# “Lies, Damned Lies and Statistics[*]”: The use of Statistical Evidence in Criminal and Civil Law

Mary M. THOMSON[**]



---

[*] The Full quotation attributed to Benjamin Disraeli is, “There are three kinds of lies: lies, damned lies and statistics.”

[**] McCarthy Tétrault, Barristers & Solicitors, Toronto, ON.

On May 25, 2000, the Supreme Court of Canada delivered its judgment dismissing leave to appeal applications in a class action certification commenced in Ontario in 1997.[1] The class certification followed a well-publicized outbreak of Hepatitis B amongst patients of several electroencephalographic clinics in Toronto which were owned and run by a neurologist and an encephalography technician. The class now certified consists of all patients who contracted Hepatitis B after being given an EEG at one of the clinics, or the estates of those now deceased, together with their derivative and *Family Law Act* claimants and all patients who did not contract Hepatitis B but who nonetheless were sent notice by health officials to be tested and were tested, or their estates.

What is notable about this case is the sheer numbers of potential claimants involved. By the end of the investigation by Public Health, 18,567 former patients of the clinics had been identified and efforts had been made to contact them. A number of initial information packages were returned unopened, but some 15,610 mail packages were not. Contact was made with those patients. Eventually, data was secured from some 10,244 patients. Of these, the plaintiffs allege that 962 meet the definition of "possible" cases. They allege further that 87 patients had had an EEG at one of the clinics and had tested positive as being Hepatitis B carriers. 9,286 patients were notified by Public Health and tested negative. They claim as part of the class. The defence takes a very different view of the number of patients to be included.

The *Anderson* case together with the Hepatitis C class actions against the Canadian Red Cross Society have heralded a new era of litigation in Canada. In civil suits, the claims of many individuals whether in contract or tort are pursued together with members of a class numbering in the thousands. The Walkerton tainted water case is likely to proceed by way of class action. The evolution of class proceedings means that courts and

---

1 *Anderson* v. *Wilson* (1999), 175 D.L.R. (4th) 409, 44 O.R. (3rd) 673 (Ont. C.A.).

lawyers will be increasingly asked to consider statistical evidence in weighing the various aspects of cases coming before them.

This paper will review, in simple terms, i) basic tenets of statistical methodology; ii) how those principles assist in considering evidence before a court or tribunal and iii) historically, how statistical evidence has been used to impress, to befuddle, and at times, to inveigle the trier of fact.

## I. THE APPLICATION OF STATISTICAL DATA

To understand statistical evidence and its application to the litigation process, certain phrases must be understood. For those trained in statistics or epidemiology, these concepts are the most elementary of building blocks. For many lawyers, however, often trained in the social disciplines such as history or politics, statistical principles seem completely foreign. Let's start with some basic definitions.[2]

## II. ASSOCIATION

Perhaps the most simple of statistical definitions, an association exists between a condition and an agent when the condition occurs in the presence of the agent more often than would be expected by chance. For example, if a particular condition such as asbestosis occurs more frequently in coal miners than in an unexposed population,[3] there is an association between the asbestosis and coal mining.

## III. RELATIVE RISK

To determine the strength of the association, the relative risk must be addressed. Relative risk is the ratio of the incidence of occurrence in the subject population to the incidence of occurrence in the general population. The relative risk is determined by dividing the incidence in the "at risk" population by the incidence in the general population. A relative risk of 1.0 means that the incidence of the condition is the same in the subject and the general population. A relative risk of 2.0 means that a member of the subject population is twice as likely to suffer from the condition than a member of

---

2 T.A. Sheehan & C.M. Cameron, "Epidemiologic Proof of Causation" (1999) 41 For The Defence. These definitions are taken liberally from the noted article but can be found in any basic statistics text.

3 A "population" is any defined aggregate of objects, persons or events. A "sample" is any subaggregate drawn from the population.

the general population. A relative risk of less than 1.0 creates a negative association in the subject population.

## IV. P-VALUES AND STATISTICAL SIGNIFICANCE

In most cases when a scientific witness determines that the relationship between an event and an outcome is "statistically significant", all questioning stops. The finding of "statistical significance" seems to silence debate. It is important for the lawyer or trier of fact, however, to understand what is meant by "statistical significance" and, in appropriate situations, to press beyond the stated conclusion.

In the area of epidemiological or statistical research, a non-representative element may enter a study regardless of how careful the researcher has been. The researcher may have inadvertently chosen a sample population which by sheer chance has an inordinately high incidence of asbestosis despite its member (-s') never having been coal miners. This can lead to what is called a "sampling error."

In order to determine whether the observed association is the result of a random sampling error as opposed to a true result, the researcher calculates the *p-value*. If the p-value is determined at less than 5 %, it means that there is less than a 5 % chance that the observed association was due to a sampling error, or stated conversely, that there is more than a 95 % chance that the observed association represents a true condition of association within the subject population. If a specific p-value is reached, the result is said to have statistical significance.[4]

The most commonly accepted p-value in the scientific community is 5 %. This percentage has been arbitrarily selected, but through usage has come to be accepted as the minimum allowable to create a statistically significant outcome.

[4] A number of American cases (*Wade-Greaux* v. *Whitehall Laboratories Inc.*, 874 F. Supp. 1451 at 1483 (D.V.I. 1994); *Whelan* v. *Merrell-Dow Pharmaceuticals Inc*., 117 F.R.D. 299 (D.D.C. 1987); *Re Joint Eastern and Southern Asbestos Litigation*, 827 F. Supp. 1014 (S.D.N.Y. 1993), rev'd on other grounds, to cite a few) have held that epidemiological studies must show a statistically significant increase of risk in a subject population before they can be relied upon by experts to draw causal conclusions.

## V. CONFIDENCE INTERVAL

Statistics are used to establish the probability or likelihood of events being related. We are, however, often dazzled by numbers and percentages. We tend to accept the statements upon which they are based as being reliable. The mathematical approach can create a sense of reliance in the result which is false.

As noted above, the outcome of the research must be shown to be statistically significant to be persuasive. Moreover, the relative risk which we have looked at above must fall within the confidence interval.

The relative risk is based on a sample population being compared with the general population. The resulting ratio is the relative risk of being in the subject population. Because the relative risk is only an estimate of the true risk which exists within a given population, it is necessary to find a measure of reliability of the stated relative risk. How reliable is the relative risk in any given situation?

Determining such reliability is done through application of the confidence interval. Simply stated, the confidence interval provides a range within which the relative risk is said to be statistically reliable. Thus, if an estimated relative risk is 2 with the subject population being twice as likely to have the condition as the general population and we apply our p-value of 5 %, the confidence interval will range from 0.8 to 2.4. The researcher can then express a 95 % confidence rating that the real relative risk lies somewhere between 0.8 and 2.4 rather than the estimated relative risk of 2.

## VI. ATTRIBUTABLE RISK

For lawyers and judges, the most useful measure of risk in determining causation is attributable risk. Attributable risk is defined as being that proportion of the condition, such as asbestosis, in the subject population which is attributable to the exposure, such as coal mining.

The attributable risk is determined by subtracting the incidence in the general population from the incidence in the proposed population and then expressing the result as a proportion in the subject population. Thus, if the incidence in the subject population is 40 people out of 100 as opposed to the incidence in the general population at 20 people out of 100, the proportionate risk is $0.4 - 0.2 \div 0.2 = 0.5$. Otherwise stated, the attributable risk in the subject population is 50 %. There is a 50 % risk that being a member of a subject group such as coal miners will cause asbestosis.

## VII. BUT HOW DO WE MAKE ANY SENSE OF THIS?

It is axiomatic to our legal system that the standard of proof in any given case must be met to persuade the trier of fact. The civil standard is “on a balance of probabilities.” The criminal standard is “beyond a reasonable doubt.” Evidence is received and weighed against those two standards. From Bentham onwards, the law has tried to equate those two standards with a precise mathematical computation. Does “on the balance of probabilities” mean greater than 51 %? Does “beyond a reasonable doubt” mean 99.9 % certainty?

The temptation to apply mathematical or statistical weight to evidence is an old one. In 1899, the trial of Alfred Dreyfus, a Captain in the French General Staff, accused of treason, revolved around the likelihood that a handwritten letter which had fallen into German hands came from Captain Dreyfus. The letter purportedly contained statements of treason. A number of expert witnesses were called to compare individual letters within the words in the treason letter to handwriting in other letters which had come from his home. The experts did so by citing mathematical and statistical probabilities about the similarity of the handwriting—of the shape, slant, slope and size of the individual letters. After a reasonably brief trial, Captain Dreyfus was found guilty. The finding of treason against Captain Dreyfus became infamous.

A number of years later, the conviction was challenged and overturned. When the “Dreyfus Affair” was revisited, a panel of experts showed that “there was nothing statistically remarkable” about the existence of close matches in penmanship in some words in the leaked document and some in the correspondence from Dreyfus’ home despite the evidence of the earlier witnesses. What is more notable is that counsel for Captain Dreyfus and for the government Commissioner who acted as prosecutor both testified later that they had understood nothing of the original witnesses’ mathematical demonstrations linking the letters in the words in various documents. Subsequent scholars have opined that although the judges who convicted Dreyfus were in all likelihood equally mystified by the mathematical analysis, they nonetheless “allowed themselves to be impressed by the scientific phraseology of the system.”[5]

---

[5] A. Charpentier, “The Dreyfus Case 52-53” (1935) J. May trans. as quoted in L.H. Tribe, “Trial By Mathematics: Precision and Ritual in The Legal Process” (1971) 84 Harv. L. Rev. 1329 at 1333-1334.

In one of the very few academic works dealing with mathematical evidence, Professor Lawrence Tribe, a professor of Law at Harvard University, writes:

> "[…] the very mystery which surrounds mathematical arguments—the relative obscurity that makes them at once impenetrable by the layman and impressive to him—creates a continuing risk that he will give such arguments a credence they may not deserve and a weight they cannot logically claim."[6]

In the criminal setting, a seminal case in California addressed the use of statistical evidence touching on the identity of the accused.[7] A robbery had occurred. The victim, an elderly woman, testified that her assailant was a young white woman with blond hair. A neighbour testified that he had seen a caucasian woman with a dark blond ponytail run from the scene. He also reported that she had climbed into a yellow automobile driven by a black male with a beard and moustache. Several days later, a couple fitting the description were arrested but positive identification could not be made by the witnesses.

The prosecutor called evidence from a college mathematics professor to the effect that if the robbery had been committed by a white woman with a dark blond ponytail accompanied by a black man with a moustache and beard, there was an overwhelming probability that the accused couple was guilty because of the number of details in the description. The jury convicted the couple on the strength of that statistical evidence.

The California Court of Appeal later reversed the findings, holding that the mathematical testimony and the prosecutor's associated arguments were inadmissible. The Court of Appeal went on to warn:

> "[M]athematics, a veritable sorcerer in our computerized society, while assisting the trier of fact in the search for truth, must not [be allowed to] cast a spell over him."[8]

---

[6] L.H. Tribe, "Trial By Mathematics: Precision and Ritual in The Legal Process" *supra* note 5 at 1334.

[7] *People* v. *Collins*, 68 Cal. 2d 319, 438 P.2d 33; 66 Cal. Rptr. 497 (1968).

[8] "Trial By Mathematics: Precision and Ritual in The Legal Process", *supra* note 5 at 1334-1335.

## VIII. THE TRADITIONAL APPROACH

Despite the reluctance to draw specific conclusions based on statistics alone, not all statistical evidence is of concern. Indeed, without perhaps ever analysing the reason why, we are all quite comfortable with the use of statistical evidence in certain aspects of our cases. When calculating damages, we rely on life expectancy tables, on calculations of likely employment earnings, on tables reporting on the effect of a prosthesis on employability, on statistics detailing illnesses and diminution of life associated with smoking and so on. These tables detail certain probabilities and usually meet our comfort level if put into evidence by an expert who is familiar with them. We rarely insist, although entitled to do so, on calling the compiler of the data or the author of the statistical report as a witness. If the table or graph bears the StatsCan label, we tend to accept it into evidence.

We have come to accept statistical evidence being introduced to establish the probability of a future event. This may be a practical response. Without a statistical analysis of past events, there would be no useful projection of outcomes. Without statistical evidence, there would be no evidence on the point, and efforts to predict what might occur in the future would be pure conjecture.

Our instincts, however, bridle hard at statistics being used to authenticate past events, to establish liability or causation, be it through identification or through the probability of the event having occurred. The use of a statistical model to establish that an event, more likely than not, has occurred or to establish that a particular player, more likely than not, was involved in the event causes discomfort. Although the standard of proof test in a civil setting is much lower than that in a criminal context, it would be shocking to us if the statistical or mathematical evidence alone was taken as satisfying the test of causation on the balance of probabilities. In the criminal context, reliance on statistical or mathematical evidence alone would fall well short of the mark of beyond a reasonable doubt. Interestingly, though, there is no inherent distinction mathematically between past and future events, if looked at through probability concepts.[9]

---

[9] "Trial By Mathematics: Precision and Ritual in The Legal Process", *supra* note 5 at 1346.

The real question, however, is whether, in the circumstances of any particular case, it is acceptable to combine the mathematical evidence with the non- mathematical evidence in such a way as to allow the trier of fact to rely on it. Making use of statistical evidence requires transforming statistics from their generalities to the specifics or particulars of the case before us. Simply stated, statistics and epidemiology may be useful on the issue of general causation—does coal mining as an occupation increase the likelihood of asbestosis—as opposed to being used to determine actual causation—did John Smith acquire his asbestosis because he was a coal miner?

The frailty of applying statistical evidence of general causation to establish a specific finding or outcome in a case was addressed by the Ontario Court of Appeal in its decision certifying the class proceedings for patients of the encephalography clinic.[10] Although the class definition includes all patients who contracted Hepatitis B after being given an EEG at one of the clinics, the court held that causation was not a common issue for the class. In the circumstances of the case, an infected patient may or may not be able to relate their Hepatitis B positivity to having been a patient at the clinic. The Court noted that "[c]ausation is an individual issue with respect to every infected member of the class." Presumably, this means that each plaintiff must establish before the Court that an attendance at the clinic caused his or her Hepatitis B status rather than simply saying: "I was a patient at the clinic. I have Hepatitis B. Therefore, I acquired it at the clinic." The epidemiology of the Public Health Report indicates a 600-fold increase of Hepatitis B in the clinic population but that does not establish that an individual patient acquired it there.[11]

---

[10] *Anderson* v. *Wilson, supra* note 2.

[11] Despite limitations in establishing a causal relationship between having had an EEG at one of the Clinics and being Hepatitis B positive, Public Health drew some notable conclusions in their study. They reported that the rate of Hepatitis B at the EEG Clinics far exceeded that in the general population. The rate among all Clinic attendees in 1995 was 1,733 per 100,000 population. This was reported as being 600 times more than the annual rate of known Hepatitis B in Ontario in 1995 which was 2.9 per 100,000. This conclusion however fails to note that there has never been a similar study for Hepatitis B positive individuals within the general population in Ontario. We therefore have no way of knowing the true prevalence rate within the general population of Ontario. If the denominator is unknown, it is doubtful that the purported multiplier can have any veracity.

## IX. RELATED CONCEPTS AND LEGAL THEORIES

***Similar Fact Evidence***

The notion of many similar cases being more persuasive than one case is not new to our courts. Epidemiological evidence is like similar fact evidence on a very grand scale. If the comparisons are taken further, we can expect that the courts will be more likely to give weight to epidemiological evidence in civil cases with continued limitations on the applicability of similar fact evidence in criminal matters.

The admissibility of similar fact evidence in civil cases was considered by Lord Denning in *Mood Music Publishing Co*. v. *De Wolfe Ltd*.[12] The case involved allegations of infringement of copyright. Lord Denning outlined the test of admissibility of the similar fact evidence in civil cases at p. 766:

> "The criminal courts have been very careful not to admit [similar fact] evidence unless its probative value is so strong that it should be received in the interests of justice: and its admission will not operate unfairly to the accused. In civil cases, the courts have followed a similar line but have not been so chary of admitting it. In civil cases, the courts will admit evidence of similar facts if it is logically probative, that is, if it is logically relevant in determining the matter which is in issue; provided that it is not oppressive or unfair to the other side; and also that the other side has fair notice of it and is able to deal with it."

Thus, if statistical or epidemiologic evidence is logically probative and relevant, it is likely to be similarly admitted in civil proceedings unless it is oppressive or unfair.

The courts are likely to hold that epidemiologic evidence meeting that standard can be adequately tested through cross-examination.

Recently, however, in *R*. v. *Arp*, the Supreme Court of Canada confirmed the Courts' reluctance to see similar fact evidence admitted in a criminal setting. The court instructed any trial judge faced with a similar fact issue to evaluate carefully the degrees of similarity being advanced and to determine whether the objective improbability of coincidence was established. In assessing the similarity of certain acts, the trial judge should consider only the manner in which the acts were committed and not evidence

---

[12] [1976] All E.R. 763 (C.A.).

as to the accused's involvement in each act. The Court cautioned that similar fact evidence must be linked to the accused by evidence disclosing more than the mere possibility that the accused committed the alleged similar act.[13] This approach seems to confirm a hesitation to draw conclusions from the general to the specific, especially on issues of identity or causation.

***Res ipsa loquitur***

At first glance, the topics of statistical evidence and *res ipsa loquitur* may seem far removed from each other. The link, however, is discernible.

We have reviewed above the role of statistical evidence establishing findings of general causation. Statistical analysis by definition will convey general findings. Arguably, unless the defence can find a reason to distinguish these generalizations, the statistics combined with relevant non-statistical evidence can persuade the Court. The introduction of statistical evidence is aimed at relieving the plaintiff or Crown from proving the specifics of his case, relying instead on the generalities that a statistical analysis of the facts can bring. The burden of proof perceptibly shifts to the opposing party in the face of strong statistical evidence. Similarly, the old legal maxim of *res ipsa loquitur* shifts the burden of proof from the plaintiff. In the face of strong epidemiologic evidence or a claim of *res ipsa loquitur*, the defendant may face the burden to establishing why the obvious is not so obvious.

It is therefore paradoxical that as complex matters before the Courts are increasingly being tried on evidence of scientific certainty with inferences of mathematical precision and statistical analysis, the Supreme Court of Canada[14] has ruled that the maxim of *res ipsa loquitur* should be put to rest once and for all.

In the *Fontaine* decision, the Court was asked to consider that negligence had been established against the driver of a truck which was found badly damaged in a swollen creek bed beside a mountain highway. The plaintiff, the widow of the passenger in the truck, claimed that the truck could only have rolled off the road into the creek if the driver had been negligent. She claimed against the insurer in British Columbia. The Court however was unpersuaded and noted that severe weather conditions create situations where accidents may occur and vehicles may leave the road regardless of the care shown.

---

13 *R.* v. *Arp*, [1998] 3 S.C.R. 339 (S.C.C.).

14 *Fontaine* v. *Insurance Corporation of British Columbia*, [1998] 1 S.C.R. 424.

Beyond its own facts, the case marks a critical juncture in the evolution of tort law. It effectively expunges the maxim of *res ipsa loquitur*. Mr. Justice Major for the Court began the decision by quoting the definition of *res ipsa loquitur*:

> "The doctrine applies (1) when the thing that inflicted the damage was under the sole management and control of the defendant, or of someone for whom he is responsible or whom he has the right to control; (2) the occurrence is such that it would not have happened without negligence. If these two conditions are satisfied, it follows, on a balance of probability, that the defendant, or the person for whom he is responsible, must have been negligent. There is however a further negative condition: (3) there must be no evidence as to why or how the occurrence took place. If there is, then appeal to *res ipsa loquitur* is inappropriate, for the question of the defendant's negligence must be determined on that evidence."[15]

After analyzing the degree to which a finding of *res ipsa loquitur* is dependent on the facts of any given case, Mr Justice Major affirmed that the burden of proof remains on the plaintiff. If the defendant produces a reasonable explanation that is as consistent with no negligence as the *res ipsa loquitur* inference was consistent with negligence, the plaintiffs' case will fail.

Mr Justice Major then went on to deliver the death knell to the doctrine:

> "Whatever value *res ipsa loquitur* may have once provided is gone. Various attempts to apply the so-called doctrine have been more confusing than helpful. Its use has been restricted to cases where the facts permitted an inference of negligence and there was no other reasonable explanation for the accident. Given its limited use, it is somewhat meaningless to refer to that use as a doctrine of law.
>
> It would appear that the law would be better served if the maxim was treated as expired and no longer used as a separate component in negligence actions. After all, it is nothing more than an attempt to deal with circumstantial evidence. That evidence is more sensibly dealt with by the trier of fact, who should weigh the circumstantial evidence with the direct evidence, if any, to determine whether the

---

15 J.F. Clerk & W.H.B. Lindsell, *Clerk & Lindsell on Torts*, 13th ed. (London: Sweet & Maxwell, 1969, at para. 967 quoted with approval in *Fontaine*, *supra* note 14.

> plaintiff has established on the balance of probabilities a *prima facie* case of negligence against the defendant. Once the plaintiff has done so, the defendant must present evidence negating that of the plaintiff or necessarily, the plaintiff will succeed."[16]

In light of the language of the Supreme Court in *Fontaine*, it is the role of the trier of fact to determine whether strong statistical evidence, which is circumstantial, is capable of shifting the burden from the plaintiff to the defence. Courts must address whether strong statistical evidence of general causation shifts the burden to the defence to address specific causation matters or whether the burden still rests on the plaintiff. This will be a case-by-case decision.

## CONCLUSION

American jurisprudence has taken great measures in the past ten years to expose trial decisions based on "junk science." For our judicial system, lawyers and judges must be aware of cases based heavily on statistical evidence.

Statistical analysis has the shiny allure of objectivity and precision which makes it difficult to attack or negate. We have been schooled to believe that numbers are more reliable than narrative; that the mathematical is more reliable than the anecdotal. We are convinced that evidence is more persuasive and impartial if reduced to calculations and statistics. Yet, mathematical formulae and models are variable. Rational men do differ on statistical methodologies and depending upon the assumption made and the questions asked, statistics can lie. Disraeli was not wrong!

We must therefore be vigilant. When presented with a case which is founded on epidemiologic or statistical evidence, we must analyze the evidence carefully. While evidence of a statistically significant attributable risk may sound persuasive, the conclusion must be unpackaged and examined for that upon which it is really based. Not every statistic must be proven anew but a healthy dose of skepticism must be applied to the use to which the statistic is being put. Once in evidence, vigilance must be maintained about the mystery of mathematical arguments; about the allure and yet, the ambiguity that makes those arguments often impenetrable and impressive.[17] In particular, lawyers and judges must not appear to understand

---

16 *Fontaine, supra* note 14 at 435.

17 "Trial By Mathematics: Precision and Ritual in The Legal Process", *supra* note 5 at 1334.

statistical evidence if indeed they do not. Like all evidence, statistics and epidemiology must be tested and scrutinized. The challenge will be ours to keep mathematics, the veritable sorcerer of our time,[18] in its proper place.

[18] See *supra* note 7.