

TOXIC TORTS AND CAUSATION: TOWARDS AN EQUITABLE SOLUTION IN AUSTRALIAN LAW

PART II: MEANS-ENDS ANALYSIS[†]

PAOLO F RICCI* AND NATALIE J GRAY**

I. INTRODUCTION

In toxic tort and environmental law, scientific evidence bears on causation through emission-exposure and dose-response models. The latter consists of clinical studies, animal *in vivo* studies, epidemiological studies, and tests on lower organisms. Scientific and legal causation merge in an uncertain network of feedbacks: variable inputs and outputs surrounded by different degrees of knowledge. It is doubtful that a trier of fact could adequately decide between conflicting forms of evidence on 'common sense' and 'ordinary experience'. However, otherwise meritorious claims cannot fail because the causal chain is not within ordinary experience.¹ Thus, a coherent and uniform treatment of uncertain evidence, scientific theories and variable data in toxic tort cases is timely in Australia.

Symmetry of information, discussed in Part I of this paper, is a way to restore the traditional 'economically efficient' solutions, such as Pareto optimality, that do not hold without it.² Although non-cooperative dynamic game theory³ can help when there is an asymmetry of information, legal fairness demands that the

† Part I of this article appeared in Volume 21(3) of the *University of New South Wales Law Journal*.

* University of California, Berkeley, USA; PhD, LL.M, MPA, MA, MSc. Ricci & Molton, 685 Hilldale Ave, Berkeley, CA 94708, USA; Professor, Faculty of Law, University of Wollongong, Australia.

** BSc, LLB (Hons) (Woll).

1 The injustice of requiring a deterministic description of the disease process which science has not yet been able to deliver is discussed by JB Brennan and A Carter, "Legal and Scientific Probability of Causation of Cancer and other Environmental Diseases in Individuals" (1985) 10 *Journal of Health, Politics, Policy and Law* 33.

2 B Holmstrom and R Meyerson, "Efficient and Durable Decision Rules with Incomplete Information" (1983) 51 *Econometrica* 1799.

3 See, generally, A Roth (ed), *Game Theoretic Models of Bargaining*, Cambridge University Press (1985).

parties and the decision maker have access to the information and treat it according to well demonstrated methods for the analysis of causation. Symmetry of information fits with both economic efficiency and with liability rules formulated in terms of 'durability', rather than efficiency. 'Durability' means that the parties would not unanimously reject a decision if they took a vote after pooling all of their private (undisclosed) information. Such analysis needs an informal framework for uncertain cause-in-fact and proximate cause with:

- numerical representations of vague ('possible'), uncertain ('probable') or probabilistic statements;
- a list of symbols (eg, 'for some');
- axioms and postulates (eg, 'due process');
- rules for combining symbols (eg, 'If ... , If ..., If, ... Then ...');
- rules for inference;
- data protocols;
- a contingent process for overall evaluation by all parties; and
- a jurisdictionally appropriate social calculus (eg, 'risk-cost-benefit' balancing).

The next sections provide some substance to this framework, exemplifying its potential for judicial reasoning in toxic tort and environmental law. The focus is on human health. Extending the framework we propose to other scientific and technical areas does not present difficulties and is not discussed.

A. Context

In toxic torts, neither scientific nor legal causation requires, nor can hope for, certainty. An American case, *Allen et al v US*⁴ illustrates how a court deals with scientific uncertainty. The plaintiffs sued the US Government, attempting to recover for leukemia allegedly caused by fallout from testing nuclear devices. The court discussed "natural and probable" consequences, the "substantial factor", "but for" and "preponderance of the evidence" tests; and opted for the "substantial factor" test.

However, the court established that each plaintiff should show three things with the preponderance of the evidence test. First, that the probability of exposure to fallout results in doses in significant excess of 'background'. Secondly, that the injury is consistent with those known to be caused by ionising radiation. Thirdly, that individual claimants resided near the Nevada Test Site for at least some of the years when the tests were conducted. The court also found that lack of statistical significance does not necessarily mean that the adverse effect is not present.⁵

4 *Allen et al v US* 588 F Supp 247 (1984); reversed on other grounds, 816 F 2d 1417 (10th Cir 1988); certiorari denied, 484 US 1004 (1988).

5 *Ibid* at 416-17.

Although the prevailing standard of proof in tort law is the preponderance of the evidence, the admissibility of epidemiological and other 'medical' evidence is governed by the "reasonable medical probability" standard.⁶ If this standard is higher than the preponderance of the evidence, its stringency affects the plaintiff more so than the defendant, and vice versa.

When the scientific basis of causation (eg, that a particular dose-response model was biologically more plausible than another) is construed as a 'possibility', it is insufficient to demonstrate legal causation by the 'preponderance of the evidence'.⁷ These difficulties, and the often conjectural model of dose-response, increasingly force the legal system to ask science for a 'proof' of causation that allows for the symmetric treatment of uncertain and heterogeneous scientific information.⁸

Such a demand is correct. The means are available closely to approximate an accurate assessment of the uncertainties. Modern formal causal probabilistic reasoning can result in a more just resolution of tortious and environmental disputes. The end is the just apportionment of tortious liability.

There are, however, some difficulties. The judicial process requires reasoning with uncertain data and incomplete models. Statistical methods are used to extrapolate to an area far removed from the 'relevant range' of the observations. The second difficulty is the need to reconcile legal reasoning about causation with probability weighted factual patterns. The difficult problem of whether any interpretation of probability number (such as logical or personal-subjective) can truly guide behaviour is unresolved. Fortunately, we do not need to resolve these problems because the frame of reference that always improves the state of knowledge is preferable as a matter of legal fairness.

B. Probabilities

Causation and probabilities coexist. A physical regularity becomes causal when it can be identified and assessed. The total process may include a number of regularities or 'laws'. Is the regularity (the 'law' mathematically described through, say, differential equations) that determines the uncertainties, with statistical analysis merely providing reliable numbers for the coefficients of the differential equation? Or are probabilities fundamental measures inherently and inexorably part of that and any physical and behavioural law?

Briefly, there are different views of probabilities. They have been understood as either subjective measures of belief or objective as the result of the long-term replication of the same physical experiment. Epistemic probabilities describe ignorance about a known deterministic process. Frequencies are justified as probabilities by the indefinite replications of the same generating process. They

6 *Cottle v Superior Court* 3 Cal App 4th 1367 (1992) at 1384-5.

7 See *Green v American Tobacco Co* 391 F 2d 97 (5th Cir 1968); rehearing (en banc), 409 F 2d 1166 (5th Cir 1969). *Bowman v Twin Falls* 581 P 2d 770 (Id 1978) at 774 held that: "to require certainty when causation itself is defined in terms of statistical probability is to ask for too much".

8 H Harris, "Toxic Tort Litigation and the Cause Element: Is There Any Hope of Reconciliation?" (1986) 40 *Southwestern Law Journal* 909.

are the empirical probabilities. Quantum probabilities represent irreducible natural randomness.⁹

C. Some Reasons for These Differences

The imperfect knowledge of the true mechanical state of a system led to epistemic probabilities. The probability attached to the representation of the system, in Maxwell's view, is the a priori "measure of the degree of ignorance of the true physical state of the system".¹⁰ Boltzmann's probability represents an average state of the system, during a specific interval of time; Maxwell's results from considering a very large number of equal systems, each with different initial states.¹¹

For Einstein, in 1909 and later, probabilities were determined by the limit of relative frequencies: they were objective measures of the state of a system.¹² Probabilities, in his view, are time-average measures of the evolution of the physical process. Einstein also stated the principle that: "only the theory decides what can be observed".¹³ In 1922, Schrödinger believed that physical processes are statistical: deterministic causation is merely commonplace thinking. He found determinism "prejudicial".¹⁴ Reichenbach, sometime later, developed the nexus between the continuous causal evolution of the system and its probabilistic interpretation. He held the view that probabilities are relative frequencies.¹⁵

Borel's probability measures were defined by convention, not essential properties.¹⁶ A probability is a degree of belief: the individual's choice of a bet that makes it subjective when the probability values are sufficiently away from 0 or 1. He explains the concept of "practical certainty" as an event characterised by a sufficiently large probability number, and vice versa.

Probability, as an inherent part of physical laws, through 'becoming' (rather than 'rigidly existing') is a physical property that is probabilistic.¹⁷ Gibbs' work through the random drawing from a large number of items that are, at least in principle, fully described was based on epistemic probabilities.¹⁸

For von Mises, probabilities apply to homogeneous events that are characterised by large number of repetitions.¹⁹ He rejected the idea of subjective probabilities, preferring the random choice of elements out of the "collective" of those elements.²⁰ If these collective related conditions for determining a probability number do not obtain, then there is no probability.

9 J von Plato, *Creating Modern Probability*, Cambridge University Press (1994) p 167.

10 *Ibid*, p 75 (footnote and emphasis omitted).

11 *Ibid*, p 76 (footnote omitted).

12 A Einstein, "Zum Gegenwärtigen Stand des Strahlungsproblems" (1909) 10 *Physikalische Zeitschrift* 185.

13 Von Plato, note 9 *supra*, p 158.

14 *Ibid*, p 150.

15 H Reichenbach, "Stetige Wahrscheinlichkeitsrechnung" (1929) 53 *Zeitschrift für Physik* 274.

16 Von Plato, note 9 *supra*, p 43.

17 H Weyl, *Philosophie der Mathematik und Naturwissenschaft*, Oldenbourg (1927).

18 JW Gibbs, *Elementary Principles of Statistical Mechanics*, Dover (1960) p 17.

19 R von Mises, *Mathematical Theory of Probability and Statistics*, Academic Press (1964).

20 *Ibid*, pp 183-97.

Kolmogorov developed a set-theoretic interpretation of probability in the 1920s²¹ in contrast to von Mises'. According to this interpretation, a probability admits the concept of dependence: the practical "causal connection".²² In the 1960s, in a reversal, Kolmogorov's probabilities used events with long sequences of repetitions. A probability number is a 'constant' with objectively determined numerical characteristics. It has both probabilistic (the fluctuation about the constant value) and frequentistic (law of large numbers) aspects. Kolmogorov believed that empirical studies determined the nature of the probability laws (that is, the *distribution* of elementary events) inductively.

Should, then, probabilities be related to imperfect measurement, considering that the laws of nature are unchangeable, or not? The former vision is deterministic and epistemic. De Finetti held to the contrary: the laws of nature are statistical regularities, thus requiring the "sharing" of properties between the two polar views. Probabilities are primitive qualitative characterisations of human behaviour that can be numerically represented. De Finetti's concept of *coherence*, which requires that a rational person would not engage in bets whereby he or she would surely lose, is most compelling.²³

De Finetti demonstrated that coherence means that any probability measure associated with that bet satisfies the additivity axiom for finite events. This is a critical aspect of decision making under uncertainty. De Finetti's work transcends epistemic and other probabilities; natural processes are indeterministic with subjective but coherent probability numbers characterising their outcomes. These probabilities are not influenced by the reasons for ignorance, they are independent of deterministic or indeterministic assumptions about any process. Subjective probabilities quantify degrees of belief through coherence. Using axioms, he also developed the qualitative basis of probability: the calculus of probability is the result of a primitive statements such as 'not less probable than' linking common expressions to a formal system of analysis.²⁴ He stated:

For ... an objective meaning to probability, the calculus ... ought to have an objective meaning, ... its theorems ... express properties that are satisfied in reality. But it is useless to make such hypotheses. It suffices to limit oneself to the subjectivist conception, and to consider probability as a degree of belief a given individual has in the occurrence of a given event. Then one can show that the ... theorems of the calculus of probability are necessary and sufficient conditions²⁵ for the given person's opinion not to be intrinsically contradictory and incoherent.

21 A Kolmogorov, "Logical Basis for Information Theory and Probability Theory" (1968) 14 *Institute of Electrical and Electronic Engineers: Transactions of Information Technology* 662.

22 Von Plato, note 20 *supra*, p 205.

23 *Ibid*, p 272.

24 *Ibid* at 273-6.

25 B de Finetti, "Fondamenti Logici del Ragionamento Probabilistico" (1930) 5 *Bollettino Unione Matematica Italiana* 258 at 260.

(i) Remark

Probabilities (numbers between 0 and 1 including these two limits) are not 'cardinal'. It cannot be said that the probability of 1 is twice as large as the probability of 0.5. If the probability of response in a year t is p , then the probability scale can be transformed to yield a cardinal measure that permits comparisons such as 'twice as large'. The transformation is $r(\text{probability number}) = \{-[\text{natural logarithm}(1-\text{probability number})]\}$. At low probability numbers, the transformation is unnecessary: $r(pr)$ is approximately equal to the probability number (pr) itself.

II. REJECTION OF PROBABILISTIC REASONING AND IMPLICATIONS FOR TOXIC TORT LAW

Causation in toxic tort law is synthesised as the process characterised by heterogeneous and uncertain events:

[source(s) of concern; mass, magnitude, probability, time, type] => [transport, fate; concentration, time, probability] => [exposure; type, duration, context, probability] => [dose; type, duration, context, probability] => [response; cancer, mutagenicity, teratogenicity, toxicity, probability] => [acceptability of risks; individual, population, type, magnitude, probability].

The requisite, before judicially limiting a toxic torts causal network, is the formal concatenation of the essential elements, accounting for uncertainty.

Australian courts have been reluctant to adopt probabilistic reasoning in determining causation, particularly in negligence cases, preferring instead to allow the trier of fact to temper the inadequacies of the 'but for' test with the application of 'common sense'.²⁶ The basis of this rejection may be found in the need of the courts to ensure that causation is grounded in realism and experience. In the words of Dixon J, in *Briginshaw v Briginshaw & Anor*:

The truth is that, when the law requires the proof of any fact, the tribunal must feel an *actual persuasion* of its occurrence or its existence before it can be found. *It*

26 A leading case establishing the common sense test is *Bennett v Minister of Community Welfare* (1992) 176 CLR 408 at 412, per Mason CJ, Deane and Toohey JJ: "In the realm of negligence, causation is essentially a question of fact, to be resolved as a matter of common sense. In resolving that questions, the 'but for' test, applied as a negative criterion of causation, has an important role to play but is not a comprehensive and exclusive test of causation; value judgments and policy considerations necessarily intrude". This statement stems from the majority judgment in *March v E & MH Stramare* (1991) 171 CLR 506, and has since been approved in *Steele v Twin City Rigging Pty Ltd* (1993) 114 FLR 99 at 109, and *Medlin v State Government Insurance Commission* (1995) 182 CLR 1 at 6. An argument for probabilistic reasoning was, however, made by Murphy J in *TNT Management Pty Ltd v Brooks* (1979) 23 ALR 345, which was discussed in detail in Part I of this article.

cannot be found as a result of a mere mechanical comparison of probabilities independently of any belief in reality.²⁷

This is naive in light of the complexity of causal patterns, their probabilistic nature and our discussion of probability numbers. Reliance on “actual persuasion” is inappropriate and leads to unjust results. The belief of causation “in reality” confuses legal causation with scientific and statistical causation. The lawyer must demonstrate that one causal pattern, for a set of given circumstances, is more likely than not. Legal causation is one probable outcome out of many; it is the single causal path sufficient to resolve the dispute.²⁸ The scientist searches for verifiable results, given the state of knowledge,²⁹ but faces numerous alternatives, each of varying probability and refutation.

As the legal system attempts to determine factual cause, what appears to be an objective search for truth is corrupted by normative judgments about facts.³⁰ Consider factual causation stopped, somewhere in its logical continuum, by proximate cause,³¹ the judge-made idea of justice, legal and political policy. This most subjective criterion can determine legal liability; it is evident in the Australian ‘common sense’ test.

‘Proximate cause’ and ‘cause-in-fact’ form legal causation in tort law. This distinction is more pronounced in negligence than in product liability law. Cause-in-fact is the factual chain of events leading to ultimate injury.³² Recollect the dissent by Andrews J in *Palsgraf v Long Island RR*,³³ where he stated:

What we do mean by the word ‘proximate’ is that, because of convenience, of public policy, of a rough sense of justice, the law arbitrarily declines to trace a series of events beyond a certain point. It is all a question of expediency.³⁴

It is based on: “[t]he foresight of which the court speaks assumes prevision.”³⁵ We think that expediency is an excuse, not a reason.

27 (1938) 60 CLR 336 at 361 (emphasis added). This was followed in the mesothelioma case of *Wintle v Conaust* [1989] VR 951 at 953. This statement may also be interpreted as raising the required standard of proof (on the balance of probabilities) above 51 per cent.

28 I Macduff, “Causation, Theory and Uncertainty” (1978) 9 *Victoria University of Wellington Law Review* 87 at 93.

29 E Adeney, “The Challenge of Medical Uncertainty: Factual Causation in Anglo-Australian Toxic Tort Litigation” (1993) 19 *Monash University Law Review* 23 at 24.

30 J Cohen, “The Value of Value Symbols in Law” in Smith and Weisstib (eds), *The Western Idea of Law*, Butterworths (1983) 1 at 8.

31 See *Chapman v Hearse* (1961) 106 CLR 112 at 122.

32 The California Supreme Court “disapproved” the use of “proximate cause” in favour of the “substantial factor”. See *Mitchell v Gonzales* 54 Cal 3d 1041 (1991), cited in *Pamela Lee v Heydon* (1994) CDOS 4265 at 4265-6. In the *Restatement of Torts*, II § 431 ‘substantial’ relates to the defendant’s conduct to the extent it results in harm which “reasonable men ... regard ... as cause, using the word in the popular sense”.

33 162 NE 99 (NY 1928).

34 *Ibid* at 103.

35 *Ibid* at 105.

The more removed or unusual the occurrence of factors, the more likely it is that the causal chain is either conjectural or not foreseeable. What is normal is foreseeable.³⁶ But, what is foreseeable is in the eye of the beholder. The California Supreme Court stated:

Experience has shown that ... there are clear judicial days on which [a jury] can foresee forever and thus determine [causation,] but none on which that foresight alone provides a socially and judicially acceptable limit on recovery of damages for injury.³⁷

This, we suggest, holds for all fact finding in toxic torts because critical facts often can be remote from each other, separated perhaps by years of latency.³⁸

The resolution of toxic tort disputes seems to: (a) confuse deterministic with probabilistic causation; and (b) use tests couched in probabilistic language and yet resist mathematical probabilistic balancing. There is a pernicious reduction of complex causal paths to a minimal path resolved by common sense arguing that probabilistic methods do not adequately link the individual cause (or defendant) with the individual effect (or plaintiff).³⁹ It is analytical escapism. The result is unjust for the defendant, who faces unpredictable heuristics and bizarre scientific theories; and unjust for the plaintiff, who may have no recourse for an arbitrary finding of lack of causation.

A. Means-Ends Analysis: Symmetric Treatment of Causal Events and Uncertainty

A just process requires a symmetric treatment of causal facts between plaintiff and defendant. The parties to the dispute, the fact finder and the legal decision maker would then have equal access to the information and its processing. Naturally, the parties will assign different weights to the evidence and the factual links. But consistency and coherence is guaranteed using probabilistic measures and by analytically combining evidence. Errors would be discovered in the judicial proceedings; symmetry is unaffected.⁴⁰

This symmetry would fit well with evidence expressed in terms of health numbers. Take, for example, a mathematical representation of risk: the relative risk, which indicates the ratio of the incidence rate of disease in a group exposed

36 Foreseeability concerns the determination of duty. See *Roland v Christian* 69 Cal 2d 108 (1968).

37 *Thing v La Chusa* 48 Cal 3d 644 (1989) at 668.

38 Law and economics theorists argue that legal causation should be replaced by 'social efficiency' via a form of Judge Learned Hand's test, *US v Carroll Towing* 159 F 2d 169 (2nd Cir 1947). Because this test yields an expected value, probabilistic causation remains.

39 Adeney, note 29 *supra* at 59.

40 *Hyatt v Sierra Boat Co* 79 Cal App 3d 325 (1981) at 337-9, rejecting expert's testimony not being reasonably with the field of expertise; *Pacific Gas and Electric Co v Zuckerman* 189 Cal App 3d 1113 (1987) at 1135: "the value of opinion evidence rests not only in the conclusion reached but in the factors considered and reasoning employed". In accord: *De Luca v Merrell Dow Pharmaceutical Inc* 791 F Supp 1042 (DNJ 1992) at 1047.

to a toxic chemical to the incident rate of disease in a non-exposed group.⁴¹ Where relative risk is high, legal and scientific causation can be established with little difficulty. But when the relative risk is close to 1.0 and is statistically significant, the results for the plaintiff can be whimsical or unjust, or both. This is the result of judicially demanding a legal balancing based on the 51 per cent rule. The relative risk that the plaintiff must demonstrate is much larger than the toxic tort related relative risk, which is often about 1.2. The US Ninth Circuit Court of Appeals has held that for "an epidemiological study to show causation under a preponderance of the evidence standard, the relative risk ... will, at the minimum, have to exceed '2'".⁴² However, some state courts have recognised that a lower relative risk number can be acceptable under suitable circumstances.⁴³

III. FACTUAL CAUSATION: EXPOSURE AND RESPONSE

Uncertainty may be analytical and conjectural, where formal analyses can be applied, and be 'transcientific', where scientific constructs are put forth as plausible hypotheses but cannot be answered by science. This raises two considerations that affect the law of toxic torts:

- 1 Scientific conjectures arise when causal or explanatory models are unknown or incomplete and the supporting data are either missing, imprecise, or even contradictory.
- 2 Mathematical objects (such as probabilities) describe uncertainty, understood as the combination of the *variability* of the data and the *specification* of the model. The former results from natural sampling variability. The latter refers to the choice of mathematical form (linear or non-linear) of the dose-response function and to the inclusion of relevant independent variables.

Ad hoc qualitative choices, which limit the set of possible events to be considered, reduce the combinations of inputs and methods achieving manageable problem statements.⁴⁴ This creates the danger that consensus on the elements and events to be included in the choices will merge the objectives of the analysis with its cost, and can be influenced by who pays for mitigating risks. The conclusions contain the premises: an inductive fallacy.

41 EK Christie, "Toxic Tort Disputes: Proof of Causation and the Courts" (1992) 9 *Environmental and Planning Law Journal* 302 at 310; PF Ricci and LS Molton, "Risk and Benefits in Environmental Law" (1981) 214 *Science* 1096.

42 *Daubert v Merrell Dow Pharmaceuticals Inc* 43 F 3d 1311 (9th Cir 1995) at 1321, (citations omitted).

43 *Hall v Baxter Healthcare Corp* 947 F Supp 1387 (D Oreg 1996) at 1403.

44 The problem is that the number of options may reach into the thousands.

A. Context

The prototypic exposure relationship (a subset of the fuller emission-exposure-response paradigm) in toxic tort law is:

[site (or component); emissions; transport and fate; uptake] \Rightarrow [exposure time series of concentrations].

Each bracketed element includes probabilistic inputs, models and outputs. The observed output (eg, ozone concentrations measured in Sydney) is one of many possible samples generated by a complex physico-chemical process. The US Environmental Protection Agency (EPA) has used, for environmental risk calculations, the Reasonable Maximum Exposure measure. It indicates the highest exposure that could reasonably be expected to occur for a given exposure pathway:⁴⁵ the upper 95 per cent confidence limit on the normally distributed concentrations.⁴⁶ The upper 95 per cent confidence limit on the normal (or log-normal) distribution is also used for the duration and frequency of exposure. If the data are insufficient for statistical analysis, then the highest modelled or measured concentration is used.⁴⁷

Currently, the guidelines issued by the US EPA call for the distribution of exposure, achieved by propagating the uncertainty using probability distributions associated with each input and output in the chain, often through Monte Carlo simulation.⁴⁸ This partially causal chain includes emission, transport and uptake across several components of the exposure assessment.

(i) Probability of Response

The risk to an individual of developing an adverse health response, such as cancer, is defined as the probability that the individual will develop the disease, in a period of time, given survival until then. Individual risks alone do not show the impact of exposure on the population at risk. The full representation requires considering the population (aggregate) risk and the distribution of risk over those affected. This is provided by the distribution of 'remaining life-years in the population'.

B. Cancer as an Example of a Causal Network Leading to Probabilities

The biological processes leading to cancer yield an accurate mathematical description of the cancer process through a dose-response model. The simplest is the 'one-hit' function: it describes how a single chemical interaction between the chemical and the DNA results in a probability of cancer. The multistage

45 US EPA, *Risk Assessment Guidance for Superfund, Vol 1 Human Health Evaluation Manual (Part A)*, EPA/540/1-89/002 (December 1989).

46 US EPA, *Supplemental Guidance to RAGS: Calculating the Concentration Term*, PB92 - 963373 (May 1992) at 2.

47 *Ibid* at 3.

48 US EPA, *Guidelines for Exposure Assessment*, 57 FR 22888-938 (1992).

model is biologically more realistic because it describes the number of stages⁴⁹ through which an affected cell line must pass, and the number of sequential hits it must suffer, without repair, before it becomes tumorigenic. The results of a choice between one model and another can be astounding: differences in the dose, from the same level of risk, vary by several orders of magnitude.

A simple view of the cancer process consists of initiation (when an irreversible lesion of the DNA occurs), promotion (the biochemical process accelerating the tumorigenic process), and progression (describing the now precancerous cell's progression toward malignancy). If a carcinogen does not affect the DNA directly, it is likely that the particular dose-response function has a threshold: a level below which exposure does not trigger cancer. More biologically plausible cancer dose-response models, which describe the interaction of a chemical with a cellular target, and the birth and death processes of cells, are being advanced and used.

These considerations raise some statistical issues. The first, relevant to causation regardless of the type of model, is the propriety of extrapolations outside the relevant range of the data to very low exposure, or doses. These are the doses that the toxic tort plaintiff encounters. Even with the same data set, different forms of the multistage model yield different low dose relationships, but very similar results in the relevant range of the data. The second is that the natural pharmaco-kinetic biological processes eliminates some of the mass of the original chemical to which the plaintiff has been exposed, thus reducing the mass of the original chemical that reaches the DNA.⁵⁰ Thirdly, these models follow the assumption that cells at risk are transformed independently of one another. This can be questionable because evidence suggests that there is loss of intercellular control.

This discussion suggests two things. First, there are competing theories resulting in ad hoc reasoning to simplify complex scientific matters for regulatory and tort law. Secondly, there is a logical discontinuity between scientific evidence adduced to approximate causation, accounting for incomplete knowledge and variable data. Following Richard Jeffrey, the appropriate legal view of causation must assign a probability number conditioned on the available evidence either to accept or to reject that causation.⁵¹ This contrasts with empirical philosophers such as Churchman, Braithwaite and Rudner who believe that, as Richard Jeffrey summarises, "ethical judgments are essentially involved in decisions as to which hypothesis should be included in the body of

49 The actual definition of a 'stage' in cancer process is difficult, "[a] rough general rule is that if a change is not likely to have happened within 10 years of a cell being ready for it, then it would count as a stage, but if it is likely to take less than a year it would not": M Kiah, JD Watson and H Winsten (eds), *Origins of Human Cancer*, vol 4 Cold Spring Harbor Laboratory NY (1977) p 1403.

50 One way for overcoming this problem is to use physiologically based pharmaco-kinetic (PB-PK) models. These yield the concentrations of the ultimate by-products of biochemical reactions, from the original chemical, to the target tissue, cell or DNA. This is the dose, often measured in milligrams per kilogram of body weight per day.

51 R Jeffrey, *Probability and the Art of Judgement*, Cambridge University Press (1992). See, in particular: ch 2, p 15.

scientifically accepted propositions".⁵² In judicial proceedings about scientific causation, the opposite should be true. An example is the use of probability values⁵³ to justify the acceptance or rejection of a statistical finding. There is judgment about a final choice, but it is factual and based on probabilistic reasoning that can be refuted scientifically.

(i) *Animal Studies*

In vivo animal studies are used to study disease⁵⁴ for many different reasons. Apart from the practical consideration that animal bioassays can provide faster results than epidemiological studies, their primary advantage is that they are well controlled, unlike epidemiologic studies. However, there are a number of shortcomings in these studies. Animals in experimental studies are often exposed through routes, eg gavage, which differ from human exposure. The biochemical and physiological make-up of experimental animals can be different, requiring interspecies conversions. The assumption of intraspecies homogeneity is questionable because of genetic differences. The exposure among animals in the same dose group may vary, animals may gain and lose weight at different rates and undetected infections may occur. There also is increasing concern with oncogenes (cancer causing genes) in some commonly used animal strains.⁵⁵

Although animal studies are critical in dealing with causation, their use in environmental law can result in questionable practices. One is related to the conjectural aspects of dose-response models outside the relevant range of the data which has led to averaging different animal results. For example, in the early stages of regulating benzene, a known *leukaemogen*, animal data were used with different dose-response models. The human risk estimates were developed from the geometric mean of the slopes from these models as estimated from CFW mice, who developed forestomach, larynx and oesophagus squamous cell papillomas and carcinomas;⁵⁶ SWR/J mice; and Sprague-Dawley male and female rats. To the extent that humans do not have forestomachs, it is unclear how that information is relevant to regulating human leukemic risks, considering that the animal cancers are solid cancer, and the leukemias are not.

Some of these uncertainties have lead the US EPA to associate letter "weights of evidence" with each cancer potency factor used in regulatory risk

52 *Ibid*, p 14 (footnote omitted). Jeffrey cites Rudner's view that: "for, since no scientific hypothesis is ever completely verified, in accepting a hypothesis the scientist must make the decision that the evidence is sufficiently strong or that the probability is sufficiently high to warrant the acceptance of the hypothesis ... [which] is going to be a function of the importance, in the typical ethical sense, of making a mistake in accepting or rejecting the hypothesis" (emphasis in original).

53 The probability value is the proportion of events, out of the total number of events, which do not support the null hypothesis of no effect.

54 Office of Technology Assessment, *Assessment of Technologies for Determining Cancer Risks in the Environment*, 1981; TJ Gill, GJ Smith, RW Wissler and HW Kunz, "The Rat as an Experimental Animal" (1989) 245 *Science* 269.

55 Gill et al, note 54 *supra* at 272. See also S Reynolds, S Stowers, R Patterson, R Maronpot, S Aaronson and M Anderson, "Activated Oncogenes in B6C3F1 Mouse Liver Tumors: Implications for Risk Assessment" (1987) 237 *Science* 1309 at 1310.

56 The data are cited in the US EPA data base IRIS, 11 February 1994.

assessment.⁵⁷ For instance, for lead, the weight of evidence classification is B2, “probable human carcinogen”, which is a way of stating that cancers developed at multiple sites in animal studies. Human evidence of cancer is “inadequate”. Thus, EPA’s recommendation was that numerical (cancer) risk estimates should not be used for lead.

The US EPA’s assessment of carcinogenic risk results from considering “[t]he reliability of conclusions ... from the combined strength and coherence of inferences appropriately drawn from all of the available evidence”.⁵⁸ The “science policy position” is that “tumor findings in animals indicate that an agent may produce an effect in humans.”⁵⁹ The absence of tumors in “well-conducted, long-term animal studies in at least two species provides reasonable assurance that the agent may not be a carcinogenic concern for humans”.⁶⁰

C. Biological Causation

The US EPA states that the most realistic dose-response models are those using *in vivo* data, although short-term mutagenic tests are also relevant as are the activation of oncogenes by types of mutations (eg, chromosomal translocations).⁶¹ This information can clarify the linearity or non-linearity of the dose-response model. The strength of the causal inference depends on the data⁶² and on the theoretical biological mechanism that determines the type of cancer dose-response model to be adopted. The US EPA states that:

The carcinogenicity of a direct-acting mutagen should be a function of the probability of its reaching and reacting with DNA. The activity of an agent that interferes at the level of signal pathway with many receptor targets should be a function of multiple reactions. The activity of an agent that acts by causing toxicity followed by compensatory growth should be a function of the toxicity.⁶³

The cancer process is an example of scientific conjectures about the shape of the dose-response, the experimental or epidemiologic data, or both. The process involves such steps as: cellular growth, differentiation, replication and death, including feedbacks that can result in non-linear dose-response. A carcinogenic substance can interfere with normal genetic and biochemical processes in different ways and transform a normal cell into a tumorigenic one. Such transformation can occur through faulty enzymatic repair of a heritable chemical

57 The “weights” are: (A) human carcinogen; (B1) probable human carcinogen from “limited” human data; (B2) probable human carcinogen with sufficient evidence in animals and inadequate or no evidence in humans; (C) possible human carcinogen; (D) not classifiable as to human carcinogenicity; and (E) evidence of noncarcinogenicity for humans. See, *Risk Assessment Guidance for Superfund, Part A*, EPA/540/1-89/002 at 7-11.

58 *Draft Revisions to Guidelines for Carcinogen Risk Assessment*, EPA 600/BP-92/003, (1992) at 9.

59 *Ibid* at 21.

60 *Ibid*.

61 See, for discussion, JM Bishop, “Oncogenes and Clinical Cancer” in RA Weinberg (ed), *Oncogenes and the Molecular Origins of Cancer*, Cold Spring Press (1989) 327.

62 Note 58 *supra* at 7.

63 *Ibid* at 3.

damage.⁶⁴ The factors include 'signalling' and control of genetic transcription, hormone chemistry, and changes that can affect the structure of the cell (eg the permeability of the cell's wall).

Furthermore, the linkage between predisposition and exposure to environmental factors is becoming increasingly clear. For example, genes YP1A1 and *Ha-ras* are involved with exposure to tobacco smoke in causing lung cancer (with *Ha-ras* causing non-adenocarcinomas in African-Americans); gene CYP2D6 and tobacco smoke, asbestos, and PAHs cause lung cancer; and GST1 and aromatic hydrocarbons also induce lung cancer (adenocarcinoma).⁶⁵

Understanding the biological processes leading to cancer helps to develop causally plausible dose-response models. Electrophilic chemicals that bind with the DNA and are potential mutagens, are analysed through structure-activity models to determine how analogous they are to known mutagens. The rates of formation of adducts, such as the hemoglobin adducts, is demonstrable at very low dose levels,⁶⁶ the functional relationship may be either linear or non-linear, depending on the chemical.⁶⁷ Radioimmunoassay can detect chemical adducts at extremely low concentration.⁶⁸

Enzyme systems have been found to have several functions. They co-operate in cell replication, controlling the cell cycle, and the expression of genes via transcription.⁶⁹ Thus, Koshland concluded that: "spontaneous errors ... from intrinsic DNA chemistry in the human body are usually many times more dangerous than chance injuries from environmental causes".⁷⁰ There may be differences among the species used to establish whether a chemical is carcinogenic in humans. Aspirin, for instance, is known to cause birth defects (not cancer) in rabbits but not in humans. Humans have an enzyme repair mechanism that rabbits do not have.⁷¹

Finally, chemicals can interact to cause cancer. The US EPA, for Polycyclic Aromatic Hydrocarbons (PAHs), assumed that the mixture of like chemicals are equally potent and has used an approximate upper bound on the estimated coefficient of the Linearized Multistage model (LMS) dose-response function to

64 *Ibid* at 4.

65 National Research Council, *Science and Judgment in Risk Assessment*, National Academy Press, Washington DC (1994) at H-2-5, Table H-3.

66 MA Pereira and LW Chang, "Binding of Chloroform to Mouse and Rat Hemoglobin" (1982) 39 *Chemico-Biological Interactions* 89.

67 E Bayley, T Connors, P Farmer, S Gorf, and S Rickard, "Methylation of Cysteine in Hemoglobin Following Exposure to Methylating Agents" (1981) 41 *Cancer Research* 2514.

68 IC Hsu, MC Poirier, SH Yuspa, RH Yolken, and CC Harris, "Ultrasensitive Enzymatic Radioimmunoassay (USERIA) Detects Femtomoles of Acetylaminofluorine-DNA Adducts" (1980) 1 *Carcinogenesis* 455.

69 E Culotta and DE Koshland, "The Molecule of the Year: How DNA Repair Works its Way to the Top" (1994) 266 *Science* 1926 at 1927.

70 DE Koshland, "Molecule of the Year: DNA Repair Enzymes" (1994) 266 *Science* 1925 at 1925.

71 *Ibid*.

set acceptable exposure levels for that mixture.⁷² The biochemical and cellular paths taken by PAHs, in general, and by B(a)P in particular, are:⁷³

[exposure \Rightarrow partition and distribution \Rightarrow metabolic transformations \Rightarrow adduct formation OR cell proliferation \Rightarrow mutation \Downarrow repair \Rightarrow end, OR \Rightarrow cell proliferation OR death \Rightarrow cell death \Rightarrow end, OR proliferation \Rightarrow adduct formation \Rightarrow mutation \Downarrow repair \Rightarrow end, OR second mutation \Rightarrow cell proliferation \Rightarrow tumor].⁷⁴

The metabolism of PAHs in rodents is similar to that of humans. However, the rate of formation of specific adducts and the elimination of products, (the formation and disappearance of an epoxide) varies from species to species. A key aspect of the carcinogenicity of B(a)P is that it is a genetic toxicant that requires metabolic activation before it becomes a carcinogen. The metabolic by-product binds with the DNA, forming a covalent adduct, which, if unrepaired, can initiate the cancer process.

IV. QUANTITATIVE MODELS OF DOSE-RESPONSE

An added difficulty is that the exposure in toxic tort disputes is generally orders of magnitude lower than the experimental exposure, requiring either predictions or extrapolations to values very near zero. Paradoxically, it is often difficult to discriminate among the alternative models in the relevant range of the data. A multistage model generally fits the data better than the single-hit model. The number of stages must have biological basis because too many stages (which would result in better fit) may not be biologically plausible for a specific cancer.

The quantal (ie presence or absence of the cancer) form of the multistage model has been the mainstay of regulatory work in American environmental law.⁷⁵ The US EPA used the Linearized Multistage dose-response model (LMS) for most cancers.⁷⁶ The LMS accounts for cellular changes occurring through the transitions from the normal stage to the preneoplastic stage, and from that

72 *Comprehensive Environmental Response, Compensation, and Liability Act* 1980, (CERCLA/Superfund), 42 USC § 9601 et seq.

73 International Agency for Research on Cancer, *Polynuclear Aromatic Compounds, Part 1, Chemical, Environmental, and Experimental Data*, vol 32 (1983); US EPA, *Health Effects Assessment for Polycyclic Aromatic Hydrocarbons (PAH)* ECAO, EPA 540/1-86-013 (1986).

74 The meaning of the symbols is as follows: \Rightarrow means 'results in', \Downarrow means 'results in a repair' and OR means 'an alternative' (either a repaired DNA adduct or a mutation which goes unrepaired, but not both).

75 US Environmental Protection Agency, *Risk Assessment Guidance for Superfund: Volume 1 - Human Health Evaluation Manual, Part C*, (Risk Evaluation of Remedial Alternatives) Interim, Publication 9285.7 01C (December 1991), Office of Emergency and Remedial Response, Washington DC.

76 The duration of exposure is lifetime. The form of the LMS is $\Pr(d) = 1 - \exp[-(q_0 + q_1d + q_2d^2 + \dots + q_n d^n)]$. $\Pr(\cdot)$ is the lifetime probability of cancer, d is the lifetime dose, and q_i are the parameters of the model estimated from experimental data. See KS Crump, "An Improved Procedure for Low-Dose Carcinogenic Risk Assessment from Animal Data" (1984) 5 *Journal of Environmental Pathology and Toxicology* 339.

stage to the cancerous stage, through transition rates linearly dependent on exposure.⁷⁷

The Moolgavkar-Venzon-Knudsen model (MVK) is a more biologically realistic cancer dose-response model than the LMS because it consists of a probabilistic birth-death staged cancer model.⁷⁸ The MVK model yields age-specific cancer incidence and can describe initiation, promotion, inhibition, and other aspects of exposure to a carcinogen. The essence of the model is that two cellular transformations are required to change a normal cell into a tumorigenic one; each transformation is a stage toward forming a tumor. The birth of a cell occurs in the same stage where its parents reside. A heritable, unrepaired transformation before cell division results in a transition to another stage. The MVK model accounts for cytotoxic effects through the difference between birth rate and death rate. The prevalence and incidence data are developed through epidemiologic studies. Thus, the predictions from the MVK model can be compared with existing data. This merges cellular models with epidemiological data.

The shape of the function that relates dose to response in toxicological risk assessment represents the non-linear cumulative distribution of responses for an adverse health endpoint. It is important because its choice affects the identification of the Lowest Observed Adverse Effect Level (LOAEL) and the No Observed Adverse Effect Level (NOAEL). These phrases represent exposures or doses that are used with factors of safety to determine acceptable exposure (or dose) levels for toxicants.⁷⁹ Thus, a NOAEL/100 would mean that the safety factor is 100.

A. Epidemiological Studies

Epidemiology is the study of the pattern of disease in human populations to discover the incidence (rate of occurrence of new cases per period of time) or prevalence (existence of cases in a suitable period of time) of a particular disease, in order to predict the risk of others developing the same condition in either similar or different settings. These results are particularly relevant in establishing causal associations because the unit of analysis is the human being and the study explains the changes in the prevalence or incidence rates.

77 Following LA Cox, "Assessing Cancer Risks: from Statistical to Biological Models" (1990) 116 *Journal of Energy Engineering* 189 at 199: "Explanation of carcinogenesis is organized around a few key parameters ... [which] provide the basic input data for the model, from which cancer hazard rates are predicted Many of the key parameters [or surrogates for them] can potentially be measured directly in the laboratory [or in cellular systems], rather than being estimated statistically from whole-animal bioassay response data. Thus, the MVK model can potentially use the empirical data from molecular epidemiology".

78 Following SH Moolgavkar, "Carcinogenesis Modeling: From Molecular Biology to Epidemiology" (1981) 7 *Annual Review of Public Health* 151, a simple formulation of this model yields the age-specific incidence rate for the cancer, given: the initial population of normal cells, the rates of cell transformation from the normal stage (containing stem cells) to the initiated stage (containing initiated cells) and from the initiated stage to the malignant stage (containing malignant cells), the average rates of cell formation (birth) and the average rates of cell death or differentiation.

79 Risk Assessment Guidance for Superfund, *Human Health Evaluation Manual*, vol 1, EPA/540/1-89/002 (December 1989) at 7-2.

However, reliable causal results often require lengthy and costly prospective studies, perhaps over a generation.

The evidence of adverse effects in human studies has been sufficient to justify intervention to eliminate the source of the problem, even if that evidence is circumstantial and the biological mechanism is not yet fully known. This is because:

unless epidemiologists have studied a reasonably large, well-defined group of people who have been heavily exposed to a particular substance for two or three decades without apparent effect, they can offer no guarantee that ⁸⁰continued exposure to moderate levels will, in the long run, be without material risk.

Briefly, epidemiological risk models are generally based on one of two hypotheses: the additive (absolute risk) or the multiplicative (relative risk) hypotheses. The choice between these hypotheses depends on knowledge about the incidence of the disease and a exposure-response model of the disease process. For instance, the absolute risk model is based on the assumption that the probability that the irreversible transformed cell will cause leukemia is proportional to exposure, while the background probability of cancer is independent of exposure. The relative risk is predicated on the hypothesis that cellular transformation is proportional to background and exposure.

The US EPA also discusses causation statistically, in terms of meta-analysis, and epidemiologically from Hill's criteria, with "temporal relationship" being "the only criterion that is essential", with none of the criteria being conclusive by themselves.⁸¹ We review these aspects in the next section.

V. UNCERTAIN CAUSATION

The brief review of probabilities (epistemic and so on) and their uses as weights associated with statements, theories and judgments had the purpose of leading to reasoning about uncertain causation.

A. Fundamental Aspects

All uncertain quantities are treated as random variables and functions. Uncertainties about functions, values, parameters, variables, and sampling are handled in a uniform computational framework based on conditional probabilities. Prior knowledge, information and beliefs can be represented by prior probability distributions and by conditional probabilities, given the variables influencing them. Legal causal reasoning suggests the ordered heuristic:

80 R Doll and R Peto, "The Causes of Cancer: Quantitative estimates of Avoidable Risk of Cancer in the United States Today" (1981) 66 *Journal of the National Cancer Institute* 1195 at 1219.

81 Note 58 *supra* at 15. The criteria are: temporal relationship, consistency, magnitude of the association, biological gradient, specificity of the association, biological plausibility, and coherence.

[Past experience \Rightarrow Empirical facts \Rightarrow Causal network] = [Legal cause-in-fact]

Legal reasoning is consistent with Richard Jeffrey's "radical probabilism"⁸² in which the assessment of the legal outcome goes beyond purely Bayesian empirical reasoning⁸³ to include updating rules for prior beliefs and updating based on the empirical findings. Specifically, it adds the principle that retractions are allowed. Jeffrey has called traditional Bayesian analysis "rationalistic Bayesianism".⁸⁴

The reason for Jeffrey's radical Bayesianism is that it is virtually impossible, other than in the simplest of cases, truly to capture, in the 'prior' distribution, required by rational Bayesianism because one is often unable "to formulate any sentence upon which we are prepare to condition ... and in particular, the conjunction of all sentences that newly have probability of one will be found to leave too much out".⁸⁵ Fortunately, following de Finetti:

Given a coherent assignment of probabilities to a finite number of propositions, the probability of any proposition is either determined or can coherently be assigned any value in some closed interval.⁸⁶

(i) *Empirical Evidence Represented by Likelihood Functions*

In developing ways to account for probabilistic evidence, empirical data are summarised by likelihoods or likelihood functions. The likelihood is a probabilistic measure of the support, given by the data, to a statistical hypothesis about the possible values of a population parameter.⁸⁷ The likelihood principle states that the effect, if any, on estimates of the parameters of the model depends only on its likelihood.⁸⁸ The maximum likelihood estimate is the value that maximises this likelihood, given the observed data and model.

B. Updating, Priors and Coherent Beliefs⁸⁹

A natural measure of uncertainty is the conditional probability, used in Bayes' theorem,⁹⁰ where what is sought is the posterior probability. Conditioning is

82 Jeffrey, note 51 *supra*, p 3.

83 C Glymour, *Theory and Evidence*, Princeton University Press (1980) p 69.

84 Note 51 *supra*, pp 2-3.

85 *Ibid*, p 83.

86 *Ibid*, p 100.

87 D Clayton and M Hills, *Statistical Models in Epidemiology*, Oxford Science Publications (1994).

88 We let y be a vector of uncertain quantities to be predicted (eg, health responses), x be a matrix of explanatory variables (eg, one or more forms of exposures, socio-economic and other independent variables), and $p(y; x, b) = pr(Y = y | X = x; b)$ be a prior conditional probability model for the relation between the two random variables X, Y , and b the vector of parameters that needs to be estimated. Given a probability model symbolised by $pr(y; x, b)$, the likelihood function for b is $pr(y; x, b)$ considered as a function of b instead of as a function of x and y . Then $L(b; x, y, pr)$ denotes the likelihood function for b based on observed data vectors, for a probability model $pr(\cdot)$.

89 LA Cox and PF Ricci, "Dealing with Uncertainty: From Health Risks Assessment to Environmental Decisionmaking" (1992) 118 *Journal of Energy Engineering* 77.

understood as follows. If $F(b)$ denotes the prior (subjective) joint distribution of uncertain quantities, $(F | L)$, read 'F given L', is the posterior probability distribution (developed by Bayes' Theorem using the prior distribution and the likelihood functions) for the uncertain quantities in b , obtained by conditioning F on the information summarised in L . When the calculation of the updated probabilities $(F | L)$ becomes computationally burdensome, there are methods, based on belief networks, to minimise the number of combinations and the number of events.⁹¹

Evidence from multiple sources is combined by successive conditioning. Thus, if L_1 and L_2 are the likelihood functions generated by two independent data sets, then, provided that L_1 and L_2 are based on the same underlying probability model, the posterior probability for the parameters, after incorporating evidence from both sources, is $[(F | L_1) | L_2]$.

(i) *Alternative Models and Assumptions*

The calculation of $(F | L)$ from prior beliefs formalised (perhaps through elicitations of expert opinion) in $F(\cdot)$ and assumptions and evidence forming $L(\cdot)$ depends on the probability model, pr [short for $pr(y; x, b)$], and on the data. It is the uncertainty about the correct model, pr_i (out of several alternative models) that is often the greatest in causal models. To include it, let $\{pr_1, \dots, pr_n\}$ denote the set of alternative models that are mutually exclusive and collectively exhaustive. Also, L_1, \dots, L_n denote the corresponding likelihood functions for these competing models, and w_1, \dots, w_n are the corresponding judgmental probabilities, or weights of evidence, that each model is correct. If the models are mutually exclusive and collectively exhaustive, these weights should sum to 1. Then, the posterior probability distribution obtained from a prior $F(\cdot)$, data, and model weights of evidence w_1, \dots, w_n is the weighted sum $[w_1(F | L_1) + \dots + w_n(F | L_n)]$.

(ii) *Advantages*

There are several advantages to a formal probabilistic method because they result in a coherent⁹² approach. These include:

- 1 Showing the range of possible values. Probability density functions can be used to show how much statistical uncertainty can potentially be reduced by further research which would narrow the distributions.
- 2 Distinguishing the contributions of different sources of evidence, uncertainty and information, identifying where additional research is most likely to make a significant difference in reducing final

90 A demonstration of Bayes' theorem follows. From $pr(A \text{ and } B) = pr(A)pr(B|A)$ and $pr(B \text{ and } A) = pr(B)pr(A|B)$ equate the right hand sides to obtain $pr(A|B) = [pr(B|A)pr(A)]/pr(B)$. See DV Lindley, *Bayesian Statistics: A Review*, SIAM (1984).

91 J Pearl, "Bayesian and Belief-Functions Formalisms for Evidential Reasoning: A Conceptual Analysis" Proceedings of the 5th Israeli Symposium on Artificial Intelligence, December 1988.

92 The coherence means that the axiomatic properties of the system are established first, and the methods follow.

uncertainty. A measure used to resolve this issue is the expected value of information.⁹³

- 3 Consistency with sensitivity analyses to show changes to the output (ie, of the mean and mode of the probability density function), to uncertain data and to the choice of a prior distribution.
- 4 Inclusion of the most recent information, without forcing the adoption of default values or models, through Bayesian updating rules.⁹⁴

(iii) Disadvantages

The approach is only appropriate for a single decision maker, when the distinctions among subjective and objective uncertainties are largely irrelevant. All that matters is the joint posterior density function for all uncertain quantities. Other limitations include:

- 1 Probability models cannot adequately express ignorance and vague or incomplete thinking. Other measures and formal systems may have to be used to characterise these forms of uncertainty such as 'fuzzy' numbers, arithmetics and logic.
- 2 A problem related to the need to specify a prior probability is that the assumed probability model, $pr(y; x, b)$, often requires unrealistically detailed information about the probable causal relations among variables. The Bayesian answer to this problem is that an analyst normatively should use knowledge and beliefs to generate a probability model. An opposing view is that the analyst has no such justification, and should not be expected or required to provide numbers in the absence of hard facts. This view has led to the Dempster-Shafer belief functions to account for partial knowledge.⁹⁵ Belief functions do not require assigning probabilities to events for which they lack relevant information: some of the 'probability mass' representing beliefs is uncommitted.
- 3 An assessment that gives a unit risk estimate of 1×10^{-6} for a chemical through an analysis that puts equal weights of 0.5 on two possible models, one giving a risk of 5×10^{-7} and the other giving a risk of 1.5×10^{-6} , might be considered to have quite different implications for the 'acceptability' of the risk than the same final estimate of 1×10^{-6} produced by an analysis that puts a weight of 0.1 on a model giving a risk of 1×10^{-5} and a weight of 0.9 on a model giving a risk of zero.
- 4 Probability modelling makes a 'closed world assumption' that all the possible outcomes of a random experiment are known and can be

93 Lindley, note 90 *supra*, pp 72-3.

94 Pearl, note 91 *supra*.

95 AP Dempster, "A Generalization of Bayesian Inference (with discussion)" (1968) 30 *Journal of the Royal Statistical Society* (B) 205; G Shafer, "Belief Functions and Parametric models (with discussion)" (1982) 44 *Journal of the Royal Statistical Society* (B) 322.

described (and, in the Bayesian approach, assigned prior probability distributions). This can be unrealistic: the true mechanism by which a pollutant operates, for example, may turn out to be something unforeseen. Moreover, conditioning on alternative assumptions about mechanisms only gives an illusion of completeness if the true mechanism is not among the possibilities considered.

The next section is a review of methods that are used in regulatory and toxic tort law when dealing with multiple studies and uncertain causation. These are used to:

- summarise independent results (such as the potency of a carcinogen) that were developed independently from one other (meta-analysis);
- cumulate the uncertainty in a causal chain (using methods such as Monte Carlo simulations); and
- deal with lack of knowledge about the distribution of the population (perhaps using methods such as the 'bootstrap resampling method', discussed below).

C. Empirical Ways to Deal with Multiple Studies and Uncertainties

Recently, meta-analysis⁹⁶ has been imported from the psychometric literature into health assessments (in particular, epidemiology) quantitatively to assess the results from several independent studies on the same health outcome.⁹⁷ This statistical method does so by generating meta-distributions of the quantitative results reported in the literature.⁹⁸ Meta-analysis is a pooling device that "goes beyond statistical analyses of single studies; it is the statistical assessment of empirical results from independent and essentially identical sampling".⁹⁹ It has been used to develop the empirical distribution of the estimated coefficients of exposure-response models (ie, the estimated parameters form a very large number of regression equations), from several independent studies, to develop a meta-distribution of the selected values. The focus of meta-analysis is the explicit, structured and statistical review of the literature with the expressed intent to confirm the general thrust of the findings reported in that literature. Meta-analytic studies are second-hand because the researcher has no control over the data themselves. In a very real sense the researcher takes the data as he or she finds them.

96 WG Cochran, "Problems Arising in the Analysis of a Series of Similar Experiments" (1937) 1 *Journal of the Royal Statistical Society* (Supplement 4) 102; L Hedges and I Olkin, *Statistical Methods for Meta-Analysis*, Academic Press (1985).

97 R Fisher, "Combining Independent Tests of Significance" (1948) 2 *The American Statistician* 30.

98 Peer review of heterogeneous material is assumed to be unbiased and to be able to separate the wheat from the chaff.

99 This summarises the ideas by GV Glass, "Synthesizing Empirical Research: Meta-Analysis" in SA Ward and LJ Reed (eds), *Knowledge Structure and Use: Implication for Synthesis and Interpretation*, Temple University Press (1983) 24.

The judicial acceptability of meta-analysis appears established. In *In re Paoli RR Yard PCB Litigation*,¹⁰⁰ the trial court had excluded meta-analysis results holding that those studies were not related to the plaintiff's health. The appeal admitted the meta-analyses as probative findings that exposure to PCBs could result in future risk of cancer.

Another aspect of a causal chain is the propagation of the uncertainty of each variable in that chain. Take, for instance, the casual chain that begins with the emissions of toxic pollutants and ends with the dose to the DNA. This chain consists of very many components that require mathematical operations to link them. If we simplify this causal chain to the function: $R = f(x, y, z)$, each of these variables (x , y and z) have a known distribution, say uniform, triangular and Poisson, to represent their uncertainty. Then the question is: what is the shape of the distribution of R , given that of the variables x , y and z ?

A class of methods for developing the answer is Monte Carlo (probabilistic) computations. They replace an integrand with a stochastic simulation based on sums of random numbers: integration is replaced by a probabilistic simulation which returns the unbiased expected value of the estimator and its variance.

Another class of numerical probabilistic methods (that can generally be used without having to assume or otherwise know the shape of the distribution function for the population) is the bootstrap resampling method. The essence of this computational scheme is the random sample of observations, the empirical distribution, from which a large number of bootstrap samples is obtained to approximate large sample estimates of central tendency, dispersion and other estimators, including 'confidence' limits. The approach uses empirical distribution functions which are taken to be the simple estimate of the entire distribution.

The bootstrap generates a relatively large number of independent samples drawn at random with replacement, from the single original sample, to develop an empirical cumulative distribution approximating the unknown population distribution. A practical use of bootstrap methods is to develop confidence intervals to represent the variability of statistical estimators such as the median, often a difficult matter.

D. Inference

Inference can be deductive: the conclusion is not false, given that the premise is true. If the conclusion is false when the premises are true, then the inference is nondemonstrative. A characteristic of scientific reasoning is that it uses hypotheses and initial conditions, from which the prediction flows. This is the hypothetical-deductive construct where the initial conditions are taken as true at least until observations about the prediction forces a retraction.¹⁰¹ Statistical inference is inductive because the evidence flows from the observed data to the hypotheses.

100 916 F 2d 829 (3d Cir 1990) at 29.

101 WC Salmon, *The Foundations of Scientific Inference*, University of Pittsburgh Press (1966) p 18.

The way plausibly to proceed is through induction, avoiding semantic and syntactic vagueness (perhaps a tall order in legal reasoning) or using the methods suggested here. The inductive apparatus is plausible if it:

- relies on coherent probabilistic reasoning or other measures of uncertainty;
- admits retractions without logical fallacies;
- is dynamic, allowing new information to be added through Bayesian updating or other formal rules; and
- is mechanistic (eg, biologically-motivated at the fundamental level).

Probabilistic results are not a matter of 'true' or 'false' answers. They are, rather, an honest and accurate balancing of uncertain data and theories adduced in a causal argument. On these notions, a legally admissible scientific result must satisfy each of the elements of the calculus on which it is developed and must be verifiable. Following Carnap, the probabilistic degree of confirmation can be legal balancing,¹⁰² but not necessarily at the 'more likely than not' level, as we argued in Part I. For, what is this level that can amount to an impenetrable and often incoherent barrier for the plaintiff?

VI. CONCLUSION

Part II has laid the heuristic foundations for reasoning with uncertain facts and models as a means-ends framework for judicial reasoning. The framework uses probabilistic measures and methods consistent with judicial and legislative reasoning. Although we have used probability and statistical theory as the means to achieve a just expression of causation, there are other, equally formal, methods to represent uncertainty and variability.

We have heuristically demonstrated that the management of causation in toxic tort law would greatly benefit from using the methods and principles that we have described and exemplified. To do otherwise invites chaotic and unjust allocations of liability.

A most recent case casts some light on the issues discussed in this paper. A critical aspect of *Kumho*, just decided by the US Supreme Court, as a measure of *Daubert*-reliability, is that the Court confirms for the third time in five years a "judicial broad latitude" as a means to allow scientific gatekeeping by trial judges.¹⁰³ *Kumho* holds that *Daubert* applies to all contexts to which FRE 702 applies,¹⁰⁴ regardless of whether scientific, technical, or other specialised

102 R Carnap, *The Continuum of Inductive Methods*, University of Chicago Press (1952)

103 *Kumho Tire Co Ltd v Carmichael* [1999] US Lexis 2189 at 2195, US Supreme Court, decided 23 March 1999.

104 FRE 702 states that: "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 thereto in the form of an opinion or otherwise".

knowledge is being introduced into the controversy. The FRE 702 “latitude” allowed to experts’ opinions is balanced by the *Daubert*, *Joiner*, and *Kumho* “latitude” given to trial judges to admit that evidence.

As discussed in Part I, *Daubert* stands for the admissibility of scientific, knowledge requiring expert testimony, under FRE 702, through four factors.¹⁰⁵

The critical term is “knowledge”,¹⁰⁶ not the adjectives modifying it.¹⁰⁷ Those four factors “may be” required to establish the reliability of *expert knowledge*. These are not the sole factors: the context of the case determines the appropriate number of factors to be set up by the trial judge.¹⁰⁸ This reflects the inherent “flexibility” of the wording of FRE 702.¹⁰⁹

The argument used by the US Supreme Court to ascertain whether or not expert evidence is admissible at trial is that:

- it is *not* the “reasonableness” of the expert’s use of a specific form of inference to determine the causation that is unreliable; but rather
- it *is* the “reasonableness” of a specific form of inference *and* the expert’s methods for analysis *and* deriving conclusion about the cause of failure (of tyre damage leading to a car crash that killed one person and injured three more) that is unreliable.¹¹⁰

In other words, if there are relevant but unaccounted initiating and causal events, and the causal analysis is purely judgmental, then the expert’s testimony can be insufficient to pass through the judicial gates. The process is inductive, or more precisely, empirico-inductive.¹¹¹

Kumho correctly extends *Daubert* beyond scientific knowledge and requires trial judges to establish a flexible protocol capable of filtering out dubious evidence. Reliable and admissible expert evidence must be cleared according to a coherent frame of reference for the parties and the judge. The “leeway” allowed by *Kumho* is precisely matched by the analytical aspects developed in Parts I and II of this work. The *Kumho* Court appears to believe that the Federal Rules of Evidence both expedite trials, and seek the “truth” and the “just determination” of the judicial process.¹¹² The trilogy of *Daubert*, *Joiner* and *Kumho* still provides no operational or coherent framework for balancing the

105 *Daubert v Merrell Dow Pharmaceuticals Inc* 509 US 579 (1993) at 592-4. The factors are whether (a) a “theory or techniques ... can be (and has been) tested”; (b) its “peer review and publication”; (c) its “known or potential rate of error” and whether there are “standards controlling the technique’s operation”; and (d) its “general acceptance in the relevant scientific community”. The factors “may bear on the judge’s gatekeeping determination” and are not exhaustive: *Kumho*, note 103 *supra* at 2194.

106 *Kumho*, note 103 *supra* at 2193, citing *Daubert*. *Daubert* dealt with scientific knowledge, not with the other forms of evidence listed in the FRE 702.

107 *Ibid*, citing *Daubert*, note 105 *supra* at 589-90.

108 *Ibid* at 2195.

109 *Ibid*.

110 *Ibid*.

111 *Ibid* at 2193, citing Learned Hand’s view that experts’ “general truths are derived from ... specialized knowledge” (citation omitted).

112 *Ibid* at 2195, referring to FRE 102.

latitudes give to experts and to trial judges. We are now even more convinced than ever that our suggestions are worth further development.