Does (Sample) Size Matter?

Identifying and Challenging Unreliable Methodologies Employed in Statistical Extrapolation

By Mark D. Shifton and Benjamin S. Wilner

Statistics is a science, and expert opinions seeking to extrapolate results based on statistical sampling must adhere to well-recognized statistical sampling methodologies.

Statistical sampling and extrapolation are important—yet often ignored—concepts in many cases involving claims of construction or product defect. Adverse technical experts will sometimes inspect small, cherry-picked samples of those units most likely to support a plaintiff's claims and then paint their opinions with a broad brush by concluding that their results may be extrapolated across the entire case, based solely on the examination of a select few units. Such opinions, if left unchallenged, can often significantly increase the potential damages. When the methodologies of such opinions are aggressively probed, however, the flaws underlying the ultimate conclusions that extrapolate the opinions across an entire population often become clear. By challenging the reliability of the extrapo-

■ Mark D. Shifton is a senior partner in the Princeton, New Jersey, and New York City offices of Seiger Gfeller Laurie LLP. Mr. Shifton represents owners and developers, contractors, and design professionals in construction and professional liability disputes and advises and represents clients on sustainable development and green-building issues. He currently serves as the program co-vice chair of the DRI Construction Law Committee and is the immediate past chair of the Building Products SLG of the DRI Product Liability Committee. Benjamin S. Wilner, Ph.D., a Managing Director at Alvarez & Marsal, is an economist and statistician. He regularly serves as a consultant and expert witness on economic, financial, and statistical issues around the world.

lation, the potential damages may be limited significantly.

While most experienced attorneys would be able to probe the weaknesses in such technical experts' opinions qualitatively, using a statistical expert (along with your own technical expert) can significantly support demonstrating the flaws (and potential unreliability) in an adverse expert's extrapolation conclusion. Many attorneys have little understanding of these concepts, however, and thus often fail to challenge expert opinions that are predicated on unreliable sampling methodologies properly. As one court has noted, "[l]awyers and judges working with statistical evidence generally have only a partial understanding of the selection processes they seek to model, they often have incomplete or erroneous data, and are laboring in an unfamiliar terrain." *Waisome v. Port Authority of N.Y. & N.J.*, 948 F.2d 1370, 1372 (2d Cir. 1991). Having a basic understanding of statistical sampling and extrapolation mechanics provides a powerful tool for litigators who are willing to become involved with the necessary concepts and terminology.

As any attorney defending claims of construction or product defect knows, nearly every such claim must be supported by the opinion of a qualified expert (generally as pertains to both liability and damages). In most cases, inspecting every product at issue is impossible or impractical, or it is not done. For instance, in a class action involving an allegedly defective product, where millions of units have entered the stream of commerce, there is often no practical method to inspect all relevant units. Likewise, in a construction-defect action, not every window in a thousand-unit condominium complex can be subjected to tests for water infiltration. Even from a purely objective standard, where the search for "truth" about whether a product is defective is the goal, a sampling analysis will require inspection of a sample of the product, followed by a statistical analysis and opinion deciding whether the results of the inspection may be extrapolated further than the sample actually inspected. This is the very nature of statistical extrapolation: after testing a sample, the expert might find that the number of units determined to be defective were statistically significantly excessive. Based on that opinion, the finder of fact may conclude that a large percentage of the products in question are indeed defective, without requiring the plaintiff to establish that defect in each unit. As the Fifth Circuit has noted, "[t]he essence of the science of inferential statistics is that one may confidently draw inferences about the whole from a representative sample of the whole." *In re Chevron U.S.A., Inc.*, 109 F.3d 1016, 1019–20 (5th Cir. 1997).

Several courts around the country have endorsed the use of statistical sampling to support claims or defenses—in many types of cases—including mass torts, building product defect, and wage-and-hour claims. *See, e.g., Cimino v. Raymark Indus.*, 151 F.3d 297, 319–20 (5th Cir. 1998) (mass torts); *Duran v. U.S. Bank Nat'l Assoc.*, 325 P.3d 916, 937–38 (Cal. 2014) (wage and hour claim); *Kadas v. MCI Systemhouse Corp.*, 255 F.3d 359 (7th Cir. 2001) (Age Discrimination in Employment Act claim). Other courts, however, have raised doubts about using statistical sampling to determine certain issues. *See, e.g., Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 348, 367 (2011) (reversing a grant of class certification, for which the plaintiffs relied on statistical sampling and extrapolation to establish back-pay claims in a wage-and-hour class action, on the grounds that statistical extrapolation would deprive the defendant of its right to assert statutory defenses to individual claims); *United States v. Jones,* 641 F.3d 706, 712 (6th Cir. 2011) (vacating a criminal defendant's prison sentence on the grounds that evidence based on statistical sampling proffered by the government during sentencing was not sufficiently representative to be reliable).

Few bright-line rules have been established to determine whether an expert's opinion relying on sampling should be admissible. Nevertheless, expert opinions increasingly offer extrapolation opinions, even when those opinions are offered without regard to well-accepted principles of statistical sampling. Expert opinions that seek to extrapolate the results from testing a limited sample, or misapply statistical sampling (whether intentionally in an effort to bolster a plaintiff's claims, or due to ignorance of proper methods) may be unreliable—and if challenged appropriately—vulnerable to preclusion in trials. Accordingly, as mentioned, having a working knowledge of the basic concepts underlying statistical sampling will often prove an extremely potent weapon in defense counsel's arsenal when faced with an expert opinion that draws from statistical sampling.

This article highlights the basic concepts of statistical sampling, identifies some of the more common flaws in statistical sampling methodologies, and discusses how defense counsel can identify and exploit such flaws. By understanding these issues, defense counsel will be in a better position to evaluate the strengths (or weaknesses) in an adversary's expert's opinion, identify the flaws underlying those opinions, and defend their clients when an unreliable statistical sampling methodology is used in cases.

> **Several courts** around the country have endorsed the use of statistical sampling to support claims or defenses—in many types of cases—including mass torts, building product defect, and wage-and-hour claims.

Statistical Sampling: The Basic Concepts

Statistical sampling is the process of collecting and analyzing data from a subset of an overall population, and then extrapolating the results of that sample across the entire population. For the average litigator's purposes, understanding the concepts underpinning statistical sampling and extrapolation does not require advanced study of statistics, but having an understanding of the basic terminology and concepts is critical. Statistical sampling and extrapolation is, typically, outside the realm of knowledge familiar to the average juror, and thus any opinion based on a statistical sampling must be offered by an expert qualified to offer such an opinion. With extrapolation, a statistical expert is often paired with a "technical" expert; the technical expert will examine the samples and generate a conclusion regarding the samples. The statistical expert will then offer an opinion pertaining to the statistical significance of the technical expert's findings and offer an opinion about whether the results of the underlying inspection can be extrapolated across a broader population.

When confronted with an opinion that offers a conclusion based on a statistical extrapolation, one must probe how closely the samples tested represent the entire population.

Performing statistical sampling and extrapolation requires a detailed, step-by-step procedure, a full discussion of which is far beyond the scope of this article. Three critical concepts, however, are essential to understanding the process of statistical sampling, and in recognizing whether an expert's opinion is sufficiently reliable to support the expert's conclusion.

Representativeness

The *representativeness* of a sample describes whether the sample being studied accurately reflects the overall population. When confronted with an opinion that offers a conclusion based on a statistical extrapolation, one must probe how closely the samples tested represent the entire population. The more representative the sample is of the population, the greater the likelihood that the conclusion can be reliably extrapolated to the population. Samples that are "cherry-picked" are less likely to be truly representative of the population, and any opinion based on an analysis of such a sample is less likely to be reliable.

Margin of Error

A statistical extrapolation's precision is often measured by its margin of error, which measures the amount by which the results of the analysis likely would be different if the analysis were repeated. A statistical analysis by its very nature results in a conclusion about an overall population based on an examination of a less-than-complete subset of the population. Accordingly, there is the possibility that any statistical analysis, if repeated, would yield a different result. Stated differently, the margin of error represents the "plus or minus" range in the sampling result and describes the magnitude by which the result likely could differ if the analysis were to be repeated. A small margin of error implies a greater likelihood that the analysis's results are reliable, while a large margin of error reflects the increased chance that if the analysis were repeated, the result would be sizably different.

To illustrate this important concept, assume that 100,000 units of a certain product are manufactured. Even if an expert inspects 99,950 units (99.95 percent of the entire population) and finds all of them to be free of defects, the expert cannot necessarily conclude that the remaining 50 units are also free of defects. The expert may, however, conclude that a very small number of units, at most, could have a defect. Despite the expert having found no defects in the units inspected, if the analysis were to be repeated, (and were the expert again to examine 99,950 units out of the 100,000 manufactured), there remains a chance that a small percentage of the total units could be found defective; the margin of error effectively measures this percentage.

Mathematical Modeling

Different sampling methodologies may be employed, each of which requires different mathematical modeling. Once the sample to be inspected is chosen, a mathematical model must be applied to analyze the sample and form a conclusion. While a discussion of the various mathematical models used during statistical analysis is well beyond the scope of this article, it is sufficient to note that using a mathematical model that is inconsistent with the sampling methodology will often lead to an improper (and thus likely unreliable) conclusion. Very few technical experts are sufficiently familiar with these mathematical models that are used in any given analysis, and thus, without input from a statistician, they are prone to drawing unreliable conclusions.

The Steps Involved in Sampling Extrapolation

The three core concepts explained above—*representativeness*, *margin of error*, and *mathematical modeling*—underlie seven standard mathematical steps of a typical statistical sampling analysis. While these critical concepts are often dismissed in cavalier fashion by technical experts, and often by attorneys, as "just math," statistical analyses that fail to apply these concepts properly run the risk of being found unreliable, and thus they are vulnerable to preclusion at trial. Understanding the three concepts is integral to understanding the steps involved in conducting a statistical analysis and an extrapolation, to which we now will turn.

Population

The first step in conducting a statistical analysis is to define the *population*, meaning the entire "universe" to be studied. In a product liability case, the population could contain all units of an allegedly defective product, or in a class action, the entire putative class, while in a construction-defect action, the population could be every window in a large condominium development. Because inspecting every existing product unit, testing every single window, or interviewing every putative class member would be impractical, (or at the very least, cost-prohibitive), statistical sampling is often put to use to inspect a representative sample of the population.

Sampling Frame

The second step is to define the *sampling frame*, meaning the source (or sources) from which the samples to be studied are drawn. Because the population is the entire universe of possible units, the sampling frame may differ significantly from the population. In a class action, if a putative class of plaintiffs consists of all customers of a particular company during a given year, (the population), only those individuals for whom there is valid contact information may be available to be interviewed. Accordingly, although there may be thousands of other individuals in the population, the sampling frame is limited to those samples that can actually be part of the analysis. Based on the unique facts of the study, the sampling frame may be either under inclusive or over inclusive relative to the population. Significant differences between the sampling frame and the population may call the representativeness of a study—and accordingly, any conclusions drawn from it—into question, especially if the population as a whole is sufficiently different from the sampling frame. *See, e.g.*, *Temple-Inland, Inc. v. Cook*, 192 F.Supp.3d 527, 549 (D. Del. 2016) (denying the defendant's motion for summary judgment in action to recover unclaimed property on grounds that the defendant's damages argument was based on the conclusion that extrapolated results of statistical sampling to years not in the sampling frame).

To illustrate the importance of the concept of the sampling frame, assume an expert inspects one hundred units of a certain product alleged to be defective. If the units were taken directly from the production line without any identifying features during the entire production period, the expert's conclusions might be reliably extrapolated to the full product line. In contrast, if the units inspected were only those that had been returned to the manufacturer by aggrieved customers, the sampling frame likely would be markedly different than the population, and an analysis of the population would be far more likely to result in a different conclusion. In this scenario, the expert's conclusions might only be extrapolated to units that had been returned to the manufacturer, not the entire universe of products. Given the significant differences in the sampling frames, the latter opinion would be far less reliable than the former.

Sampling Methodology

A critical step in conducting a statistical sampling is selecting the *sampling methodology*, meaning the method by which the subjects within the sampling frame are selected to be part of the study. The sampling methodology may be developed or guided by a statistician (or at least an expert who is familiar with the concepts of statistical sampling), or it may be performed without regard to any methodology at all. Statisticians (and courts) generally prefer randomized, controlled experiments. David H. Kaye & David A Freedman, Reference Guide on Statistics, Reference Manual on Scientific Evidence (Fed. Jud. Ctr., 3d ed. 2011). As one might expect, random sampling occurs when samples are selected from the sampling frame at random. In a construction-defect action alleging water infiltration through the windows of a condominium complex, a random sample might be obtained by assigning every window a number and using a random number generator to select those windows that will be part of the study.

Other, less objective sampling methodologies are often seen in litigation, including convenience sampling (in which samples are selected based on the ease or cost of analysis). Inspecting only those windows that are easily accessible without the need for mobilizing a scaffold or those windows located in units with cooperative and friendly unit-owners would fall into the convenience-sampling category. Another type is judgmental sampling, in which samples are specifically selected for inclusion in the study based on a subjective belief that such items are relevant, such as deliberately choosing to inspect only those windows that have a reported history of water infiltration and deliberately excluding any windows without such a history. Courts have often questioned these other sampling methodologies. *See Duran,* 325 P.3d at 940 ("[C]onvenience alone cannot justify procedures that substantially curtail the parties' ability to litigate their case").

Sample Size

The next step involves selecting the number of units ultimately inspected, which is referred to as the *sample size*. Generally, a statistical analysis with a larger sample size will decrease the study's margin of error and increase the study's reliability. Statisticians generally use mathematical formulas before sampling begins to determine the proper sample size that will yield a result within an acceptable margin of error. A large margin of error does not in and of itself render a sampling invalid or unreliable because a sample may be statistically valid regardless of the sample size. Even if the results of a sampling are statistically valid, however, a sufficiently large margin of error may call the results into question.

A large margin of error does not in and of itself render a sampling invalid or unreliable because a sample may be statistically valid regardless of the sample size. Even if the results of a sampling are statistically valid, however, a sufficiently large margin of error may call the results into question.

See Duran, 325 P.3d at 943 (questioning whether sampling results with a margin of error of near 50 percent could be considered reliable).

A concept closely related to margin of error is the confidence level, which measures the degree of confidence that an expert has in a given margin of error. Different confidence levels could be utilized in an analysis. Generally, the larger the margin of error, the greater confidence the expert will have in the results. To illustrate, in performing a statistical analysis, an expert might be 99 percent confident that the margin of error is less than 4 percent, roughly meaning that if the study were repeated (with different units chosen from the same sampling frame), there would be a 1 percent chance that the second study would yield results outside of the 4 percent margin of error. Alternatively, with that same data, the expert might be 90 percent confident that the analysis has a margin of error of less than 1 percent in the results of the study. This generally means would mean that there would be a 10 percent chance a second study of the same data would yield results outside of the 1 percent margin of error.

Most courts have made it clear that expert opinions offering extrapolations will be admissible only if the statistical sampling and methodologies underlying those opinions are reliable.

Sample Selection

The next step involves actually selecting the samples that will be part of the study, as determined by the sampling methodology. As noted above, the chapter on statistics in the Reference Manual on Scientific Evidence, published by the Federal Judicial Center, clearly notes a preference for randomly selected samples. Accordingly, including items for study such as named plaintiffs in class action lawsuits, units purposely selected for the study, or units that were selected solely on the grounds of cost or convenience reduces the randomization of study, and it may call the study's reliability into question. *See, e.g.*, *Chevron*, 109 F.3d at 1020 (denying the defendant's writ of mandamus from the district court, which sought approval of a trial plan that included 30 bellwether claimants (15 chosen by the plaintiff and 15 chosen by the defendant) to comprise a "unitary trial" to determine the issue of liability, on the grounds that the district court failed to find that the bellwether claimants were representative of population). Similarly, in a multidistrict litigation alleging claims of product defect, the United States District Court for the District of Minnesota precluded the plaintiffs' expert's extrapolation opinion, noting the expert had "cherry picked 6 'flawed' planks out of the billions currently on U.S. structures, but seeks to opine that they are representative [of all units of the product in existence.]" *See In Re: HardiePlank Fiber Cement Siding Litigation*, 12-md-2359, MDL No. 2359, 2018 WL 262826, at *1, *9 (D. Minn., Jan. 8, 2018).

Analysis of the Sample

The sixth step is the actual technical analysis of the samples selected for the study. In a construction-defect action alleging water infiltration through the windows of a building, the actual analysis might include applying a spray-rack test to all windows selected for the study, followed by determining which windows in the study showed signs of water infiltration. In a product liability action, the analysis could consist of testing all samples selected for the study and determining how many of those units failed to meet the test's criteria. In this step, it is the actual expertise of the technical expert—not the statistician—that is important. To ensure the reliability of the result, the methods used during the analysis should be based on any relevant industry or consensus standards (such as ASTM, ASME, or some other generally accepted guidelines). The technical expert should be able (and qualified) to explain the testing method and criteria for determining whether given units passed or failed the test, defend the testing methodology, and explain why the testing methodology was selected over other available methodologies. Finally, the testing procedure must be replicable.

Extrapolation of the Sample Results

The final step in a statistical sampling is the statistical extrapolation of the sample results to the sampling frame. After the population and sampling frame have been determined, the methodology for selecting the sample has been decided on, and the samples inspected, the statistician must evaluate the data and determine what conclusions, if any, may be drawn. For each sample, the statistician relies on the conclusions drawn by the technical expert during the analysis (such as the number of windows tested that revealed signs of water infiltration, or the number of product failures). In performing the extrapolation, the statistician must use the proper mathematical model. Using an improper mathematical model will likely lead not only to an incorrect result, but it may render the expert's opinion vulnerable to preclusion in a trial.

Ultimately, the statistical expert may conclude that based on the samples tested, a certain characteristic (such as a product defect) can be extrapolated across the entire sampling frame. Many opinions seeking to extrapolate the results of a statistical analysis—especially when offered by a technical expert unfamiliar with proper statistical sampling—contain significant flaws in sampling technique or methodology, calling the reliability of such opinions into question. Most courts have made it clear that expert opinions offering extrapolations will be admissible only if the statistical sampling and methodologies underlying those opinions are reliable. *See, e.g., In Re: HardiePlank*, 2018 WL 262826, at *9 (precluding the plaintiffs' extrapolation opinion, noting that the expert "failed to document how the samples were selected; nor did he have a sampling plan," the expert offered no "opinion whether his samples were statistically significant or their error rate," and that ultimately, the expert "provides no basis to extrapolate his results."). *See also Anghel v. Sebelius*, 912 F.Supp.2d 4, 18 (E.D.N.Y. 2012) ("the sampling utilized need not be based on the most precise methodology, just a valid methodology") (quoting *Miniet v. Sebelius*, No. 10-24127-CV, 2012 WL 2930746, *1, *6 (S.D. Fla., July 18, 2012).

When faced with an expert opinion offering an extrapolation based on an inspection of only a small population, or of samples that were clearly selected for purposes of convenience or subjective criteria, attacking the opinion on the grounds of unreliability is often an effective strategic decision—one which may have a significant, if not fatal, effect on the case as a whole.

Expert Opinions Based on Statistical Sampling: A Question of Reliability

Statistical sampling and extrapolation are often necessary, especially in cases such as construction-defect and product liability cases, where it would be effectively impossible (or at least impractical) for a party to have its expert inspect each and every unit in a given population. While there are few bright-line rules relating to the admissibility of expert opinions that rely upon statistical sampling, one unifying principle is generally clear: statistical extrapolations are expert opinions in their own right, and they will generally be admissible if they meet the criteria subject to the admission of expert opinions generally. In the federal courts, that is the standard set forth by the United States Supreme Court in *Daubert v. Merrill Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), and later written into Rule 702 of the Federal Rules of Evidence, which states:

> A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:
> a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
> b) the testimony is based on sufficient facts or data;
> c) the testimony is the product of reliable principles and methods; and
> d) the expert has reliably applied the principles and methods to the facts of the case.

While there are relatively few decisions discussing the concept of statistical sampling and extrapolation in the construction-defect or product liability context, there are numerous decisions dealing with this issue in other contexts. In many of these decisions, the primary indicator of whether an expert opinion based on statistical sampling will be held admissible is the reliability of the sampling methodology underlying the opinion. Ultimately, an expert opinion based on statistical sampling bears little difference from any other opinion offered by an expert in another field. It is subject to the same gatekeeping function to be performed by the court under *Daubert, Frye*, or other relevant admissibililty standards. Should the methodology underlying such an opinion prove reliable, and should it be helpful to the finder of fact, the extrapolation opinion will likely be admitted.

Most courts faced with the issue of statistical sampling and extrapolations have taken a relatively broad view and have admitted or precluded an expert's extrapolation opinion on the basis of its reliability as a whole. *See, e.g., Massachusetts Mutual Life Insurance Company v. Residential Funding Company, LLC*, 989 F.Supp.2d 165, 172–75 (D. Mass. 2013) (denying a motion to preclude a statistician's expert opinion on the grounds that the proposed opinion was sufficiently reliable); *Wielgus v. Ryobi Technologies, Inc.*, No. 08 CV 1597, 2012 WL 3614642, at *1, *4–5 (N.D. Ill., Aug. 21, 2012) (denying the defendant's motion to preclude the plaintiff's expert's opinion on the grounds that the defendant failed to establish sufficient indicia of unreliability of the data relied on by plaintiff's expert); *Robert Dillon Framing, Inc. v. Canyon Villas Apt. Corp.*, Nos. 55897, 57122, 57927, 2013 WL 3984885, at *1, *3–4 (Nev., Apr. 17, 2013) (affirming the trial court's admission of the statistician's opinion based on extrapolation from limited testing). *See also Duran,* 325 P.3d at 933–46 (affirming the appeals court's decertification of class in a wage-and-hour class action, on the grounds that the statistical sampling methodology of the class members was unreliable and engaging in a comprehensive analysis and discussion of the principles of statistical sampling and extrapolation).

On a small number of occasions, courts have been confronted with the issue of statistical sampling and extrapolation in the context of construction-defect cases. In *Robert Dillon Framing, Inc. v. Canyon Villas Apt. Corp.*, the Supreme Court of Nevada affirmed a trial court's admission of a statistician's expert opinion. Nos. 55897, 57122, 57927, 2013 WL 3984885, at *1, *3–4 (Nev., Apr. 17, 2013). After a jury verdict in the defendant's favor, the plaintiff appealed, on the grounds that the trial court had improperly admitted the defendant's statistician's expert opinion at trial. The opinion was based on water infiltration testing performed on two percent of windows in a residential development. In affirming the admission of the statistician's testimony, the court noted that the units tested were selected at random and an equal number of units on different floors were tested, as were different types of apartments within the development. In affirming the reliability of the statistician's opinion, the court further noted that given the random method of selecting the windows to be sampled, even increasing the sample size would have yielded the same result, within the same margin of error.

Two federal district courts have recently been confronted with *Daubert* challenges to proposed expert opinions based on statistical analyses, both of which sought to use data collected by the United States Consumer Product Safety Commission (CPSC) to prove liability for the defective design of table saws, with each reaching contrary results. In *Hilaire v. DeWalt Industrial Tool Co.*, the U.S. District Court for the Eastern District of New York granted a defendant's motion to preclude a plaintiff's expert's opinion, declaring that the plaintiff's expert's conclusion (based on a statistical extrapolation from the CPSC's data) was "without scientific rigor." 54 F.Supp.3d 223, 245–46 (E.D.N.Y. 2014). In precluding the expert's opinion, the court noted that the CPSC data on which the expert relied in forming his conclusion was inherently unreliable, and thus to the extent that the expert's opinion was based on an extrapolation of the CPSC's data, the opinion was insufficiently reliable under *Daubert* to be admissible. *Id.* at 246. In contrast, the U.S. District Court for the Northern District

Most courts faced with the issue of statistical sampling and extrapolations have taken a relatively broad view and have admitted or precluded an expert's extrapolation opinion on the basis of its reliability as a whole.

of Illinois denied a similar motion to preclude a plaintiff's expert's proposed testimony, holding that the expert's reliance on the CPSC data to support his opinion that the subject product was defective was acceptable. *Wielgus v. Ryobi Technologies, Inc.*, No. 08 CV 1597, 2012 WL 3614642, at*1, *4–5 (N.D. Ill., Aug. 21, 2012). While the defendant argued that the CPSC's data

During depositions, opposing experts should be questioned extensively about which units were part of the sample, why those units were selected to be part of the sample, and by whom.

was based on its own statistical extrapolation (and thus the plaintiff's expert opinion essentially involved an extrapolation within an extrapolation), the court nevertheless admitted the expert's proposed opinion, noting that while the defendant's motion to preclude the plaintiff's expert's opinion criticized the plaintiff's expert's use of the CPSC data, it failed to provide any evidence that the CPSC data itself was unreliable, or demonstrate how the plaintiff's expert's reliance on the data rendered his ultimate opinion unreliable. The differing results between *Hilaire* and *Wielgus* are likely explained by the manner in which the defendants' motions to preclude in the two cases were framed. In *Wielgus*, while offering criticisms of the plaintiff's expert's opinion, the defendant's motion to preclude was apparently grounded in evidentiary objections to the admissibility of the data underlying the CPSC data, and not toward the expert's use of that data in forming his opinion.

Several courts have admitted expert opinions grounded in statistical evidence provided that the underlying analysis bears a certain level of statistical significance, usually held to be a significance level of five percent. (That is, the opinion was admitted if the expert opined that there was no more than a five percent probability that a statistical correlation between the dependent and independent variable would be observed in the overall population). *See, e.g.*, *Anderson v. Douglas & Lomason Co., Inc.*, 26 F.3d 1277, 1291 n.26 (5th Cir. 1994); *Ottaviani v. State University of New York*, 875 F.2d 365–71–72 (2d Cir. 1989). While most statisticians consider a confidence level of 95 percent acceptable, the Seventh Circuit has rejected applying a "bright-line" rule in admitting or precluding an expert opinion solely on the basis of the study's confidence level, holding that the issue of whether a particular confidence level assigned to a study is sufficiently reliable to be admissible is a question of law for the court to determine, not a de facto "industry standard" that the court is obligated to accept. *See Kadas v. MCI Systemhouse Corp.*, 225 F.3d 359, 363 (7th Cir. 2001) ("It is for the judge to say, on the basis of the evidence of a trained statistician, whether a particular significance level, in the context of a particular study in a particular case, is too low to make the study worth the consideration of judge or jury"). Courts in other jurisdictions have taken a similar approach, rejecting the blanket notion that a particular significance level attached to a study should determine whether an expert's opinion is admissible. *See, e.g.*, *Heller v. Shaw Indus. Inc.*, 167 F.3d 146, 158 (3d Cir. 1999); *Waisome v. Port Authority of New York and New Jersey*, 948 F.2d 1370, 1376 (2d Cir. 1991).

Putting Theory into Practice: Confronting Expert Opinions Offering Statistical Extrapolation

Proper statistical sampling and analysis requires adherence to principles that are often outside the expertise of many technical experts. Expert opinions seeking to extrapolate the results of sampling, however, must be grounded in reliable sampling methodologies. From a review of some of the decisions of the various courts faced with this issue, we can discern which issues courts are likely to seize on when deciding whether to admit or preclude an expert opinion and develop a set of best practices that savvy defense counsel may use when confronted with such opinions.

When served with an adverse expert's opinion, defense counsel should determine whether the opinion actually offers the conclusion that the results of testing can be extrapolated. Often this conclusion is not made by a technical expert but by the adverse party's damages expert. During a deposition, a technical expert will often freely concede that an inspection was limited to certain units and will decline to offer the opinion that the conclusion may be further extrapolated. A plaintiff's damages expert's opinion, however, will be predicated on the assumption that the technical expert's opinion can (and often *should*) be extrapolated. Accordingly, securing concessions from the plaintiff's technical expert pertaining to the limits of the opinions being made may significantly call into question whether the plaintiff's damages expert may justifiably rely on those opinions in making an extrapolation.

Faced with an extrapolation opinion, and with only a basic grounding in the concepts discussed in this article, defense counsel should be able to determine whether the samples tested by an expert were either "cherry-picked" for the study, or were based on an unacceptably small sample size, thus calling the reliability of the study into question. During depositions, opposing experts should be questioned extensively about which units were part of the sample, why those units were selected to be part of the sample, and by whom. Additionally, opposing experts should be questioned about whether any extrapolation opinion resulted from an actual statistical analysis. If so, have the expert explain the details concerning the mathematical model that was use, the margin of error, and the confidence level of the study. More often than not, experts are wholly unprepared to answer such questions, and an expert's inability to explain the methodology underlying the analysis effectively should lend significant support to a motion to preclude the extrapolation opinion.

Even if sampling and testing appear to have been conducted properly, consulting with an expert statistician may reveal significant flaws in the selection of the sampling frame, the sampling process, or the margin of error, making the opinion vulnerable to a motion to preclude. Consulting with an expert statistician early in a case should be a priority, and defense counsel should be prepared to serve an expert statistician's report in rebuttal. This report should address the deficiencies in the adverse expert's extrapolation opinion and demonstrate to the court by explaining in basic, "judge friendly" language, keeping the math to a minimum, the unreliability of the adverse expert's extrapolation opinion.

In certain cases, such as class actions, when counsel know from the outset that statistical sampling and extrapolation opinions will be necessary, consulting with a statistician early is critical. A statistician retained by defense counsel can play a critical role in helping chart the course and extent of the discovery necessary, such as the methodologies by which units are selected for study, or the class members who need to be deposed.

Conclusion

Statistics has been called the science of data collection, presentation, analysis, and interpretation. Accordingly, expert opinions seeking to extrapolate the results of testing beyond those units actually tested must not only involve reliable methodologies for the underlying testing, they must also involve reliable methodologies during the sampling process and statistical analysis. In construction-defect and product liability actions, experts will often focus their attention on establishing their underlying opinions by limiting their analyses to certain samples that were selected for the sole purpose of supporting a party's claims, and they will then offer the opinion that the results from testing those samples may be extrapolated across an entire population. Expert opinions based on statistical sampling and offering opinions extrapolating from such sampling are not improper per se (and indeed, in certain types of cases, statistical sampling may be required). Statistics, however, is a science in its own right, and expert opinions seeking to extrapolate results based on statistical sampling must adhere to well-recognized statistical sampling methodologies, or risk preclusion on the grounds of unreliability. FD