reflected policy considerations, rather than being the product of a scientific judgment.

In addressing these issues, the EPA relied on a weight of the evidence approach and specifically rejected the corpuscular attack by industry interests on the evidence in front of the agency. This choice indicated that, although individual studies had limitations, they were sufficiently reliable to warrant consideration in setting the standard. At the same time, however, the weight of the evidence approach recognized that the weight assigned to various studies would depend on a number of considerations going to their reliability. In other words, some studies were more persuasive than others when it came to

163. See supra text accompanying notes 155–56.
164. See National Ambient Air Quality Standards for Ozone, 73 Fed. Reg. 16436, 16456.
determining whether the EPA had sufficient evidence to meet the statutory requirements for the risk trigger and the risk standard.

As we will develop below, this is a crucial distinction. The corpuscular approach takes the position that if a study has nontrivial defects, it cannot be used as evidence to meet an agency’s (or a plaintiff’s) burden of proof.\textsuperscript{165} The weight of the evidence test asks whether the defects are so great that the study should be disregarded.\textsuperscript{166} Thus, the weight of the evidence test recognizes the inevitability of scientific uncertainty, and it does not automatically reject scientific evidence because a study does not eliminate that uncertainty.

III. REGULATORY AND TORT APPROACHES TO SCIENTIFIC UNCERTAINTY

In both the regulatory and tort contexts, the federal courts have confronted the issue of whether to permit the use of expert judgments based on the weight of the evidence approach. In reviewing regulatory cases, the federal courts have strongly supported the use of the weight of the evidence approach by regulatory agencies.\textsuperscript{167} In tort cases, by comparison, there appears to be a trend among federal district courts to rely on the corpuscular approach.\textsuperscript{168} Bucking this trend, the First Circuit has recently held that \textit{Daubert} does not preclude district courts from allowing juries to consider expert testimony based on a weight of the evidence approach.\textsuperscript{169} The Fifth Circuit, by comparison, has

\begin{itemize}
\item \textsuperscript{165} \textit{Id}.
\item \textsuperscript{166} See supra text accompanying note 80.
\item \textsuperscript{167} E.g., Envtl. Def. v. EPA, 369 F.3d 193 (2d Cir. 2004); MBH Commodity Advisors, Inc. v. Commodity Futures Trading Comm’n, 250 F.3d 1052 (7th Cir. 2001); Dixon v. FAA, 8 F.3d 798 (Fed. Cir. 1993).
\item \textsuperscript{168} See Erica Beecher-Monas, \textit{Blinded By Science: How Judges Avoid the Science in Scientific Evidence}, 71 TEMP. L. REV. 55, 57, 69 (1998) (noting that the courts have frequently read \textit{Daubert} to require them to evaluate each study underlying an expert’s conclusion sequentially to determine admissibility); Lucinda M. Finley, \textit{Guarding the Gate to the Courthouse: How Trial Judges Are Using Their Evidentiary Screening Role to Remake Tort Causation Rules}, 49 DEPAUL L. REV. 335, 336 (1999) (“Judges have applied \textit{Daubert} to subject each item of expert proof proffered by plaintiffs to substantive causation law scrutiny, to see if it, standing alone, would prove both general and specific causation.”).
\end{itemize}
expressed strong reservations about the weight of the evidence approach, although it did not definitively reject it.170

A. Regulatory Agencies

The federal courts have had no difficulty accepting agency reliance on the weight of the evidence approach. A good example is Public Citizen Health Research Group v. Tyson,171 in which the D.C. Circuit Court of Appeals reviewed the OSHA’s health standard for ethylene oxide (EtO), a carcinogen used in sterilizing hospital equipment. The Association of Ethylene Oxide Users (AEOU) claimed that OSHA’s regulation was not supported by substantial evidence because each of the individual scientific studies on which OSHA relied contained some defect or defects and no study alone supported a conclusion that EtO caused cancer in exposed workers.172 In other words, they asked the court to reject the rule because OSHA did not adopt the corpuscular approach in evaluating the evidence before it. The court rejected that approach in the following passage:

[The] AEOU attacks each piece of evidence, suggesting that no individual piece proves a relationship between EtO exposure and various adverse health effects. This approach disregards the marginal contribution that each piece of evidence makes to the total picture. While some of OSHA’s evidence suffers from shortcomings, such incomplete proof is inevitable when the Agency regulates on the frontiers of scientific knowledge . . . . OSHA need not “prove” its assertions in the manner AEOU demands . . . . Rather, OSHA need only gather evidence from which it can reasonably draw the conclusion it has reached . . . . Our function . . . is only to search for substantial evidence, not proof positive.173

172. Id. at 1486, 1495.
173. Id. at 1495.
In American Trucking Associations v. EPA, the D.C. Circuit also showed no sympathy for a corpuscular attack by regulated entities on the 1997 revision of the EPA’s NAAQS for ozone, discussed in the previous section. Industry interests had filed comments pressing the corpuscular approach with particular force. In affirming the EPA’s interpretation of the totality of the evidence on the effects of ground-level ozone, the court instead allowed the EPA to rely upon the weight of the evidence. For example, because the record was “replete with references to studies demonstrating the inadequacies of the old one-hour standard,” the agency was entitled to rely upon the consensus of its scientific advisory committee that, based upon the weight of the evidence, an eight-hour standard was more appropriate for protecting human health than the old one-hour standard. Since the agency adopted a similar weight of the evidence approach in adopting the 2008 revision of the standard, it is unlikely that the D.C. Circuit, which has exclusive jurisdiction over challenges to NAAQS, will overturn that standard on corpuscular grounds.

B. Federal Courts in Torts Cases

It is unclear at this time whether the federal courts of appeal will sanction the use of the weight of the evidence approach. Milward is the first opinion of which we are aware to accept this methodology as consistent with Daubert. As noted below, Allen suggests the methodology is inappropriate for tort law.

174. Am. Trucking Ass’ns, Inc. v. EPA, 283 F.3d 355 (D.C. Cir. 2002). The case was on remand from the Supreme Court after various industry groups challenged the EPA’s revised NAAQS. After the Court rejected the plaintiffs’ contention that section 109 of the CAA, 42 U.S.C. § 7409 (2000), required the EPA to engage in cost-benefit balancing in setting the NAAQS and ruled that their contention was unconstitutional if there were no such requirement, the Court remanded the case back to the D.C. Circuit. See Whitman v. Am. Trucking Ass’ns, Inc. 531 U.S. 457 (2001). At this point, the D.C. Circuit took up the plaintiffs’ challenges to the scientific groundings of the agency’s new standards, including their challenge that the EPA should have used the corpuscular approach. The D.C. Circuit upheld these standards against this and all pending challenges.

175. Am. Trucking Ass’ns, Inc. v. EPA, 283 F.3d at 375.
176. Id. at 378.
177. Id.
In *Milward*, Brian and Linda Milward sued Acuity Products and other chemical companies, claiming that Brian’s routine exposure as a refrigerator technician to benzene-containing products that had been manufactured or supplied by the defendants had caused his rare type of leukemia. The district court rejected the testimony of the plaintiffs’ expert, Dr. Martyn Smith—on whom the plaintiffs relied to prove general causation—and dismissed the case.

Smith had applied a weight of the evidence methodology, first developed by Sir Author Bradford Hill in 1965, that remains the state of the art method in modern epidemiology. His testimony explained how, taking these factors into consideration, he reached the judgment that benzene exposure can cause Milward’s type of cancer. The appellate court reversed the district court and ruled that Dr. Smith’s opinion rested on “a scientifically sound and methodologically reliable foundation, as is required by *Daubert*.”

The First Circuit found that the District Court had committed three errors in rejecting Smith’s testimony. First, the district court had disagreed with some of the reasoning that he used to apply the various factors in the methodology. The appellate court objected that the district court had engaged in the weighing of the evidence, a function that properly belongs to the

---

179. *Id.*
180. *Id.* at 17; see *id.* (noting that the Bradford Hill weight of the evidence criteria ask the epidemiologist to evaluate nine aspects of the available epidemiological studies, including: “the strength or frequency of the association; the consistency of the association in varied circumstances; the specificity of the association; the temporal relationship between the disease and the posited cause; the dose response curve between them; the biological plausibility of the causal explanation given existing scientific knowledge; the coherence of the explanation with generally known facts about the disease; the experimental data that relates to it; and the existence of analogous causal relationships”); Arthur Bradford Hill, *The Environment and Disease: Association or Causation?*, 58 PROC. ROYAL SOC’Y MED. 295, 295–99 (1965). The EPA boiled these nine criteria down into three broader criteria: strength of association, consistency, and biological plausibility; *supra* text accompanying notes 138–41.
182. *Id.*
183. *Id.* at 20–25.
184. *Id.* at 20–21.
jury as fact-finder. Second, the district court had rejected Smith’s testimony because there was no reliable statistical evidence linking exposure to benzene to Milward’s type of cancer. This was to be expected, the appellate court indicated, because the rare nature of the disease meant that too few people suffered from it to develop statistically reliable evidence of causation.

Finally, the appellate court noted that the district court “at times” “treated the separate evidentiary components of Dr. Smith’s analysis atomistically, as though his ultimate opinion was independently supported by each.” This was mistaken because Smith was applying a weight of the evidence approach:

In Dr. Smith’s weight of the evidence approach, no body of evidence was itself treated as justifying an inference of causation. Rather, each body of evidence was treated as grounds for the subsidiary conclusion that it would, if combined with other evidence, support a causal inference. The district court erred in reasoning that because no one line of evidence supported a reliable inference of causation, an inference of causation based on the totality of the evidence was unreliable. The hallmark of the weight of the evidence approach is reasoning to the best explanation for all of the available evidence.

ii. Allen v. Smith Engineering

In Allen, the Administrator of the estate of Walter Allen, a hospital worker who had died of brain cancer, brought a products liability action against a manufacturer of EtO, to which Mr. Allen had been exposed while working at the hospital as a maintenance worker. Although regulatory agencies classify EtO as a carcinogen, the appellate court affirmed the district court’s ruling

---

185. Id. at 22.
186. Id. at 23.
187. Id. at 24.
188. Id. at 23.
189. Id.
that, under *Daubert*, the testimony of the plaintiff’s three witnesses
was inadmissible because of a lack of sufficient scientific
grounding.191

The appellate court agreed with the district court that
existing scientific evidence linked EtO to some forms of cancer,
but not to brain cancer.192 It noted that:

\[\text{[T]he public health agencies acted at least partly on}
\text{the basis of epidemiological studies that showed a}
\text{relationship between EtO exposure and other kinds}
\text{of human cancer, so their use of a “weight of the}
\text{evidence” methodology was grounded in stronger}
\text{probative evidence than appellants’ experts have}
\text{adduced to show a link between EtO and brain}
\text{cancer.}^{193}\]

Moreover, the plaintiff also lacked evidence of the level of
Allen’s exposure to EtO, which indicated a failure to establish
specific causation.194

Taking these points into account, the case is not necessarily
a rejection of the weight of the evidence approach. The *Milward*
court read the opinion this way.195 But the court did express its
skepticism about that approach. It observed that regulatory
agencies are entitled to use a weight of the evidence approach
because they operate under statutes that require them to act in a
preventative manner.196 By comparison, the burden of proof is
greater in a tort case:

This [weight of the evidence] methodology results
from the preventive perspective that the agencies
adopt in order to reduce public exposure to
harmful substances. The agencies’ threshold of
proof is reasonably lower than that appropriate in
tort law, which “traditionally make[s] more
We think this argument is mistaken, an issue that we will take up in the next section.

IV. THE PROPER ROLE FOR SCIENTIFIC INFORMATION

The jury is still out, so to speak, on whether the federal courts will apply divergent approaches to the evaluation of scientific evidence in regulatory and tort cases. We see no grounds for any such divergence. The weight of the evidence approach, used by regulatory agencies with the approval of the appellate courts, should also be admitted to prove general causation in toxic tort cases, assuming that the expert’s testimony on the weight of the evidence is sufficiently grounded in the scientific evidence.

While regulatory agencies and trial courts in tort cases apply different burdens of proof, this difference is not relevant to the decision of whether or not to consider expert judgments based on the weight of the evidence approach. Moreover, we see nothing in the Daubert case to preclude reliance on the weight of the evidence approach, and there is language in Daubert that suggests its approval of this approach. Finally, judicial adoption of the corpuscular approach in common law litigation is tantamount to precluding many, if not most, victims from ever recovering in toxic tort cases because it adopts a scientific standard of reliability as the legal standard of reliability.

A. The Issue is Admissibility, Not Burden of Proof

The Allen court suggested that the weight of the evidence approach would be inappropriate in tort cases because agencies have a lesser burden of proof than does a plaintiff in a toxic tort case. But the issue is admissibility, not burden of proof. In this regard, agencies and courts are engaged in the same intellectual

197. Id.
198. Compare McGarity, supra note 14, at 172, with id. at 176–77.
199. Allen, 102 F.3d at 198.
exercise: evaluating the scientific reliability of a cause-effect conclusion based on multiple scientific studies. Insofar as science is supposed to be the determining factor, the outcome should be the same in both contexts. Both common law courts and regulatory agencies should consider expert opinion based on weight of the evidence evaluations of the available scientific information in accordance with valid scientific criteria, such as the Bradford Hill criteria, for evaluating evidence. Other experts applying the same weight of the evidence criteria may evaluate the same information differently, as the NAAQS case study demonstrated. But this goes to burden of proof, not the admissibility of the scientific evaluations.

An agency must determine whether the available scientific evidence is of sufficient quality and reliability to support a conclusion that the risk posed by the chemical at issue exceeds the statutory risk trigger.\footnote{200} Congress has established risk triggers that permit an agency to act in a precautionary manner in the face of uncertainty.\footnote{201} In evaluating the scientific evidence for this purpose, regulatory agencies consider the evaluations of experts who employ the weight of the evidence approach, and the agencies themselves employ that approach in evaluating scientific evidence.\footnote{202} Note that the willingness to evaluate the evidence according to the weight of the evidence is not a conclusion in any given instance that the evidence under consideration is sufficient to meet the risk trigger. It merely reflects the agency’s refusal to ignore available studies because they are to some degree imperfect, and its acceptance of some degree of scientific uncertainty in drawing conclusions about cause and effect.\footnote{203}

A court in a toxic tort case must determine whether the plaintiff has offered sufficient evidence to establish general

\footnote{200.} Sidney A. Shapiro, \textit{OMB and the Politicization of Risk Assessment}, 37 ENVTL. L. 1083, 1087 (2007) (“Whether an agency has sufficient scientific evidence to satisfy a statutory risk trigger is a legal issue and not a scientific one. It is a legal issue because Congress intended agencies to make regulatory decisions on the basis of imperfect scientific knowledge.”).

\footnote{201.} \textit{Id.}

\footnote{202.} \textit{See supra} text accompanying note 80.

\footnote{203.} \textit{See Ethyl Corp. v. EPA}, 541 F.2d 1, 24 (D.C. Cir. 1976) (“[R]egulators . . . have not thereby been endowed with a prescience that removes all doubt from their decisionmaking. Rather, speculation, conflicts in evidence, and theoretical extrapolation typify their every action.”).
causation. This is the same question that agencies confront: whether a chemical causes people to become ill at the levels to which they are exposed. The plaintiff must prove general causation by a preponderance of the evidence, which is a higher burden of proof than the “arbitrary and capricious” (or sometimes “substantial evidence”) standard an agency must satisfy in supporting a conclusion that human exposures to a chemical exceed a statutory risk trigger. Both agencies and courts rely on experts for this purpose. The issue in Milward and similar cases is whether Daubert precludes considering expert evaluations based on a weight of the evidence approach. Note that the willingness to accept testimony using the weight of the evidence is not a conclusion that the evidence under consideration is sufficient to meet the plaintiff’s burden of proving general causation by a preponderance of the evidence. The issue for the court, as for an agency, is whether an expert, using the weight of the evidence approach, has an adequate basis for the opinion being given.

Ultimately, a nonexpert decision-maker will have to evaluate the strength of the scientific evidence to support a cause-effect conclusion. In agencies, this is the administrator; and in the courtroom, this is the jury or the judge. Administrators have the benefit of getting advice from scientific advisory committees and from their in-house scientists. Similarly, the jury or judge has the benefit of testimony from experts presented by the plaintiff and the defendant. While these processes are obviously not identical, we do not see any differences between them that would require

---

204. Johnson v. Arkema, Inc., 685 F.3d 452, 468 (5th Cir. 2012) (“[T]here is a two-step process in examining the admissibility of causation evidence in toxic tort cases. First, the district court must determine whether there is general causation. Second, if it concludes that there is admissible general-causation evidence, the district court must determine whether there is admissible specific-causation evidence.”).

205. See supra text accompanying note 179.

206. See, e.g., Newkirk v. ConAgra Foods, Inc., 727 F. Supp. 2d 1006, 1034 (E.D. Wash. 2010) (“Plaintiffs must prove by a preponderance of the evidence . . . general causation . . . .”); State ex rel. Wyo. Workers' Safety & Comp. Div. v. Madeley, 134 P.3d 281, 284 (Wyo. 2006) (“Even if sufficient evidence is found to support the agency's decision under the substantial evidence test, this Court is also required to apply the arbitrary-and-capricious standard as a 'safety net'. . . .”).


screening out testimony in a tort case that relies on a weight of the evidence approach. It may be that administrators are somewhat more capable than juries in evaluating the persuasiveness of expert testimony by virtue of their education and experience with similar issues in other regulatory contexts, but we do not believe that this justifies an approach to scientific evidence that screens out expert judgments based on the weight of the evidence in a jury trial. In any case, Daubert stressed that it was not a ruling intended to protect juries from the difficulty of evaluating admissible, but weak evidence:

Respondent expresses apprehension that... befuddled juries are confounded by absurd and irrational pseudoscientific assertions. In this regard respondent seems to us to be overly pessimistic about the capabilities of the jury and of the adversary system generally. Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.209

We are not against judges, as gatekeepers, protecting juries from testimony for which there is no justifiable basis in science. The protection of the jury, however, does not require a blanket exclusion of the weight of the evidence approach.

B. Daubert Does Not Bar the Weight of the Evidence Approach

Daubert was essentially concerned with excluding expert witnesses who were charlatans—witnesses who would testify in favor of the plaintiff even though the available scientific evidence offered little or no support for their testimony.210 An expert witness, however, does not automatically fall within this category because the expert uses the weight of the evidence approach. This is not to say that all testimony based on the weight of the evidence

210. Id. at 595.
approach should be admitted. The conclusion in Allen that the experts lacked a reasonable basis for their conclusions about general causation may well have been correct if, as the court implied, no study or group of studies of sufficient quality explored the association between EtO and brain cancer. But expert testimony in toxic tort cases should not be excluded solely because the expert relied on the same weight of the evidence approach that regulatory agencies employ.

i. Rule 401

Daubert is an interpretation of Rule 401, which requires that evidence be relevant to the issues in a trial in order to be admitted. The Court notes that relevance “is defined as that which has ‘any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.’” 212 This means, the Court emphasizes, that the “Rules’ basic standard of relevance thus is a liberal one.” 215 In other words, the issue is whether a piece of evidence is a “brick” that a party is entitled to use to build the wall necessary to meet the burden of proof. An individual brick may only be a small portion of the total wall, but as long as it contributes to the building of the wall, it should be admissible.

Since the issue is whether evidence has any tendency to make a determination more probable, it is clear that an expert can use a scientific study even if it has defects. On relevance grounds, a plaintiff is entitled to use any scientific study that makes it “more probable” that the chemical causes people to become ill “than it would be without the evidence.”

This is not to say that Rule 401 has no bite in toxic tort cases. A witness using the weight of the evidence approach must be able to explain the linkage between a judgment on general causation and the studies on which the expert is relying. 214 If the underlying studies do not offer at least some evidentiary support for the expert’s judgment, then the expert’s judgment has no

211. See Fed. R. Evid. 401; Daubert, 509 U.S. at 587.
212. Daubert, 509 U.S. at 587.
213. Id.
tendency to make the determination concerning general causation more (or less) probable. As Chief Justice Rehnquist explained:

[N]othing in either Daubert or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the ipse dixit of the expert. A court may conclude that there is simply too great an analytical gap between the data and the opinion proffered.

At the same time, the relevant issue is whether or not the testimony makes it more (or less) probable that there is general causation. As the Milward court recognized, judgments about whether the evidence ultimately meets the burden of proof belong to the fact-finder.

ii. Rule 701

Daubert also interprets Rule 701, which permits expert testimony as long as the testimony relates to “scientific, technical, or other specialized knowledge [that] will assist the trier of fact to understand the evidence or to determine a fact in issue.” Daubert establishes a series of tests that judges can use to determine whether a study about which an expert testifies can be considered as “scientific” and therefore within this scope of this rule. These inquiries include such issues as whether a study uses the scientific method, whether the studies are replicable, and whether the results have been peer reviewed.

The point of these various tests is to determine whether the expert’s testimony is “more than subjective belief or unsupported

215. See Daubert, 509 U.S. at 596.
216. Joiner, 522 U.S. at 146.
220. Id. at 593–94.
speculation.”[^221] Testimony does not automatically become subjective belief or unsupported speculation because it relies on multiple studies, some or all of which may have some defects. As long as the individual studies can be considered “scientific,” an expert can testify about them.

IV. CONCLUSION

Scientific uncertainty is a fact of life in the life sciences because natural phenomena do not lend themselves to controlled study. Nowhere is this observation truer than in the study of the etiology of human disease, where scientists are often limited to statistical analyses of cause-effect relationships undertaken after the relevant exposures have already taken place. Scientists typically demand a high degree of certainty (typically 95 percent) before they are willing to conclude that a statistical association rises to the level of a cause-effect relationship in any particular study.[^222] But they are also willing to consider the weight of the evidence in making judgments about whether a given chemical is capable of causing a particular disease in humans.[^223]

Nothing in the analytical framework established by Rules 401 and 701 of the Federal Rules of Evidence and the Supreme Court’s opinion in Daubert precludes courts from employing the weight of the evidence approach in toxic tort litigation. And nothing in that analytical framework compels courts to employ the corpuscular approach. Indeed, by requiring each piece of scientific information relied upon by an expert to meet a demanding standard of confidence in toxic tort cases, the corpuscular approach effectively replaces the preponderance of the evidence burden of proof in civil litigation with the 95 percent standard of confidence that scientists apply to individual studies. Since it is highly likely that experts for plaintiffs in most toxic torts cases will not be able to point to a single study that finds a cause-effect relationship with that degree of confidence, general adoption of the corpuscular approach closes the door on such lawsuits at the admissibility stage, long before a jury is allowed to

[^221]: Id. at 590.
[^222]: McGarity, supra note 14, at 212 n.415, 218 n.447.
determine whether the plaintiff has established by a preponderance of the evidence that the chemical is capable of causing his or her disease.

Tort law serves important social purposes, all of which will be defeated in the context of chemically induced diseases if the district courts continue to employ a corpuscular approach and refuse to allow experts to employ the weight of the evidence approach. Among other things, the corpuscular approach will preclude many, if not most, victims from recovering the compensation they deserve in toxic tort cases, and it will deprive the common law of its ability to deter companies from engaging in unreasonably dangerous activities and manufacturing unreasonably dangerous products. We understand that many oppose tort law as it is presently constituted. But any debate about the scope of tort law should be based on a discussion of the functions of tort law in promoting corrective justice and deterrence. Tort reform should not be accomplished through the back door by judicial application of a rule of evidence that demands the impossible.