

Reproduced with permission from Product Safety & Liability Reporter, 40 PSLR 50, 01/09/2012. Copyright © 2012 by The Bureau of National Affairs, Inc. (800-372-1033) <http://www.bna.com>

CAUSATION**TOXIC TORTS**

A pair of recent state supreme court rulings illustrates how courts are still struggling with the challenges presented by the limits of scientific proof of toxic causation of disease, says law Professor Steve C. Gold in this BNA Insight. The author analyzes the rulings—in Texas and Vermont—and finds they pose sharply differing implications for litigants and lower courts.

Revisiting Relative Risk Rules: *Garza, Blanchard*, And the Ever Evolving Role of Epidemiologic Proof in Toxic Tort Cases



BY STEVE C. GOLD

When causation is disputed in a toxic tort case, epidemiology—the study of associations between exposure to suspected risk factors and the incidence of disease—often figures prominently. Either plaintiffs or defendants may invoke epidemiologic studies, depending on the studies' conclusions. Whichever party finds the epidemiologic results unfavorable almost always tries, in some way, to use the circumstances of a particular case to explain away the unfa-

Steve C. Gold is an associate professor of law at Rutgers School of Law in Newark, N.J., where he teaches torts and environmental law. He can be reached at stgold@andromeda.rutgers.edu.

vorable population-based data from epidemiologic studies.

Courts are left to decide the legal significance of epidemiologic data and, when confronted with expert testimony based in part on such data, to distinguish between permissible inference and impermissible speculation, between admissible and inadmissible testimony, between possibly sufficient and legally insufficient evidence. Although courts typically dress their decisions in the language of science, speaking of reliability, analytical gaps, and statistical significance, those decisions say more about judicial philosophy and policy than they do about science.

Two state supreme courts recently had occasion to address the role of epidemiologic proof in light of their own precedents. Their approaches exemplify some of the variation in judicial views of the epidemiologic enterprise.

Texas Supreme Court: *Merck & Co. v. Garza*

In *Merck & Co. v. Garza*,¹ the Texas Supreme Court had the opportunity to revisit its 14-year-old ruling in *Merrell Dow Pharmaceuticals Inc. v. Havner*.² The plaintiffs' decedent, Leonel Garza, died of a heart attack

¹ 54 Tex. Sup. J. 1697 (Tex. 2011).

² 953 S.W.2d 706 (Tex. 1997).

after a long history of heart disease. Twenty-five days before his death, Mr. Garza complained to his doctor of intermittent numbness, pain, and weakness in his left arm. The doctor determined that Mr. Garza was not having a heart attack and prescribed Vioxx for the pain. After a stress test showed a “stable” cardiac status despite small ischemic areas at the tip of Mr. Garza’s heart, he continued to take Vioxx until his sudden death. Mr. Garza’s statutory beneficiaries, who alleged that Vioxx had caused Mr. Garza’s fatal heart attack, won a jury verdict against Merck, the drug’s manufacturer. The Texas Supreme Court reversed.³

The court’s recitation of Mr. Garza’s medical history suggests considerable and perhaps justifiable skepticism of the claim that Mr. Garza would not have suffered a heart attack had he not taken Vioxx. Yet if Vioxx and heart disease both cause heart attacks, the case would seem to present a choice among several plausible causal scenarios—*i.e.*, heart disease caused the heart attack and Vioxx was not involved, or Vioxx caused the heart attack and heart disease was not involved, or the interaction of Vioxx and heart disease caused the heart attack though either one alone would not have.⁴ Framed this way, the causation issue fairly demanded a jury’s decision.

But the Texas Supreme Court concluded that the jury should not have been given the Garzas’ case to decide. Under Texas law, the court held, the Garzas had, in fact, presented no reliable evidence that Vioxx could have caused Mr. Garza’s heart attack.⁵ This conclusion may sound surprising, in light of the Food and Drug Administration’s views regarding cardiac risks of Vioxx and similar drugs.⁶ To understand it requires an understanding of *Havner*.

Texas Precedent: *Havner*

Havner was one of many cases in which plaintiffs alleged that an expectant mother’s ingestion of Bendectin to treat morning sickness caused birth defects in her child. By the time *Havner* was decided, large epidemiologic studies had failed to detect any statistically significant increase in risk of birth defects among mothers who had taken Bendectin as compared to those who had not taken the drug. To try to prove that Bendectin caused birth defects, the plaintiffs relied on the results of *in vivo* and *in vitro* animal studies on reanalyses and interpretation of certain epidemiologic data. The court briefly rejected any inference from the animal studies to Bendectin’s effects in humans, and then engaged in an

³ *Garza*, 54 Tex. Sup. J. at *2-6.

⁴ The allegation that Vioxx caused a heart attack, even in someone with a history of heart disease, is not absurd. Merck’s studies of Vioxx revealed a higher incidence of adverse cardiovascular events “in patients with and without a history of atherosclerotic cardiovascular disease.” *McDarby v. Merck & Co.*, 949 A.2d 223, 234 (N.J. Super. 2008).

⁵ *Garza*, 54 Tex. Sup. J. at *26 (“the Garzas did not present reliable evidence of general causation and are therefore not entitled to recover against Merck”). The court defined general causation as “whether a substance is capable of causing a particular injury or condition in the general population.” *Id.* at *9 (quoting *Havner*, 953 S.W.2d at 714).

⁶ See generally, *e.g.*, Press Release, U.S. Food & Drug Admin., COX 2 Selective (includes Bextra, Celebrex, and Vioxx) and Non Selective Non Steroidal Anti Inflammatory Drugs (NSAIDs) (April 7, 2005), <http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2005/ucm108427.htm>.

extended discussion of the nature of epidemiologic research in relation to legal standards of proof.

First, the *Havner* court noted, correctly, that epidemiologic data, by definition, apply to populations rather than to individual cases of disease.⁷ The court acknowledged, however, that alleged toxic injuries like birth defects do not bear little flags identifying their causes, so it was “persuaded that properly designed and executed epidemiological studies may be part of the evidence supporting causation in a toxic tort case.”⁸ This concession turned out to be small, however, as the court also explained in labored detail what it meant by “properly designed and executed.”

The court stated three requirements quite clearly. First, even the best epidemiologic study would be considered evidence of causation only if it found a relative risk greater than two, *i.e.*, that exposure to the allegedly toxic substance was associated with a risk of disease more than double the risk faced by unexposed populations. Second, that result must be statistically significant at the $P < .05$ level. Third, the sample of people studied must be “similar” to the plaintiff—meaning, among other things, that the plaintiff’s exposure must be at least as high as the exposed study subjects’ exposure.⁹ In somewhat less categorical terms, the court also cautioned against methodological biases to which observational epidemiologic studies are prone and against reliance on a single “isolated” study when subsequent studies failed to replicate the result.¹⁰ Significantly, the court framed its standards for epidemiologic studies as necessary elements of the scientific “reliability” of expert testimony based on them, making them threshold requirements for admissibility.¹¹

Reaffirming Rigidity: The Relative Risk Threshold

The Garzas thought they had *Havner* licked. The scientific data they used to support their claim of general causation came not from observational epidemiologic investigation but from controlled clinical trials, the “gold standard” of such studies.¹² More than one of the trials, as well as a meta-analysis of multiple trials, had found statistically significant relative risks greater than two. Further, although in some ways the subjects in the clinical trials may have been at greater risk from Vioxx than was Mr. Garza (because they took higher doses for longer time periods), in other ways they may have been at less risk, because for the most part the trials excluded people with prior histories of heart disease. Therefore, they argued, their general causation evidence, overall, was sufficient to reach the jury.

The Texas Supreme Court disagreed in a short, unanimous opinion which, unlike *Havner*, was bereft of citation to scientific literature and legal scholarship (including the substantial academic commentary on *Havner* itself). *Garza* eliminated any doubt about how

⁷ *Havner*, 953 S.W.2d at 715.

⁸ *Id.* at 717 (emphasis added).

⁹ *Id.* at 715-16.

¹⁰ *Id.* at 719, 727.

¹¹ *Id.* at 714-716.

¹² Michael D. Green et al., *Reference Guide on Epidemiology*, in *Reference Manual on Scientific Evidence* 549, 555 (3d ed. 2011). Merck argued that the *Havner* rules should apply to clinical trials as well as to observational studies, but “never denied that clinical trials are superior to observational studies.” *Merck & Co. v. Garza*, Merck’s Reply Brief on the Merits, No. 09-0073, 2009 WL 2362112 (Tex. July 10, 2009).

rigid that court's requirements for epidemiologic studies really are.

A few courts had tried to find wiggle room, for example, in *Havner's* statement that a relative risk greater than two was not a "litmus test" for an epidemiologic study.¹³ In *Garza* the Texas Supreme Court emphasized that it had abjured litmus tests only to one side of neutral: a relative risk greater than two might still be insufficient, but a relative risk less than two can never be sufficient proof of causation.¹⁴ Driving the point home, the court rejected one of the studies the Garzas relied on because the statistically significant overall relative risk was only 1.92 for the narrower of two categories of cardiovascular effects described.¹⁵

Moreover, said the *Garza* court, one study meeting the threshold will not do—it takes two!¹⁶ This requirement derived from *Havner's* relatively nuanced consideration of the scientific concept of reproducibility in the context of the overwhelming amount of epidemiologic study that had failed to find an association between Bendectin and birth defects. Against that mass of evidence, *Havner* emphasized, an "isolated study" finding an association "would not be legally sufficient evidence of causation."¹⁷ *Garza* liberated this statement from its

¹³ In addition to the intermediate appellate court's opinion in *Garza* itself, *Merck & Co. v. Garza*, 277 S.W.3d 430, 435 (Tex. App. 2008), see *Cotroneo v. Shaw Envtl. & Infrastructure Inc.*, No. H-05-1250, 2007 U.S. Dist. LEXIS 79139, at *8 n.5 (S.D. Tex. 2007). Most academic commentators, myself included, harbored no such illusions. Steve C. Gold, *The More We Know, the Less Intelligent We Are? How Genomic Information Should, and Should Not, Change Toxic Tort Causation Doctrine*, 34 *Harv. Envtl. L. Rev.* 369, 377 n.35 (2010); Lucinda M. Finley, *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, 356 (1999) (stating that *Havner* and similar decisions require plaintiff to "produce the magic bullet of the requisite strength of epidemiological studies").

¹⁴ *Garza*, 54 Tex. Sup. J. at *12; *id.* at *13 ("a statistically significant doubling of the risk" is "a threshold requirement of reliability" of an epidemiologic study offered to prove toxic causation).

¹⁵ This was just one of several grounds on which the court rejected the study, but it illustrates how absurd it is for courts to apply quantitative thresholds to the results of studies that are subject to sampling error. The court, which readily described 95 percent confidence intervals when it questioned the statistical significance of the findings of other studies, did not mention the 95 percent confidence interval around the 1.92 relative risk result: 1.19 to 3.11. Robert S. Bresalier et al., *Cardiovascular Events Associated with Rofecoxib in a Colorectal Adenoma Chemoprevention Trial*, 352 *N. Eng. J. Med.* 1092, 1092 (2005). This confidence interval implies a reasonable likelihood that another identically designed study might have found a relative risk substantially higher than the court's 2.0 magic number—or even further below that number.

¹⁶ *Garza*, 54 Tex. Sup. J. at *17 ("But even if [the] VICTOR [study] qualifies under *Havner's* test, it cannot do so alone. Another study is still necessary, but lacking here.").

¹⁷ *Havner*, 953 S.W.2d at 727. *Havner* also stated that "if scientific methodology is followed, a single study would not be viewed as indicating that it is 'more probable than not' that an association exists," but the discussion and supporting authority evinced particular caution about allowing a case to reach a jury based on an anomalous result contradicted by a large body of scientific evidence. *Id.* (citing *Richardson v. Richardson-Merrell Inc.*, 649 F. Supp. 799, 702 n.10 (D.D.C. 1986) as "noting that no single study would be sufficient to exonerate or to implicate Bendectin with certainty and that stud-

context and converted it into a bright-line rule that a plaintiff must produce two studies that meet the relative risk threshold and satisfy the court's other standards—regardless of whether negative findings exist in any other studies.

Rejecting Inference: A Question of Fit

Garza also reflected the Texas Supreme Court's continued commitment to an exceptionally exacting standard for congruence between investigations of scientific hypotheses and disputes of fact in legal controversies. Federal courts treat this as a standard of "fit" between scientific results and the legal conclusions they are offered to support;¹⁸ the Texas Supreme Court has framed it as a component of a scientific study's "reliability" to be judged against the requirement that the plaintiff be "similar" to the exposed group in an epidemiologic study. However the issue is described, deciding it involves a judgment about the scope of permissible inference. The more stringent the demand for "fit" or "similarity," the less inference, extrapolation, or interpolation a court is willing to permit.

In *Garza* the Texas Supreme Court reiterated that it will not consider an epidemiologic study to be probative unless the plaintiff can show that "the conditions of the study [were] substantially similar to the claimant's circumstances."¹⁹ Thus the Garzas were prohibited from relying on a clinical trial, in which the subjects took a dose of Vioxx twice as high as that prescribed to Mr. Garza for a median duration nine or 10 times longer, to provide any support for an inference that the Vioxx Mr. Garza took could have caused his heart attack.²⁰

Nevertheless, *Garza* may have indicated a little more flexibility on the similarity requirement than *Havner* did. Where *Havner* demanded "proof . . . that the [plaintiff's] exposure or dose levels were comparable to or greater than those in the studies,"²¹ *Garza* at least acknowledged that the studied exposure "need not match the claimant's usage exactly."²² How much, if any, leeway will result from this acknowledgment will depend on future interpretations. A future court might hold that all *Garza* meant by "need not match . . . exactly" was what *Havner* said: a plaintiff could rely on a study only if the plaintiff's exposure matched or exceeded the exposure in the study. But *Garza* could also be read to allow a future plaintiff to rely on an epidemiologic study even if his or her exposure were somewhat lower than, but still "substantially similar to," the exposure of the study subjects. Then the key issue would be how much lower an exposure could be, while still being "substantially similar." Resolving this issue would seem to depend quite strongly on particular factual circumstances (the nature of the substance, its toxic effects, and the toxic dose), and to require expert

ies become 'conclusive' only in the aggregate." (emphasis added)).

¹⁸ *Daubert v. Merrell-Dow Pharms. Inc.*, 509 U.S. 579, 591 (1993); see *General Electric Co. v. Joiner*, 522 U.S. 136, 144-47 (1997) (affirming district court's conclusion that animal and epidemiologic studies could not support inference of causation).

¹⁹ *Garza*, 54 Tex. Sup. J. at *15.

²⁰ *Id.* The court also noted problems of dissimilarity with several other studies on which the Garzas attempted to rely. See *id.* at 16-17.

²¹ *Havner*, 953 S.W.2d at 720.

²² *Garza*, 54 Tex. Sup. J. at *15.

opinion about what levels of exposure are biologically comparable enough to allow inference from one to the other.

Yet there is reason to doubt that the Texas Supreme Court would tolerate any significant role for expert-guided inference. *Havner* and *Garza* both suggest that the Texas Supreme Court is at least suspicious of, and perhaps actively hostile to, a variety of inferential approaches to proof of causation. Neither decision, for example, so much as acknowledges the possibility that a dose-response relationship could connect the effects of a studied higher dose to a plaintiff's received lower dose. *Garza* further restricted the possibility of inferential reasoning when it made clear that the various requirements for an epidemiologic study are to be evaluated sequentially, taking for granted that failure to meet any one of these requirements would be fatal to a study's ability to pass the "primary reliability inquiry."²³ That approach leaves no room for a scientist to explain why a study that did not satisfy all of the requirements might nonetheless support an inference of causation, or how multiple studies might support a conclusion even though each study individually might be inconclusive. *Garza* rejected the very concept that such reasoning might be valid: "A plaintiff cannot prove causation by presenting different types of unreliable evidence."²⁴

Power to the Judges: Resurrecting *Frye's* Spirit

Even if a plaintiff presents two studies that survive the "primary" *Havner-Garza* reliability gantlet, the proof must elude one final club, or, more precisely, gavel. The reward for passing all of the tests already described—the doubling of relative risk threshold, the statistical significance standard, the similarity requirement, the tests of methodology and bias, and the two-studies requirement—is only a possibility that the evidence will reach a fact-finder. The judge is then empowered to make a "secondary reliability inquiry" and render his or her own determination of the "soundness of a study's findings using the totality of the evidence test."²⁵

In the post-*Daubert* era we take for granted that judges assess the reliability of scientific evidence, and it is not revolutionary to say that judges may dismiss cases that lack sufficient evidence to support a jury verdict. But by framing the totality of the evidence test as a matter of reliability, *Garza* explicitly conflated these roles and invited trial judges to weigh conflicting evidence.

Garza's litigation history demonstrated how completely the two concepts have merged. In the trial court, Merck made its *Havner* argument at least three times: in a motion made before trial to exclude the testimony of plaintiffs' causation expert, in a motion made during trial to strike that testimony after it was admitted, and in a motion made at the conclusion of plaintiffs' case for a directed verdict because of insufficient proof. The argument's predicate never changed. On appeal, Merck pursued only the assertion of insufficiency. The word "admissibility" appeared only once in the Texas Supreme Court's opinion (in a footnote), yet its entire le-

gal discussion was about the "scientific reliability" of the basis for plaintiffs' proffered expert testimony.²⁶

The line between evidentiary admissibility and substantive sufficiency in toxic tort cases has long been blurred.²⁷ After *Garza* it seems to have disappeared entirely in Texas, further shifting decisionmaking authority from juries to judges.

Furthermore, *Garza* made clear that, at least in toxic tort cases, the Texas Supreme Court remains determined to create clear rules for deciding what constitutes "good science." For the purpose of assessing epidemiologic evidence, the general standard that evidence must be "reliable" has been reduced to a checklist of judicially-chosen ideals. This will limit the role not only of juries but also of trial judges, whose reliability decisions will involve nothing more than holding plaintiff's evidence against the Texas Supreme Court's template to see if it matches. Reasoning about the reliability of the particular evidence presented in light of the particular facts of the case will be superfluous.

Garza thus completed a remarkable about-face on the nature of judicial evaluation of expert testimony. In *Daubert*, the U.S. Supreme Court directed federal judges to assess the "relevance and reliability" of the methodology underlying a proffered scientific expert's opinion,²⁸ rather than apply the *Frye* test that demanded "general acceptance" of any scientific principle "from which the [expert's] deduction is made."²⁹ Furthermore, the Court cautioned, judges must focus "solely on principles and methodology, not on the conclusions that they generate."³⁰ Texas adopted the *Daubert* mode of analysis, including the methodology-versus-conclusions caution, in *E.I. du Pont de Nemours & Co. v. Robinson*,³¹ a 5-4 decision³² that affirmed the exclusion of plaintiffs' expert who attempted to testify that the fungicide Benlate had damaged the plaintiffs' pecan orchard.

Garza, however, completely abandoned both the distinction between methodology and conclusions,³³ and

²⁶ See *Garza* at *11 n.21 (admissibility); *id.* at *17 (reliability).

²⁷ See Joseph Sanders, *The Controversial Comment C: Factual Causation in Toxic-Substance and Disease Cases*, 44 *Wake Forest L. Rev.* 1029, 1032-35 (2009); see also *Havner*, 953 S.W.2d at 712 ("While Rule 702 deals with the admissibility of evidence, it offers substantive guidelines in determining if the expert testimony is some evidence of probative value.")

²⁸ *Daubert*, 509 U.S. at 598.

²⁹ *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923).

³⁰ *Daubert*, 509 U.S. at 595.

³¹ 923 S.W.2d 549 (Tex. 1995); *id.* at 552 (adopting *Daubert*); *id.* at 557 (adopting distinction between methodology and conclusions).

³² The dissent argued that the Texas Rules of Evidence, which tracked the Federal Rules of Evidence, had adequately permitted exclusion of expert testimony that was based on "facts or data" not "of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject," without need to resort to the *Daubert* reliability analysis. *Id.* at 562 (dissenting opinion of Cornyn, J.) (quoting Tex. R. Civ. Evid. 703).

³³ This distinction had already been weakened by decisions—including *Havner*, see 953 S.W. 2d at 714, and then *Joiner*, see 522 U.S. at 145—that allowed judges to reject an expert's reliance on scientifically reliable studies to reach a conclusion the court deemed unjustified by those studies. See, e.g., Jean Macchiaroli Eggen, *Toxic Torts and Causation: The Chal-*

²³ *Id.* at *14.

²⁴ *Id.* at *17.

²⁵ *Id.* at *14.

the difference between a reliability test and a general acceptance test. The plaintiffs' experts in *Garza* based their opinions on clinical studies undertaken or commissioned by the drug's maker. Merck did not argue that the studies were themselves scientifically unreliable; it challenged only the propriety of inference from these studies to Mr. Garza's case. The court, however, said nothing at all about whether a reasonable *scientist* would consider such an inference reliable, or even about whether the reliability of such an inference is subject to honest scientific debate. Rather, the Texas Supreme Court, following its own precedent and the defendant's invitation, announced its own view of what scientific evidence *could* support the conclusion that Vioxx taken as Mr. Garza took it could cause a heart attack.³⁴

With its emphasis on statistical significance, magnitude of effect, tightness of "fit," and above all reproducibility, the *Havner-Garza* standard effectively means that a causation claim in Texas must be supported by evidence that (the Texas Supreme Court thinks) should lead a scientist to accept the causal conclusion; otherwise, the claim is supported by no evidence at all. The requirement of methodological reliability has mutated into something very like a requirement of general acceptance of a scientific conclusion.

A perspicacious federal judge, just weeks after the U.S. Supreme Court decided *Daubert*, remarked that the "decision . . . kills *Frye* and then resurrects its ghost."³⁵ Texas never had a live *Frye* precedent to kill,³⁶ but the Texas Supreme Court has brought the *Frye* standard to life by spontaneous generation.

The Future: Preferring False Negatives

Some other courts have recently taken an approach to toxic tort causation quite different from the Texas approach. In *Milward v. Acuity Specialty Products Group Inc.*,³⁷ the court of appeals for the First Circuit explicitly endorsed use of a "weight of the evidence" methodology to support an inference of causation even if no individual study would suffice.³⁸ In *Anderson v. Akzo Nobel Coatings Inc.*,³⁹ the Washington Supreme Court held that, even though Washington continues to apply the *Frye* standard, *Frye* does not require that "there

lence of Daubert After the First Decade, 17 *Nat. Resources & Env't* 213, 214 (2003).

³⁴ Merck argued, and *Garza* held, "that *Havner* requires a plaintiff who claims injury from taking a drug to produce two independent epidemiological studies showing a statistically significant doubling of the relative risk of the injury for patients taking the drug under conditions substantially similar to the plaintiff's (dose and duration, for example) as compared to patients taking a placebo." *Garza*, 54 Tex. Sup. J. at *10.

³⁵ *In re Joint E. & S. Dist. Asbestos Litig.*, 827 F. Supp. 1014, 1033 (S.D.N.Y. 1993). Researchers have documented that *Daubert*, although it nominally relaxed the general acceptance test, actually resulted in a tightened standard for expert testimony. Lloyd Dixon & Brian Gill, *Changes in the Standards for Admitting Expert Evidence in Federal Civil Cases Since the Daubert Decision*, 8 *Psychol., Pub. Pol. & L.* 251, 271-74, 291-92 (2002).

³⁶ The dissent in *Robinson* pointed out that Texas had never adopted the general acceptance test as a standard for admissibility of expert testimony. *Robinson*, 923 S.W.2d at 568 (dissenting opinion of Cornyn, J.).

³⁷ 639 F.3d 11 (2011).

³⁸ *Id.* at 38.

³⁹ 260 P.3d 857 (2011).

must be scientific consensus that a specific type of exposure causes a specific type of injury before expert testimony is admissible."⁴⁰

By contrast, the Texas Supreme Court has twice required that causation either "be proved directly" by "controlled scientific experiments" or be proved "indirectly" by data from epidemiologic studies (satisfying the court's vision of adequate scientific proof.⁴¹ The "indirect" nature of epidemiologic proof is taken to justify stringent standards that will often prevent a factfinder from even considering epidemiologic studies. Yet the alternative option seems chimerical. Considering that *Havner* discounted reliance on experiments conducted on animals or in test tubes, and *Garza* treated clinical drug trials—the closest thing to human toxicological experimentation that society will tolerate—as indistinguishable from observational epidemiologic studies, it is doubtful that any type of evidence available in the real world would qualify as "directly" establishing causation "through controlled scientific experiments."

It is unimaginable that the Texas Supreme Court does not know that few if any cases will be able to meet the standards it has set. The inescapable conclusion is that the Texas Supreme Court has made a policy judgment that in toxic tort claims, erroneous failures to attribute causation are preferable to erroneous findings of causation.⁴²

Vermont Supreme Court: *Blanchard v. Goodyear Tire & Rubber Co.*

The Vermont Supreme Court first considered the use of epidemiology in toxic tort causation much more recently, but nevertheless it has already had occasion to revisit its prior holding. *Blanchard v. Goodyear Tire & Rubber Co.*⁴³ gave the court an opportunity to think about its 2010 decision, *Estate of George v. Vermont League of Cities and Towns*.⁴⁴

Mr. Blanchard had a type of non-Hodgkin's lymphoma (NHL). He alleged that his cancer resulted from exposure to benzene that Goodyear had discharged from a manufacturing plant onto a baseball field where Mr. Blanchard had spent many hours as a teenager.⁴⁵ The plaintiff's expert witness testified that "a substantial body of new epidemiology research" showed a relative risk greater than two "to the link between benzene and non-Hodgkin's lymphoma."⁴⁶ The Vermont Supreme Court unanimously affirmed the trial court's entry of summary judgment against Mr. Blanchard—even though he argued that he, unlike the claimant in *Estate of George*, had epidemiology on his side.

Vermont Precedent: *Estate of George*

In *Estate of George*, a firefighter's estate alleged that he had contracted NHL because he had been exposed to

⁴⁰ *Id.* at ¶ 16.

⁴¹ *Garza*, 54 Tex. Sup. J. at *9 (quoting *Havner*, 953 S.W.2d at 714-15).

⁴² See Steve C. Gold, *The "Reshaping" of the False Negative Asymmetry in Toxic Tort Causation*, 37 *Wm. Mitchell L. Rev.* 1507, 1520-32 (2011).

⁴³ 30 A.3d 1271 (Vt. 2011).

⁴⁴ 993 A.2d 367 (Vt. 2010).

⁴⁵ *Blanchard*, 30 A.3d at 1273-74.

⁴⁶ *Id.* at 1278. The trial court admitted plaintiff's expert testimony, unlike the trial court in *Estate of George*.

carcinogens while fighting fires. The worker's compensation agency denied the claim for failure to prove that Mr. George's employment caused his illness. The estate appealed to the trial court, which excluded the estate's proffered expert causation testimony. By a 2-1-2 vote, the Vermont Supreme Court affirmed.⁴⁷

The estate's experts relied on eight published studies that reported that firefighters had elevated risk for NHL as compared to the general population. Only some of the results were statistically significant, but two of the studies reported statistically significant relative risks greater than two. One of the claimant's experts performed a meta-analysis which "found the summary risk estimate for NHL to be 1.51."⁴⁸

The problem with this evidence, the majority held, was that Vermont "law requires claimant to show, not merely that firefighting increased the likelihood of injury, but that it more likely than not caused *his* disease,"⁴⁹ and the claimant's evidence was insufficient to establish specific causation. Quoting *Havner*, the court stated that "the very use of epidemiological evidence to show specific causation reflects a compromise, given that epidemiological studies 'cannot indicate the actual cause of a given individual's disease or condition.'"⁵⁰ The court, therefore, approved the trial judge's use of relative risk greater than two as a "benchmark" that "easily tied into Vermont's 'more likely than not' civil standard."⁵¹ Because the epidemiologic studies "reflected widely varying degrees of relative risk," and only two of the eight had statistically significant results showing more than a doubling of risk, the court affirmed the exclusion of claimant's causation testimony and the summary judgment that ensued.⁵²

The two dissenting justices argued that the claimant's experts should have been permitted to present to a fact-finder their views that the epidemiologic data, coupled with information about the conditions faced by this particular firefighter, could support an inference of specific causation. The dissenters questioned the majority's conclusion that the wide range of reported relative risks favored exclusion of the proffered testimony.⁵³ To demand that each study show more than a doubling of risk, the dissenters opined, "requires each study to prove that claimant should win on the merits" rather than allowing a claimant to combine pieces of supporting evidence to meet the preponderance standard of proof.⁵⁴ The claimant's experts would have testified that several factors supported an inference that Mr. George faced higher occupational NHL risk than the average firefighter studied by the epidemiologists, the dissenters noted, but the majority precluded the fact-finder from considering those factors in combination with the epidemiologic data.⁵⁵

⁴⁷ *Estate of George*, 993 A.2d at 369-70.

⁴⁸ *Id.* at 381; *id.* at 387 (Reiber, C.J., dissenting).

⁴⁹ *Id.* at 381.

⁵⁰ *Id.* at 377 (quoting *Merrell-Dow Pharms. Inc. v. Havner*, 953 S.W.2d 706, 718 (Tex. 1997)).

⁵¹ *Id.* at 375.

⁵² *Id.* at 375, 382.

⁵³ *Id.* at 397 (Reiber, C.J., dissenting). The dissent wondered how many studies, or what percentage of studies, must show a relative risk greater than two to form a basis for an expert opinion.

⁵⁴ *Id.* at 387 (Reiber, C.J., dissenting).

⁵⁵ *Id.* at 384 (Reiber, C.J., dissenting).

Reconciling *Blanchard* and *Estate of George*: The Role of Proof of Exposure

The Vermont Supreme Court's unanimous opinion affirming entry of summary judgment against Mr. Blanchard—who introduced expert opinion based on epidemiologic studies that showed a relative risk greater than two—might seem at first blush to be in some tension with the *Estate of George* majority's endorsement of doubling of relative risk as a causation "benchmark," and even more at odds with the views of the *Estate of George* dissenters. The apparent tension is easily explained, however. *Blanchard* turned, primarily, not on the link between benzene exposure and NHL in general, nor even on a link between plaintiff's particular exposure to benzene and plaintiff's specific case of disease. Rather, the key failure in Mr. Blanchard's case was his inability to prove that the defendant had exposed him to any benzene at all.

The description of Mr. Blanchard's claim highlighted the problem immediately. This was not a case of a plaintiff's exposure to a carcinogen that occurred routinely at work, or through contact with a known source such as a contaminated drinking water well, or in a readily proven short term exposure such as an industrial accident. In attempting to link his illness to benzene vaporizing from unknown wastes encountered on portions of a ball field over a five-year period in the distant past, Mr. Blanchard faced an uphill battle. Over and over again, the Vermont Supreme Court emphasized that Mr. Blanchard's evidence failed to prove that there was *any* benzene on the ball field when he played there, much less that he was exposed to benzene or to any knowable minimum dose of benzene. The court relied on this deficiency to discount several pieces of evidence Mr. Blanchard introduced in his attempt to prove causation.⁵⁶ Because the decision is dominated by the absence of good evidence of exposure, it is easy to see how the dissenters in *Estate of George* could join the *Blanchard* opinion without difficulty.

It is worth comparing the role of proof of exposure in the Vermont cases to its role in the Texas case, *Garza*. Mr. Garza's exposure was as well documented as exposure could be—the parties knew and did not dispute the exact amount of Vioxx Mr. Garza had taken. The problem in that case, according to the Texas Supreme Court, was that the degree of Mr. Garza's exposure was much less than the degree of exposure in the epidemiologic studies on which the plaintiff's experts relied. Mr. Blanchard, by contrast, could not prove that he was exposed to benzene at all. *Estate of George* lies between these extremes of proof: the inference that Mr. George had been exposed to carcinogens on the job was reasonably strong, but the extent of his exposure was unknown and unknowable because no one had sampled and recorded the constituents of what he breathed while on the job. The dissenters would have admitted evidence that Mr. George's personal relative risk was actually *higher* than the epidemiologic studies suggested. But the estate's inability to match Mr. George's exposure to the epidemiologic study subjects' exposure was *not* the basis on

⁵⁶ *Blanchard*, 30 A.3d at 1274-78 (¶¶ 5, 6, 8, 11, 15); *see also id.* at 1276 (¶ 9, stating that environmental studies of the field did not "preclude" the possibility of benzene contamination—but not that those studies provided affirmative evidence of benzene's presence).

which the majority affirmed the exclusion of the estate's expert testimony. Vermont has not gone as far down the road of bright-line rules as has Texas.

But *Blanchard* repeatedly confused proof of exposure with proof of specific causation,⁵⁷ thus also muddying the analytical distinction between general and specific causation. Exposure to the allegedly toxic agent, of course, is logically essential to any toxic tort claim—if plaintiff never encountered the accused substance, that substance cannot possibly have caused plaintiff's illness, no matter how potent the substance's disease-causing power. Absent exposure, proof of general causation is irrelevant.

A quantitative or qualitative estimate of the amount of exposure can be important to proof of specific causation if, for example, there is reason to believe that a plaintiff's exposure might have been below a threshold required to produce a toxic response.⁵⁸ In *Blanchard*, the Vermont Supreme Court observed that "hard proof" of the amount of exposure would often be unavailable, but nevertheless seemed to suggest that a court should demand "measurements" of exposure as a condition for admitting expert testimony on specific causation.⁵⁹ The court thus seemed to demand the impossible even while acknowledging its impossibility.

The confusion of exposure and causation grew worse when the court suggested that the purpose of Mr. Blanchard's proffered differential diagnosis testimony was to overcome his lack of good proof of exposure.⁶⁰ The court's definition of differential diagnosis—"a scientific analysis entailing . . . listing all likely causes of the patient's observed symptoms or injury, then eliminating all but one cause"⁶¹—belies such use, for exposure to the toxic substance at issue would also have to be ruled out as a cause upon a conclusion that plaintiff had never been exposed to it.

Specific Causation: The Significance of Idiopathic Disease

Despite the lack of the necessary predicate of exposure, the *Blanchard* court nevertheless discussed plaintiff's differential diagnosis testimony. Its discussion exaggerated the importance of the fact that medicine has

⁵⁷ *Id.* at 1275 ("plaintiffs in toxic exposure cases must demonstrate specific causation by submitting evidence concerning 'the amount, duration, intensity, and frequency of exposure'" (citing *White v. Dow Chem. Co.*, 321 F. Appx. 266, 273 (4th Cir. 2009)); see *id.* (treating exposure evidence as part of the reliability assessment for expert testimony on specific causation).

⁵⁸ Estimates of the amount of exposure may also affect the strength of an inference of causation if the toxic effect is dose-dependent (even without a threshold) and the plaintiff's exposure can be compared to the exposure levels encountered by the subjects in an epidemiologic study.

⁵⁹ *Blanchard*, 30 A.3d at 1275 ("Of course, in many, if not most, toxic tort cases it is impossible 'to quantify with hard proof—such as the presence of the alleged toxic substance in the plaintiff's blood or tissue—the precise amount of the toxic substance to which an individual plaintiff was exposed.'" (quoting *Plourde v. Gladstone*, 190 F. Supp. 2d 708, 721 (D. Vt. 2002)); "courts generally preclude experts from testifying as to specific causation without having any measurements of a plaintiff's exposure to the allegedly harmful substance." (quoting *Henricksen v. ConocoPhillips Co.*, 605 F. Supp. 2d 1142, 1157 (E.D. Wash. 2009))).

⁶⁰ *Id.*

⁶¹ *Id.* (quoting *Plourde*, 190 F. Supp. 2d at 722).

not yet explained much of the incidence of Mr. Blanchard's disease (and therefore labels the unexplained cases "idiopathic"). The court seemed to suggest that differential diagnosis can never help prove causation if most cases of a disease are unexplained.⁶² Such a suggestion is an overstatement. If an expert can rule out the (limited) other known causes of a disease, that testimony makes it more likely that the alleged toxic exposure was the cause, and therefore should be relevant to the causation issue. It may not be sufficient, on its own, to persuade a jury, but it has probative value. If a plaintiff suffering from such a disease could provide strong proof of general causation (e.g. epidemiologic or toxicology studies) and strong proof of exposure, a differential diagnosis that ruled out other known causes would certainly be helpful. To see this, one need only imagine how the defense's closing statement would address the absence of a differential diagnosis in such circumstances.

Blanchard stated that if most cases of a disease are idiopathic, something "beyond a differential diagnosis is required" to prove specific causation.⁶³ The court did not say what that something could possibly be. Whatever it is, it is very likely beyond the ability of science today to provide.⁶⁴

Specific Causation: The Role of Epidemiology and Inference

Blanchard's discussion of epidemiology was somewhat inconsistent. On the one hand, the court restated its holding, from *Estate of George*, that a relative risk greater than 2.0 is both necessary and sufficient to prove "more likely than not" specific causation in an individual case.⁶⁵ On the other hand, the court stressed that epidemiologic studies "focus on general causation . . . rather than specific causation,"⁶⁶ and seemed to suggest that because epidemiologic data are population-based, epidemiology can never be used to prove specific causation.⁶⁷

The court was of course correct that discerning the cause of an individual's disease is not the object of epidemiologic study, but it was wrong to imply that epidemiologic data cannot bear on the legal issue of specific causation. Epidemiology assesses disease incidence in populations rather than in individuals, but both courts and ordinary people, perhaps without realizing it, readily use epidemiologic information to support causal

⁶² *Id.*

⁶³ *Id.* (quoting *Henricksen*, 605 F. Supp. 2d at 1162).

⁶⁴ The district court in *Henricksen* seemed to suggest that the required "something beyond" should be molecular or chromosomal damage unique to cases of the plaintiff's disease that are caused by exposure to the substance at issue. *Henricksen*, 605 F. Supp. 2d at 1162-63. For a critique of such a requirement in general and of *Henricksen* in particular, see Gold, *supra* note 13, at 401-06.

⁶⁵ *Blanchard*, 30 A.3d at 1277.

⁶⁶ *Id.* at 1275.

⁶⁷ *Id.* ("In addition to showing general causation through epidemiological studies, plaintiffs in toxic exposure cases must demonstrate specific causation by submitting evidence concerning the amount, duration, intensity, and frequency of exposure." (internal quotation marks and citation omitted)); see *id.* at 1278 (evidence of association between benzene and NHL not sufficient to prove specific causation).

inferences about individual cases.⁶⁸ At bottom, for example, it is epidemiology that supports the conclusion that a heavy smoker's lung cancer was caused by exposure to tobacco smoke, or that a non-smoking pipefitter's mesothelioma was caused by the asbestos he cut for many years.⁶⁹

Blanchard, however, did not fully commit the Vermont Supreme Court to the position that epidemiology is irrelevant to specific causation. To take that position would have been practically inconsistent with *Estate of George*. Rather, in discussing the limited role for epidemiology in proving Mr. *Blanchard's* particular claim of specific causation, the court again emphasized the lack of proof of exposure and thus the lack of a link between the epidemiologic data and Mr. *Blanchard's* case.⁷⁰

The *Blanchard* court hinted at another issue involving the plaintiff's epidemiologic evidence. Mr. *Blanchard* was diagnosed with a particularly rare subtype of NHL. Part of his difficulty in proving general causation seemed to result from the lack of study of that subtype.⁷¹ It seems clear that Mr. *Blanchard's* expert attempted to infer a causal role for benzene in this subtype based on studies of other leukemia subtypes. The court noted the problem but did not really express a view on this type of inference.⁷² As a general matter, courts should be careful not to categorically reject this type of inference, but should consider whether or not the subtype classifications (which reflect histological differences or other classification schemes) exclude the possibility of etiologic similarity.⁷³

The Future: Unresolved Issues

In the Vermont Supreme Court's view, Mr. *Blanchard* suffered from a disease whose causes are not well understood; he had little to no evidence that he was exposed to the chemical that he alleged made him sick; and he did not have particularly strong general causation evidence linking that chemical to his disease in any case. Thus described, it is little wonder the court affirmed the trial judge's refusal to let Mr. *Blanchard's* claims reach a jury.

⁶⁸ This is especially true for "signature diseases" for which exposure confers an exceptionally high risk relative to non-exposed individuals.

⁶⁹ Today, biochemical, molecular, and genomic techniques have partly explained these carcinogenic mechanisms. See, e.g., *Tompa v. Philip Morris USA Inc.*, 362 F.3d 882, 89-94 (discussing cancer-related mutations and cellular damage caused by tobacco smoke); *In re Asbestos Prod. Liab. Litig.* (No. VI), 714 F. Supp. 2d 535, 540-41 (discussing molecular basis of asbestos-caused disease). The initial strong evidence about these carcinogens, however—which no doubt stimulated much of the later mechanistic research—came from epidemiologic studies confirming clinical observations and suspicions. See, e.g., *Allan M. Brandt, The Cigarette Century*, 153-56, 211-16 (2007) (discussing early epidemiology linking cigarette smoking to lung cancer); Irving Selikoff et al., *Relationship Between Exposure to Asbestos and Mesothelioma*, 272 *New Eng. J. Med.* 560 (1965) (documenting extremely high relative risk).

⁷⁰ *Blanchard*, 30 A.3d at 1275.

⁷¹ *Id.* at 1278.

⁷² See *id.* (stating only that expert's testimony of "an association between occupational exposure to benzene and non-Hodgen's [sic] lymphomas" was not "sufficient to support a jury finding of specific causation.")

⁷³ For an example of the latter approach, see *Milward*, 639 F.3d at 19.

Yet *Blanchard* is best understood in the context of its particular facts. The court did not use this flawed case to articulate a broader vision of toxic tort causation and explain clearly what proof would suffice in Vermont to establish specific causation. The court gave some hints in its discussion of differential diagnosis and of epidemiology. But it seemed to be feeling its way and left much unresolved.

For example, the Vermont Supreme Court's statements on differential diagnosis for diseases with mostly idiopathic origin could—if taken literally, read out of context, and applied strictly—be construed to doom automatically any toxic tort claim of a plaintiff with such a disease. But in *Blanchard*, the existence of a large fraction of idiopathic cases was particularly salient because of the plaintiff's minimal proof of exposure and the concomitantly doubtful pertinence of the epidemiologic data. Yet a toxin could impart a very high relative risk for a disease even though exposure to the toxin explains only a relatively small fraction of the total incidence. If a plaintiff were to offer sufficient proof that he or she had been exposed to such a toxin, it seems unlikely that the Vermont Supreme Court would apply *Blanchard* to bar the plaintiff's claim; to do so would be illogical and unjust.

Similarly, as described above, the Vermont Supreme Court's description of epidemiology could be construed to suggest that a plaintiff must always produce more than epidemiology to reach a jury on the question of specific causation. Yet *Blanchard* reaffirmed, although it did not extend or even apply, the *Estate of George* holding that doubling of relative risk is a "benchmark" for epidemiologic proof of causation. The "benchmark" holding itself was equivocal, implying that even the relative risk threshold was not absolute (although the court applied it, in *Estate of George*, as though it were). Again in *Blanchard*, the court stopped short of establishing a bright-line rule. *Blanchard* at least left open the possibility that very strong evidence of general causation (such as strong epidemiologic results), coupled with very strong evidence of exposure, could be enough to allow a jury to infer specific causation.⁷⁴

Conclusion

These recent cases show that courts still struggle with the challenges presented by the limits of scientific proof that can be adduced in support of or in opposition to a claim of toxic causation of disease. They also reflect sharply differing responses to those challenges with different implications for litigants and lower courts. Both the Texas and Vermont decisions express a concern for incorrect attributions of causation, but the Texas Supreme Court's approach more clearly reflects a policy preference for false negative findings over false positive findings of causation. Texas's reaffirmed, even sharpened, bright-line rules for all seasons of *Garza* offer the benefit of ease of application. Yet Vermont's murky fact-bound tea leaves of *Blanchard* have other advantages: they leave more room for expert judgment and a more nuanced consideration of available scientific evidence taken as a whole.⁷⁵

⁷⁴ *Accord, Restatement (Third) Of Torts: Liab. For Physical & Emotional Harm* § 28 cmt. c(3) (2010).

⁷⁵ See *Milward*, 639 F.3d at 23.

The challenges courts face will not get easier. Science's rapid advance will complicate rather than simplify the issues, as epidemiology morphs into molecular epidemiology, as geneticists parse the genome for inherited susceptibility to toxins and diseases, and as researchers find biomarkers of disputed validity and sig-

nificance.⁷⁶ In a rapidly changing scientific landscape, courts will do well to retain doctrinal flexibility.

⁷⁶ For a description of toxicogenomics and molecular epidemiology, and an assessment of how courts hearing toxic tort causation disputes should think about these scientific advances, see Gold, *supra* note 13.