

Turning Statistics into Evidentiary Facts: Ruse or Reality

by
David W. Elrod
Elrod, PLLC

I. Introduction

“There are three kinds of lies: lies, damn lies and statistics.”¹ Most would agree that statistics can be expressed and interpreted in many different ways, often grossly skewing the facts they purport to describe. This distortion may arise unwittingly or intentionally to advance a specific agenda. Politicians are notorious for using statistics to demonstrate the positive or negative effects of particular policies on the economy, health insurance, unemployment or other areas of interest to their constituents. Commercial manufacturers of consumer goods, from automobiles to pharmaceuticals, utilize statistics to advertise the superior performance, safety or effectiveness of their products.

Despite the seemingly inherent subjectivity of statistics, our judicial system sometimes permits litigants to utilize such information in the prosecution and defense of civil and criminal actions. In such a context, however, proponents generally define statistics less derogatively as “the science and art of describing data and drawing inferences from them.”² Statistical data and analysis are typically introduced in the form of expert testimony and may relate to scientific studies or tests. As the United States Supreme Court has cautioned, however, “statistics are not irrefutable; they come in infinite variety and, like any other kind of evidence, they may be rebutted. In short, their usefulness depends on all of the surrounding facts and circumstances.”³

¹ Statement attributed to British politician and author, Benjamin Disraeli (1804-1881), and popularized in the United States by Samuel Langhorne Clemens, a.k.a. Mark Twain.

² Michael O. Finkelstein, Bruce Levin, *Statistics for Lawyers* (2d ed. 2001).

³ *Teamsters v. United States*, 431 U.S. 324, 340, 97 S.Ct. 1843, 1856, 52 L.Ed.2d 396 (1977).

The context in which litigants have attempted to use statistics include: establishing and refuting causation in toxic tort and product liability actions; using statistical extrapolation to shorten trials; establishing and rebutting discriminatory practices in employment and civil rights actions; calculating future earnings or lost profit damages; demonstrating illegal conduct and its effects in antitrust litigation; determining paternity; and identifying criminal suspects using DNA testing. The following is a brief discussion of the first three of these examples, along with general considerations regarding the reliability concerns surrounding the use of statistics.

II. Fundamental Evidentiary Considerations

A. **Rule 702, Texas Rules of Evidence**

As with all expert testimony, that which addresses statistical data and analysis must meet the basic requirements of the Texas Rules of Evidence. The testimony of an expert is generally *opinion* testimony. Whether it rises to the level of *evidence* is determined under the Texas Rules of Evidence, including Rule 702, which requires courts to determine if the opinion testimony will assist the jury in deciding a fact issue.⁴ Rule 702 provides as follows:

If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise. Tex.R.Evid. 702.

The Texas Supreme Court further clarified this Rule 702 determination as follows:

To say that the expert's testimony is some evidence under our standard of review simply because the expert testified that the underlying technique or methodology supporting his or her opinion is generally accepted by the scientific community is putting the cart before the horse. * * * The underlying data should be independently evaluated in determining if the opinion itself is reliable.⁵

⁴ *Merrell Dow Pharmaceuticals, Inc. v. Havner*, 953 S.W.2d 706, 712 (Tex.1997).

⁵ *Id.* at 712-13.

The evidentiary requirements set forth in Rule 702 of the Federal Rules of Evidence are similar to, but somewhat more comprehensive than, those set forth in Tex.R.Evid. 702.

B. *Daubert* and *Robinson* Considerations

The determination of whether expert testimony is admissible as evidence under Tex.R.Evid. 702, however, is not the end of a court's inquiry. It must also carry out its duty as evidentiary gatekeeper by examining numerous factors, including those set forth by the Texas and U.S. Supreme Courts in *Daubert* and *Robinson*, respectively, to determine whether the testimony is sufficiently reliable.⁶ Even if evidence is admissible under these cases, however, it may still be legally insufficient to withstand summary judgment.⁷

Under *Daubert*, admissible expert testimony must be based on “scientific knowledge”, that is, knowledge grounded in and based upon the “methods and procedures of science” and “supported by appropriate validation.”⁸ The following factors set forth in *Robinson* identify what Texas courts should consider in looking beyond the bare opinion of the expert:

- (1) the extent to which the theory has been or can be tested;
- (2) the extent to which the technique relies upon the subjective interpretation of the expert;
- (3) whether the theory has been subjected to peer review and publication;
- (4) the technique's potential rate of error;
- (5) whether the underlying theory or technique has been generally accepted as valid by the relevant scientific community; and
- (6) the non-judicial uses that have been made of the theory or technique.⁹

Moreover, in *Exxon Pipeline Co. v. Zwahr*, 88 S.W.3d 623, 629, the Texas Supreme Court held that expert testimony is unreliable if: (1) it is not grounded in the methods and procedures of science

⁶ *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 113 S.Ct. 2786, 125 L.Ed.2d 469 (1993); *E.I. du Pont de Nemours & Co. v. Robinson*, 923 S.W.2d 549 (Tex.1995).

⁷ *Havner*, 953 S.W.2d at 713, citing *Conde v. Velsicol Chem. Corp.*, 24 F.3d 809, 813 (6th Cir.1994).

⁸ *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. at 589-90, 113 S.Ct. 2786.

and is thus no more than subjective belief or unsupported speculation; or (2) there is too great an analytical gap between the data upon which the expert relies and the opinion he offers.

As revealed in the following discussion of various applications of statistics, courts have imposed even further requirements or reliability thresholds upon particular types of studies and tests commonly used in statistical analysis in an effort to ensure that unreliable evidence is not considered.

III. Statistics Relating to Causation in Toxic Tort/Product Liability Actions

A. **Direct and Indirect Proof of Causation**

In *Havner*, the Texas Supreme Court examined in great detail the use of statistical analysis, based largely on epidemiological studies, that was offered to show exposure to a toxic agent resulted in a particular injury or condition. That case involved a child born with a limb reduction birth defect whose mother ingested the prescription drug Bendectin during pregnancy. The Havner parents prevailed against the drug's manufacturer in a jury trial and received actual and exemplary damages. In addressing the issue of whether there was any evidence that Bendectin caused the Havner's child to be born with a birth defect, the Supreme Court held that the evidence offered was legally insufficient to establish causation. The following discussion stems, in large part, from the court's extensive opinion in that case regarding epidemiological studies and the plaintiffs' "more likely than not" burden of proof. In short, the court states that statistics and related studies must be reliable to be relevant or admissible, and goes to great lengths to define minimum standards of reliability.

In toxic tort cases, causation is often discussed in terms of general and specific causation. General causation is whether a substance is capable of causing a particular injury or condition in the general population, while specific causation is whether a substance caused a particular individual's

⁹ *Robinson*, 923 S.W.2d at 557. The issue in *Robinson* was admissibility of evidence, but the same factors may be applied in a no-evidence review of scientific evidence. *Merrell Dow Pharmaceuticals, Inc. v. Havner*, 953 S.W.2d 706, 714 (Tex.1997).

injury. In some instances, controlled scientific experiments can be carried out to determine if a substance is capable of causing a particular injury or condition, and there will be objective criteria by which it can be determined with reasonable certainty that a particular individual's injury was caused by exposure to a given substance. However, in many toxic tort cases, direct experimentation cannot be done, and there will be no reliable evidence of specific causation.¹⁰ Both general and specific causation are required to establish causation in toxic tort cases.¹¹

In the absence of direct, scientifically reliable proof of causation, claimants may attempt to demonstrate via epidemiological or other studies that exposure to the substance at issue increases the risk of their particular injury. The finder of fact is asked to infer that because the risk is demonstrably greater in the general population due to exposure to the substance, the claimant's injury was more likely than not caused by that substance. This theory concedes, however, that science is unable to determine exactly what caused a particular plaintiff's injury.¹²

B. Epidemiological Studies – General Causation

Epidemiological studies examine existing populations to determine if there is an *association* between a disease or condition and a factor suspected of causing that disease or condition.¹³ Commentators in this area uniformly acknowledge that epidemiological studies cannot establish that a given individual contracted a disease or condition due to exposure to a particular drug or agent.¹⁴

¹⁰ *Havner*, 953 S.W.2d. at 714-15.

¹¹ *Id.*; *See also, Frias v. Atlantic Richfield Co.*, 104 S.W.3d 925 (Tex.App.-Houston [14th Dist.] 2003).

¹² *Id.* at 715.

¹³ *Id.*, *citing* Bert Black & David E. Lilienfeld, *Epidemiologic Proof in Toxic Tort Litigation*, 52 FORDHAM L.REV. 732, 750 (1984).

¹⁴ *Id.* at 715, *citing* Michael Dore, *A Commentary on the Use of Epidemiological Evidence in Demonstrating Cause-In-Fact*, 7 HARV. ENVTL. L. REV. 429, 431-35 (1983); Steve Gold, *Causation in Toxic Torts: Burdens of Proof, Standards of Persuasion, and Statistical Evidence*, 96 YALE L.J. 376, 380 (1986).

Following an exhaustive review of the principles and basic concepts underlying epidemiological studies, and the scientific community's general consensus regarding their reliability, the Supreme Court established minimum standards that such studies must meet, otherwise they cannot be considered as reliable scientific evidence of general causation. First, they must reflect that the risk of an injury or condition in the exposed population is more than double that in the unexposed or control population (i.e., "relative risk" >2). Second, they must have a "confidence level" of at least 95%.¹⁵ Third, an isolated study finding a statistically significant association between a substance and a disease or condition is not legally sufficient evidence of causation because any conclusion about causation can be reached only after an association is observed in studies among different groups and the association continues to hold when the effects of other variables are taken into account.¹⁶

1. Relative Risk >2.0

The relative risk that a person exposed to an agent will develop a particular disease is examined via a particular type of epidemiological study known as a cohort study, or incidence study. This is a prospective study that identifies groups and observes them over time to see if one group is more likely to develop disease.¹⁷ Relative risk is the ratio of the incidence of the disease among the exposed population to its incidence among the general, or control population. For the result to indicate a doubling of the risk, the relative risk must be greater than 2.0.¹⁸ The *Havner* court cautioned, however, that "[w]e do not hold [however] that a relative risk of more than 2.0 is a litmus test or that a single epidemiological test is legally sufficient evidence of causation. Other factors must be considered."¹⁹

¹⁵ *Frias*, 104 S.W.3d at 928, citing *Havner*, 953 S.W.2d at 717-24.

¹⁶ *Frias*, 104 S.W.3d at 928, citing *Havner*, 953 S.W.2d at 727.

¹⁷ *Havner* at 721, citing Bailey et al., *Reference Guide on Epidemiology*, in FEDERAL JUDICIAL CENTER, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 169 (1994) at 173, 176.

¹⁸ *Havner*, 952 S.W.2d at 721, citations omitted.

¹⁹ *Id.* at 718.

2. Confidence Level of 95%

A confidence level in epidemiological studies is used to establish the boundaries of the relative risk. These boundaries are known as the confidence interval. A confidence interval demonstrates whether the results of a study are statistically significant at a particular confidence level. A confidence interval shows a range of values within which the results of a study sample would be likely to fall if the study were repeated numerous times.²⁰ Additionally, if, at a 95% confidence level, the lower end of the confidence interval is 1.0 or less (regardless of the upper level), the results are statistically insignificant because a 1.0 relative risk suggests no greater incidence of the disease in the exposed population than in the unexposed.²¹ The Texas Supreme Court has held that studies with confidence or significance levels below 95% cannot be considered reliable scientific evidence.²² A 95% confidence level means that if a study were repeated numerous times, the confidence interval would indicate the range of relative risk values that would result 95% of the time.²³

3. Additional Indicia of Reliability

Even if a statistically significant association is found, that association does not equate to general or specific causation. For example, there is a demonstrable association between summertime and death by drowning, but summertime does not cause drowning.²⁴ There are numerous other factors to consider in evaluating the reliability of a scientific study, including, but not limited to, the sample size of the study, the power of the study, confounding variables, whether there was selection bias, and

²⁰ *Id.* at 723, citations omitted.

²¹ *Frias*, 104 S.W.3d at 928, *citing Havner*, 953 S.W.2d at 723.

²² *Havner*, 953 S.W.2d at 724, citations omitted.

²³ *Id.* at 723.

²⁴ *Id.* at 724.

whether the study has been published or subjected to other peer review.²⁵ All of these possible shortcomings must be examined carefully to determine if such studies are relevant or reliable.

C. Epidemiological Studies -- Specific Causation

In *Havner*, where the court conducted a legal sufficiency review, it set forth the types of evidence needed to establish specific causation and thereby survive such a review.²⁶ First, a claimant must do more than simply introduce into evidence epidemiological studies that show a substantially elevated risk. A claimant must show that he or she is similar to those in the studies, including proof that the injured person was exposed to the same substance, that the exposure or dose levels were comparable to or greater than those in the studies, that the exposure occurred before the onset of injury, and that the timing of the onset of injury was consistent with that experienced by those in the study. Further, if there are other plausible causes of the injury or condition that could be negated, the plaintiff must offer evidence excluding those causes with reasonable certainty.²⁷ Absent such evidence, epidemiological studies are of little, if any, relevance.

D. Reliability of Each Step of Analysis

The Texas Supreme Court further emphasized in *Havner* the level of scrutiny with which courts must examine expert opinions based on statistics by demanding reliability at each step of analysis leading to the opinion:

If the foundational data underlying opinion testimony are unreliable, an expert will not be permitted to base an opinion on that data because any opinion drawn from that data is likewise unreliable. Further, an expert's testimony is unreliable even when the underlying data are sound if the expert draws conclusions from the data based on flawed methodology. A flaw in the expert's reasoning from the data may render reliance on a

²⁵ *Id.* at 724, 726.

²⁶ *Id.* at 720, citations omitted.

²⁷ *Id.*

study unreasonable and render the inferences drawn therefrom dubious. Under that circumstance, the expert's scientific testimony is unreliable and, legally, no evidence.²⁸

Moreover, the court added: "Our legal system requires that claimants prove their cases by a preponderance of the evidence. In keeping with this sound proposition at the heart of our jurisprudence, the law should not be hasty to impose liability when scientifically reliable evidence is unavailable."²⁹

IV. Statistical Extrapolation To Shorten Trials - Class Actions, Trial Plans and Unique Damage Calculations

Statistical evidence has been promoted as a means to address manageability concerns in class action cases and cases involving complex damage calculations.³⁰ Two common shortcomings of statistical use under these circumstances, however, are that it fails to pass muster under the *Daubert* and *Robinson* standards for reliability, and that it unduly infringes upon a defendant's right to due process.³¹

Although the use of statistical evidence may reduce the complexity or length of large class action cases and make for an easier trial plan, one must temper the temptation to resort to such statistical shortcuts with the requirement to keep intact a defendant's rights to due process and cross examination. In balancing these competing interests, courts have permitted statistical extrapolation of

²⁸ *Id.* at 714.

²⁹ *Id.* at 728.

³⁰ *See, In re Ford Motor Co. Ignition Switch Prods. Liab. Litig.*, 194 F.R.D. 484 (D. N.J. 2000) (discussing plaintiffs' proposed trial plan that required the use of statistical evidence to prove causation and damages).

³¹ *See Id.* (rejecting use of trial plan because statistical evidence might be unreliable and would not make case more manageable because defendant would still be entitled to mini trials to cross examine witnesses and rebut causation); *see also, Cimino v. Raymark Indus., Inc.*, 151 F.3d 297, 319-20 (5th Cir. 1998) (reversing trial court's ruling because proposed trial plan's use of statistical extrapolations).

causation or damages in some cases, but rejected it in others.³² The decision whether or not to allow such statistical extrapolations often turns on whether their proponent can establish that they will help make the trial manageable and not improperly affect the defendant’s due-process rights.³³

A. Statistical Extrapolations Not Permitted Where The Extrapolations Infringed Defendant’s Due Process Rights Or Otherwise Unreliable

Proof of causation and damages can be challenging in class actions—especially those involving personal injury claims.³⁴ Consequently, plaintiffs frequently resort to statistical extrapolations in an effort to “simplify” the causation or damage phases of class-action trials.³⁵ In addition, statistical extrapolations are often offered as a “short-cut” to complicated damage questions or calculations. However, such extrapolations often violate the defendant’s right to cross examine each claimant or otherwise improperly infringe upon the defendant’s due-process rights.³⁶ In other cases, the statistical extrapolations are based on subjective data or not verifiable. In either situation, the use of statistical extrapolations is improper.

1. Statistical Extrapolations Interfere With A Defendant’s Due Process Rights

³² See, e.g., *Blue Cross and Blue Shield of N.J. v. Phillip Morris, Inc.*, 113 F.Supp. 2d 345, 372-76 (E.D.N.Y. 2000), *In re Chevron U.S.A., Inc.*, 109 F.3d 1016 (5th Cir. 1997).

³³ See *Blue Cross and Blue Shield of N.J.*, 113 F.Supp. 2d at 372-76 (permitting the use of statistical extrapolations because case involved a single plaintiff “advancing its own subrogated interests” and would not involve the extrapolation of liability determinations from representative plaintiffs to the remaining plaintiffs); see also *In re Chevron*, 109 F.3d at 1020 (recognizing that the results of properly selected—truly representative—bellwether cases could be extrapolated to determine damages for the entire class).

³⁴ See *Southwestern Ref. Co. v. Bernal*, 22 S.W.3d 425, 436 (Tex. 2000) (recognizing that “[p]ersonal injury claims will often present thorny causation and damage issues”).

³⁵ See *Cimino*, 151 F.3d at 319-20 (reversing because trial plan improperly extrapolated damages for all class members from the results of “bellwether” trials); see also *Southwestern Ref. Co.*, 22 S.W.3d at 437 (noting that plaintiffs’ proposed trial plan relied heavily of the use of “models, formulas, and damage brochures”).

³⁶ See *In re Fibreboard Crop.*, 893 F.2d 706, 711-12.

In certain circumstances, the use of statistical extrapolations can fundamentally alter the nature of the trial. In such circumstances, the use of statistical extrapolations should never be permitted.³⁷ For example, in *In re Fibreboard Crop.*, the plaintiffs proposed the use of statistical extrapolations to address questions of causation and damages.³⁸ In rejecting the plaintiffs' trial plan, court of appeals recognized that the plaintiffs' proposed "changes in procedure work a change in the very character of a trial."³⁹ Indeed, the court of appeals noted that the procedures proposed by plaintiffs had so altered the traditional notion of a trial the procedures "comprise[d] something other than a trial . . . [i]t is called a trial, but it is not."⁴⁰ Underlying the Court's decision was the fundamental principle that the elements of products liability cases focused on individuals and not on groups—but that the use of statistical extrapolation improperly focused the elements on groups.⁴¹

Statistical extrapolations should not be used to alter the individual nature of the elements of a particular claim. In *Cimino v. Raymark Industries*, for example, the Fifth Circuit rejected a trial plan that relied on statistical extrapolations to determine damages because, under Texas law, "causation and damages are determined respecting plaintiffs and individuals, not groups."⁴² Similarly, in *Southwestern Refining Co.*, the Texas Supreme Court reversed a lower court's class-certification decision because the proposed trial plan, which relied on statistical extrapolations, improperly prevented the defendant from challenging "the credibility of and responsibility for each personal injury claim individually."⁴³ Similarly, plaintiffs have been precluded from focusing on a "representational

³⁷ See *In re Fibreboard*, 893 F.2d at 711-12; *Cimino*, 151 F.3d 297 (5th Cir. 1998); *Southwestern Ref. Co.*, 22 S.W.3d at 437.

³⁸ See *Fibreboard*, 893 F.2d 706 at 710-12.

³⁹ See *id.* at 711.

⁴⁰ See *id.* at 712.

⁴¹ See *id.* at 710-12.

⁴² See 151 F.3d 297, 313.

⁴³ See *Southwestern Ref. Co.*, 22 S.W.3d at 437.

sample” of plaintiffs or employing “statistics to cure the individualized nature of their breach of contract cause of action or the amount of damages each plaintiff may be entitled to recover.”⁴⁴ Even in non-class action cases, a plaintiff cannot employ statistical extrapolations to avoid the elements of a particular claim.⁴⁵

Courts will also refuse to allow the use of statistical extrapolations when it would result in the creation of a composite or fictional plaintiff.⁴⁶ In *Broussard*, the court of appeals reversed the trial court’s class-certification decision because the use of statistical evidence had provided the plaintiffs with “the practical advantage of being able to litigate not on behalf of themselves but on behalf of a ‘perfect plaintiff’ pieced together for litigation.”⁴⁷ Importantly, the rejection of statistical extrapolations is almost always fatal to the proponent’s efforts to certify a class.⁴⁸

2. Statistical Extrapolations Are Subjective Or Otherwise Unreliable.

Statistical extrapolations are impermissible when they are based on an improper sampling method or are otherwise unreliable. For example, plaintiffs often propose the use of “bell-whether” trials as a means to simplify the issues of causation or damages in mass-tort or class-action cases.⁴⁹ However, in order to extrapolate the results from the “bell-whether” trials, the basis for the statistical

⁴⁴ *See Basco v. Wal-Mart Stores, Inc.*, 216 F.Supp.2d 592, 603 (E.D. La 2002).

⁴⁵ *See In re Craig’s Stores, Inc.*, 247 B.R. 652, 656 (S.D. Tex. Bankr. 1000) (noting plaintiff’s proposed statistical extrapolation did not meet the requirements of Texas law, which requires that a plaintiff bringing a breach of contract claim must “prove that the breach of contract by mishandling of accounts . . . caused specific damages.”).

⁴⁶ *See Broussard v. Meineke Discount Muffler Shops, Inc.*, 155 F.3d 331 (4th Cir. 1998).

⁴⁷ *See id.* at 344-45.

⁴⁸ *See Broussard*, 155 F.3d at 344-45 (decertifying class); *Fibreboard*, 893 F.2d at 711-12 (reversing trial court and decertifying class); *In re Ford Motor Co. Ignition Switch*, 194 F.R.D. at 485 (denying class certification).

⁴⁹ *See id.*, *see also Fibreboard*, 893 F.2d at 708-09.

extrapolation must be sound.⁵⁰ Indeed, unless the selection of the “bell-whether” trials is based upon “competent, scientific, statistical evidence that identifies the variables involved and provides a sample of sufficient size so as to permit a finding that there is a sufficient level of confidence that the results obtained reflect results that would have been obtained from trials of the whole” the use of “bell-whether” trials will not be permitted.⁵¹ Without the “competent, scientific, statistical evidence” the use of “bell-whether” trials lacks the necessary safeguards to ensure that the issues of liability and causation “are determined in a proceeding that is reasonably calculated to reflect the results that would be obtained if those claims were actually tried.”⁵² Accordingly, courts have recognized that extrapolating the results of 41 “bell-whether” trials to the remaining 2,990 class members was an improper statistical extrapolation.⁵³ Similarly, extrapolating the results of 160 “bell-whether” trials to the 2,298 remaining class members was also held to be improper. Due to the absence of the necessary “competent, scientific, statistical evidence” courts have generally rejected this approach—which ultimately results in the denial of class-certification motions.⁵⁴

If the statistical extrapolation is not based on objectively verifiable data, the extrapolation is unreliable and should be disregarded. Often, this analysis turns not on the application of statistical formulas to the available data, but the underlying data itself. For example, in *In re Ford Motor Company Ignition Switch Products Liability Litigation*, plaintiffs offered a detailed statistical analysis regarding ignition switch fires based upon Ford’s Master Owner Relations System (“MORS II”)

⁵⁰ See *In re Chevron U.S.A., Inc.*, 109 F.3d 1016, 1019-21 (5th Cir. 1997) (reversing trial court’s adoption of a trial plan that called for the results of 30 “bell-whether” trials to be applied to more than 3,000 other plaintiffs).

⁵¹ See *id.* at 1020.

⁵² See *id.* at 1020.

⁵³ See *Fibreboard*, 893 F.2d at 708-09.

⁵⁴ See *id.*; but see *See Blue Cross and Blue Shield of N.J.*, 113 F.Supp. 2d at 372-73 (discussing the appropriate use of statistical evidence).

database.⁵⁵ MORS II was created by Ford customer service personnel taking down individual customers' complaints sent by way of letter or telephone.⁵⁶ Information gathered by MORS II included the model/year of the owner's vehicle and the owner's complaint.⁵⁷ The MORS II database, however, did not consistently record whether there was a fire or if the fire was caused by the ignition switch.⁵⁸ As a result, Plaintiffs included in their statistical analysis all MORS II entries that involved a fire and mentioned the ignition switch as well as entries that reflected a fire in or near the steering column.⁵⁹ The Court rejected Plaintiffs' statistical extrapolations out of a concern that the underlying data contained too many unverified and vague allegations regarding the origin of fires to be reliable.⁶⁰ In short, the Court rejected the statistical extrapolations because it remained "doubtful that all of the MORS II reports included in the database actually represent actual ignition switch fires."⁶¹

Similarly, statistical extrapolations regarding the number of mishandled accounts receivable and the damages from the purported mishandling was rejected because the underlying data was predicated on subjective evaluations made by the expert and his screening committee.⁶² In *In re Craig's Stores*, the plaintiff alleged that the defendant had mishandled the accounts receivable it was contracted to recover, driving down the value of the accounts receivable.⁶³ Plaintiff sought to introduce statistical extrapolations that evidenced the mishandling and the amount of damages due to

⁵⁵ See *In re Ford Motor Co. Ignition Switch Prods. Liab. Litig.*, 194 F.R.D. 484, 493 (D. N.J. 2000).

⁵⁶ See *id.*

⁵⁷ See *id.*

⁵⁸ See *id.* at 493-94.

⁵⁹ See *id.*

⁶⁰ See *id.* at 494.

⁶¹ See *id.* at 495.

⁶² See *In re Craig's Stores, Inc.*, 247 B.R. 652, 655-56 (S.D. Tex. Bankr. 1000).

⁶³ See *id.* at 654-55.

the mishandling.⁶⁴ The Court prevented the use of the statistical extrapolation because the underlying data was not objective—but rather subjectively created by the plaintiff’s expert and the expert’s screening committee.⁶⁵ Moreover, the plaintiff’s expert never articulated a “standard that could be tested.”⁶⁶ Thus the statistical evidence was unreliable and properly excluded.

In addition, even if a particular statistical extrapolation is reliable and scientifically sound, the extrapolation may only create a “rebuttable presumption” of causation—not the direct proof of causation that may be required.⁶⁷ As a result, the statistical extrapolations may not necessarily be the sole evidence of causation relied upon by the plaintiffs.⁶⁸

B. Statistical Extrapolations Permitted

Courts have permitted the use of statistical extrapolations in trial plans involving the use of bell-weather trials in limited circumstances.⁶⁹ Unique factors are often present that permit the plaintiff to overcome the defendant’s objections to the use of statistical extrapolations. The specific procedural posture of the class representative or individual⁷⁰, the size of the class⁷¹, and the specific objections raised by the defendant⁷² are factors that have overcome defendant’s objections to the use of statistics.

⁶⁴ *See id.*

⁶⁵ *See id.* at 655-56.

⁶⁶ *See id.* at 656.

⁶⁷ *See In re Ford Motor Co. Ignition Switch Prods. Liab. Litig.*, 194 F.R.D. 484, 491-92 (D. N.J. 2000).

⁶⁸ *See id.*

⁶⁹ *See id.*; *see also Hilao v. Estate of Marcos*, 103 F.3d 767, 788 (9th Cir. 1996); *In re Chevron*, 109 F.3d at 1020 (rejecting trial court’s use of 30 “bellwether” cases as proper but recognizing that the results of truly representative “bellwether” cases could be extrapolated to determine the damages of remaining class members).

⁷⁰ *See Blue Cross and Blue Shield*, 113 F.Supp. 2d at 372-73 (non-class action involving individual plaintiffs pursuing subrogation claims).

⁷¹ *See In re Chevron*, 109 F.3d at 1020 (noting that properly selected bell weather results could be extrapolated to remaining class of about 3000 plaintiffs).

⁷² *See Hilao*, 103 F.3d at 785 n.12 (noting that defendant “waived any challenge to the computation of damages”).

V. Statistics in Employment Discrimination Actions

Litigants often resort to statistical evidence in an effort to prove or rebut prima facie cases under federal discrimination statutes. The purpose of such statistical proof is to draw comparisons between the representation of protected classes in an employer's work force and their representation in the relevant population or labor force.⁷³ The theory is that, in the absence of discrimination, all groups will be distributed in all employment classifications in approximately the same proportion as they are distributed in the relevant labor market as a whole.⁷⁴ The probative value of such statistics often depends on the type of statistical analysis applied and on how the relevant labor market is defined.⁷⁵ In rebutting a plaintiff's prima facie case, a defendant may offer its own statistical analyses, demonstrate that the plaintiff's proof is based on improper data or is otherwise unreliable, or show that disparities between employer's work force and the compared population result from job-related requirements or other nondiscriminatory factors.⁷⁶

A. **Age Discrimination in Employment Act of 1967**

Before discussing statistical analyses in actions brought under the Age Discrimination in Employment Act of 1967, 29 U.S.C. §623, et seq. ("ADEA") or Title VII of the Civil Rights Act of 1964, 42 U.S.C. §2000e, et seq. ("Title VII"), it is instructive to review the evidentiary burdens borne by the parties in such cases. To establish a prima facie ADEA case, a plaintiff must establish (1) membership in the ADEA-protected class (at least 40 years of age), (2) discharge from employment;

⁷³ 8 Emp. Coord. Employment Practices §107:94.

⁷⁴ *Id.*, citations omitted.

⁷⁵ *Id.*

⁷⁶ *See, e.g., Page v. U.S. Industries, Inc.*, 726 F.2d 1038 (5th Cir.1984), *Pouncy v. Prudential Ins. Co. of America*, 499 F.Supp. 427 (S.D.Tex. 1980), *affd*, 668 F.2d 795 (5th Cir.1982), *E.E.O.C. v. Datapoint Corp.* 570 F.2d 1264 (5th Cir.1978).

(3) qualification for the position held, and (4) replacement by someone outside the protected class, someone younger, or otherwise discharged because of age.⁷⁷

The burden of proof faced by plaintiffs in ADEA claims under the “pretext” analysis is described by the U.S. Supreme Court in *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 804-05, 93 S.Ct. 1817, 36 L.Ed.2d 668 (1973). First, the plaintiff must establish a prima facie case; if the plaintiff meets that burden, the defendant must produce a legitimate, non-discriminatory reason for its decision to terminate the plaintiff; if the defendant meets its burden of production, the plaintiff then has the opportunity to demonstrate that the defendant’s proffered reason for termination is merely pretextual.⁷⁸ The United States Court of Appeals, 5th Circuit has held that although statistical evidence alone may be sufficient to support a finding of age discrimination under the ADEA, such cases are rare.⁷⁹

In *Rachid v. Jack In The Box, Inc.*, 376 F.3d 305, 311-12 (5th Cir. 2004) the Court merged the *McDonnell Douglas* “pretext” approach with the “mixed-motives” approach that was described in *Price Waterhouse v. Hopkins*, 490 U.S. 228, 109 S.Ct. 1775, 104 L.Ed.2d 268 (1989) and clarified in *Desert Palace, Inc. v. Costa*, 539 U.S. 90, 123 S.Ct. 2148, 156 L.Ed.2d 84 (2003). Additionally, it held that the mixed-motives analysis, previously available only in Title VII actions, was available in ADEA claims, and that direct proof of discrimination by a plaintiff is not necessary to receive a mixed-motives analysis under the ADEA. As discussed above, statistical evidence is a frequently used type of indirect evidence. The Court described this merged analysis as follows:

Our holding today that the mixed-motives analysis used in Title VII cases post-*Desert Palace* is equally applicable in ADEA represents a merging of the *McDonnell Douglas* and *Price Waterhouse* approaches. Under this integrated approach, called, for simplicity, the modified *McDonnell Douglas* approach: the plaintiff must still

⁷⁷ *Russell v. McKinney Hosp. Venture*, 235 F.3d 219, 223-24 (5th Cir.2000), citations omitted.

⁷⁸ *McDonnell Douglas Corp. v. Green*, 411 U.S. at 804-05, cited with approval in *Rachid v. Jack In The Box, Inc.*, 376 F.3d 305, 308 (5th Cir. 2004), citation omitted.

⁷⁹ *Walther v. Lone Star Gas Co*, 952 F.2d 119, 125-26 (5th Cir.), rehearing denied 977 F.2d 161 (1992).

demonstrate a prima facie case of discrimination; the defendant then must articulate a legitimate, non-discriminatory reason for its decision to terminate the plaintiff; and, if the defendant meets its burden of production, “the plaintiff must then offer sufficient evidence to create a genuine issue of material fact’ either (1) that the defendant’s reason is not true, but is instead a pretext for discrimination (pretext alternative); or (2) that the defendant’s reason, while true, is only one of the reasons for its conduct, and another “motivating factor” is the plaintiff’s protected characteristic (mixed-motive[s] alternative).’

* * *

If a plaintiff demonstrates that age was a motivating factor in the employment decision, it then falls to the defendant to prove “that the same adverse employment decision would have been made regardless of discriminatory animus. If the employer fails to carry this burden, plaintiff prevails.”⁸⁰

B. Title VII Actions

In actions brought under Title VII, claimants sometimes resort to statistics to establish discrimination. When such statistical evidence falls short of proving a prima facie case of discrimination, it is often offered to examine whether an employer’s proffered reason for its actions is pretextual, or to raise a triable issue that a job requirement has a disparate impact.⁸¹

Although a Title VII plaintiff is not required to prove discrimination with scientific certainty, statistics may be so incomplete as to be wholly irrelevant. Statistical analyses must be meaningful and reliable. The appropriate degree of statistical refinement often depends on the quality and control of available data.⁸²

The most frequent use of statistics in class-based Title VII cases involves claims of disparate treatment pattern and practice, and disparate impact.⁸³ To establish a prima facie case of pattern and practice disparate treatment discrimination, plaintiffs often resort to statistics in an effort to show that

⁸⁰ *Rachid*, 376 F.3d at 312 (citations omitted).

⁸¹ *International Broth of Teamsters v. U.S.*, 431 U.S. 324, 339, 97 S.Ct. 1843, 1856, 45 L.Ed.2d 280 (1977); 8 Emp. Coord. Employment Practices §107:97, citing *Bodnar v. Motorola, Inc.*, 967 F.2d 584 (9th Cir. 1992) (unpublished).

⁸² *Id.*

⁸³ *Id.*, citing *Hazelwood School Dist. v. U.S.*, 433 U.S. 299, 97 S.Ct. 2736, 53 L.Ed.2d 768 (1977) and *Griggs v. Duke Power Co.*, 401 U.S. 424, 91 S.Ct. 849, 28 L.Ed.2d 158 (1971).

the percentage of workers of the covered group in the employer's work force is significantly lower than the percentage of qualified group members in the relevant labor market.⁸⁴

1. Proper Time Frame for Statistical Analysis

Of importance to a court in evaluating the relevancy and weight of statistical testimony are whether the statistics relate to employment decisions made before or after the applicable federal statute, and whether such decisions occurred inside or outside the time period at issue in the suit.⁸⁵

In Hazelwood School Dist. v. U.S., 433 U.S. 299,300, 97 S.Ct. 2736, 53 L.Ed.2d 768 (1977), the Supreme Court held that the circuit court erred in disregarding the possibility that the plaintiff's prima facie statistical proof might be rebutted by statistical evidence of the employer's post-Act hiring practices: "once a prima facie statistical case has been established by statistical work-force disparities, the employer must be given an opportunity to show that 'the claimed discriminatory pattern is a product of pre-Act hiring, rather than unlawful post-Act discrimination.'" (citation omitted).

The Supreme Court later held in a case involving an employer charged with not having eradicated the continuing effects of its pre-Act discriminatory pay structure that the appeals court should have examined whether the pay discrepancies shown by the statistics were merely a continuation of earlier racially discriminatory action.⁸⁶

2. Choosing the Relevant Labor Market

Selecting the appropriate statistical base against which to compare data demonstrating an employer's employment decisions is critical, if not dispositive, in cases where statistics are offered to prove or disprove discrimination. This base is referred to as the "relevant labor market", and is often

⁸⁴ *Hazlewood*, 433 U.S. 299, 307-08, 97 S.Ct. 2736, 2741, 53 L.Ed.2d 768 (1977); 8 Emp. Coord. Employment Practices §107:99

⁸⁵ 8 Emp. Coord. Employment Practices §107:95

⁸⁶ *Bazemore v. Friday*, 478 U.S. 385, 106 S.Ct. 3000, 92 L.Ed.2d 315 (1986).

defined as the “general population”, “civilian labor force”, “qualified labor market”, or the employer’s own workforce composition, depending on the nature of discrimination alleged and the qualifications required for the position at issue. Also of great strategic importance is the choice of geographic boundaries within which these labor markets are located. As a general rule, when a plaintiff relies on statistics, defense counsel should attempt to utilize the qualified labor market whenever appropriate because it typically provides a more favorable statistical base for the employer than the general population.

Title VII imposes no requirement that an employer’s work force mirror the general population.⁸⁷ However, defendants may show that a workforce composition is consistent with that of the general population in the area to rebut prima facie cases of discrimination.⁸⁸ General populations tend to be the appropriate labor market where the job skill involved is one that many people possess or can readily acquire, such as unskilled, entry-level jobs, or where many vacancies are filled by hiring from the general population.⁸⁹

The civilian labor force includes only those who are employed or actively seeking work, whereas the general population encompasses many others, such as children, students, institutional inmates, military personnel, the aged and the disabled.⁹⁰ The exclusion of population segments not

⁸⁷ *International Broth. Of Teamsters v. U.S.*, 431 U.S. 324, 97 S.Ct. 1843, 52 L.Ed.2d 396 (1977).

⁸⁸ *Jimmerson v. Kisco Co., Inc.* 404 F.Supp 338 (E.D. Mo 1975), aff’d, 542 F.2d 1008 (8th Cir. 1976); *Taylor v. Safeway Stores, Inc.*, 524 F.2d 263 (10th Cir. 1975).

⁸⁹ 8 Emp. Coord. Employment Practices §107:115, citing *Hazelwood School Dist. v. U.S.*, 433 U.S. 299, 97 S.Ct. 2736, 53 L.Ed.2d 768 (1977), *Equal Employment Opportunity Commission v. Radiator Specialty Co.*, 610 F.2d 178 (4th Cir. 1979), *Equal Employment Opportunity Commission v. United Virginia Bank/Seaboard Nat.*, 615 F.2d 147 (4th Cir. 1980), *Payne v. Travenol Laboratories, Inc.*, 565 F.2d 895 (5th Cir. 1978).

⁹⁰ 8 Emp. Coord. Employment Practices §107:116.

likely to supply job applicants makes civilian workforce data a more reliable indicator of the available labor pool.⁹¹

In cases involving employment positions with required special qualifications, general population and workforce statistics offer little, if any, probative value, and qualified market data should be used.⁹² Such special qualifications include certifications or experience requirements,⁹³ unless such qualifications are subjective or are themselves challenged as discriminatory.⁹⁴ In instances where certain positions are filled almost exclusively by promotion or transfer from within, the relevant labor market may be the employer's own labor force.⁹⁵

In addition to the above criteria, the relevant labor market is generally further defined by an appropriate geographical boundary so as to establish the area from which an employer or union would draw its employees or members absent discrimination. Exceptions to this rule are instances in which the statistical data are defined by the act of applying for a particular job, such as analyses based on pass/fail rates of applicants, or when the subject positions are routinely filled from within the employer's existing workforce.⁹⁶

⁹¹ *Id.*, citing *Alexander v. Aeor Lodge No. 735, Intern. Ass'n of Machinists and Aerospace Workers, AFL-CIO*, 565 F.2d 1364 (6th Cir. 1977); *Gay v. Waiters' and Dairy Lunchmen's Union, Local No. 30*, 489 F.Supp 282 (N.D. Cal. 1980), *aff'd*, 694 F.2d 531 (9th Cir. 1982).

⁹² *Hazelwood School Dist. v. U.S.*, 433 U.S. 299, 97 S.Ct. 2736, 53 L.Ed.2d 768 (1977); *Anderson v. Douglas & Lomason Co.*, 26 F.3d 1277, 1286 (5th Cir.1994); *EEOC v. Olson's Dairy Queens, Inc.*, 989 F.2d 165, 168 (5th Cir.1993).

⁹³ *Coble v. Hot Springs School Dist. No. 6*, 682 F.2d 721 (8th Cir. 1982); *Croker v. Boeing Co. (Vertol Div.)*, 662 F.2d 975 (3d Cir. 1981).

⁹⁴ *Crawford v. Western Elec. Co., Inc.* 614 F.2d 1300 (5th Cir. 1980); *Kaplan v. Int'l Alliance of Theatrical and Stage Emp. And Motion Picture Mach. Operators of U.S. and Canada*, 525 F.2d 1354 (9th Cir. 1975); 8 Emp. Coord. Employment Practices §107:117.

⁹⁵ *See, e.g., James v. Stockham Valves & Fittings Co.*, 559 F.2d 310 (5th Cir. 1977).

⁹⁶ 8 Emp. Coord. Employment Practices §107:119, citing *Johnson v. Goodyear Tire & Rubber Co., Synthetic Rubber Plant*, 491 F.2d 1364 (5th Cir. 1974); *Equal Employment Opportunity Commission v. Local 14, Intern. Union of Operating Engineers*, 553 F.2d 251 (2d Cir. 1977).

Factors relevant to a determination of the proper geographic scope of the labor market include commuting patterns, availability of mass transit, area of recruiting efforts by employer, recruiting practices of competing employers, residences of current employees, and the degree of skill needed for the position as issue.⁹⁷ Geographic areas that have been regarded by the courts as appropriate include the city, county and state in which the employer is located, the metropolitan area used by the Census Bureau in computing statistics about the makeup of a given geographic location, and the entire country.⁹⁸

3. Statistical Rebuttal

In disparate treatment pattern and practice claims, a defendant responds to a prima facie case of statistical evidence either by demonstrating the unreliability of the statistics or offering different evidence that explains why the results of a policy or practice do not reflect intentional discrimination.⁹⁹ Some common bases for rebutting prima facie cases of pattern and practice disparate treatment discrimination based on statistics include the following:¹⁰⁰

- (a) the special qualifications of the job in issue are not reflected in the statistical data, thus distorting the relevant labor market on which the inference of discrimination is based (e.g., the relevant labor pool for a teaching job would consist of only those with teaching certifications, not the entire unemployed population surrounding the job site);¹⁰¹
- (b) the statistical proof distorts the picture of the employer's actions because it does not include applicant flow information (i.e., the employer made no actual employment decision concerning the qualified and available members of the protected group concerned, unless they applied);¹⁰²

⁹⁷ 8 Emp. Coord. Employment Practices §107:121, citations omitted.

⁹⁸ 8 Emp. Coord. Employment Practices §107:122, citations omitted.

⁹⁹ 8 Emp. Coord. Employment Practices §107:105, citing *Hazelwood School Dist. v. U.S.*, *supra*.

¹⁰⁰ 8 Emp. Coord. Employment Practices §107:106.

¹⁰¹ *Hazelwood School Dist. v. U.S.*, 433 U.S. 299, 97 S.Ct. 2736, 53 L.Ed.2d 768 (1977)

¹⁰² *Valentino v. U.S. Postal Service*, 511 F.Supp 917 (D.D.C. 1981), *aff'd* 674 F.2d 56 (D.C. Cir. 1982).

- (c) the size of the statistical sample (number of actual employment decisions) used in presenting the inference of disparate treatment is too small to be of meaningful value;¹⁰³
- (d) the statistics do not reflect a meaningful difference between the manner in which the employer treats the protected group in question, and the way it treats others;¹⁰⁴
- (e) the time period covered by the statistical evidence includes employer practices that fail adequately to consider changes in relevant circumstances, such as changes in employment patterns after the effective date of Title VII¹⁰⁵, or the employer's recent corrective action;¹⁰⁶
- (f) the time period covered by plaintiff's statistical evidence includes a time period prior to the effective date of Title VII, when the employer was not liable for discrimination under the act.¹⁰⁷

These rebuttals are equally applicable to statistical evidence used to support a claim of disparate impact discrimination. Rebuttals to statistics must do more than simply point out potential sources of error, however. They should demonstrate that the claimed error is systematic, that it actually biased the results, and that the bias was sufficiently extensive that it would affect the validity of the plaintiffs' demonstration of discrimination.¹⁰⁸

Statistical rebuttal to a prima facie case in an individual disparate treatment action will generally be less effective, however, because demonstrating a balanced workforce does not necessarily demonstrate that a specific act was without discrimination.¹⁰⁹ Likewise, a history of nondiscriminatory

¹⁰³ *Presseisen v. Swathmore College*, 442 F.Supp. 593 (E.D.Pa. 1977), aff'd without op 582 F.2d 1275 (3d Cir. 1978);

Hagans v. Andrus, 651 F.2d 622 (9th Cir. 1981).

¹⁰⁴ *Presseisen v. Swathmore College*, *supra*.

¹⁰⁵ *Allen v. Prince George's County, Md.*, 538 F.Supp. 833 (4th Cir. 1984).

¹⁰⁶ *Blizard v. Frenchette*, 601 F.2d 1217 (1st Cir. 1979).

¹⁰⁷ *Hazelwood School Dist. v. U.S.*, 433 U.S. 299, 97 S.Ct. 2736, 53 L.Ed.2d 768 (1977).

¹⁰⁸ 8 Emp. Coord. Employment Practices §107:107, *citing Berger v. Iron Workers Reinforced Rodmen Local 201*, 843 F.2d 1395 (D.C. Cir. 1988).

¹⁰⁹ *Furnco Const. Corp. v. Waters*, 438 U.S. 567, 579, 98 S.Ct. 2943, 2950-51, 57 L.Ed.2d 957 (1978).

behavior, while instructive as to an employer's motive, cannot rule out the possibility that an employer has recently begun to discriminate.

VI. Conclusion

The use of statistical analysis in litigation is a potentially powerful tool for claimants and defendants. However, its relevance depends on whether the data on which it is based are sound and reliable, and the analysis thereof is scientifically sound. Establishing the reliability or unreliability of statistical analysis is the critical task of litigants and their expert witnesses, and the successful execution of this duty can determine a case's outcome.

A significant shortcoming of statistical analysis to prove causation in toxic tort or product liability actions is that the epidemiological studies on which they are often based reveal only *associations* between a disease or condition and a factor suspected of causing it, not direct causation. It is generally acknowledged that science is unable to determine exactly what caused a particular injury in such circumstances.

Two of the great pitfalls of using statistics as evidence in the context of extrapolating statistical data for trial planning or demonstrating causation and damages are that such use denies defendants their opportunity to cross-examine opposing witnesses and denies them due process rights to which they are entitled.

It is essential that our judiciary remain ever cognizant of the potential injustices resulting from these seductive evidentiary shortcuts and vigilantly guard against them. Similarly, it is the duty of litigators to test and critically examine any such statistical evidence to reveal whether it is reliable and relevant.