Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

Journal of Evaluation in Clinical Practice ISSN1356-1294

Interpreting risk as evidence of causality: lessons learned from a legal case to determine medical malpractice

Mathew Mercuri PhD[1] and Brian S. Baigrie PhD[2]

[1]Assistant Professor, Department of Medicine, Division of Emergency Medicine, McMaster University, Hamilton, Ontario, Canada
[2]Professor, Institute for the History and Philosophy of Science and Technology, University of Toronto, Toronto, Ontario, Canada

Keywords
Bayesian, causality, evidence, frequentist, risk

Correspondence
Dr Mathew Mercuri
Division of Emergency Medicine
McMaster University
Hamilton General Hospital
237 Barton Street East, Hamilton
ON L8L 2X2
McMaster Wing, Rm. 242
Canada
E-mail: matmercuri@hotmail.com

Accepted for publication: 15 May 2016

doi:10.1111/jep.12580

Abstract

Translating risk estimates derived from epidemiologic study into evidence of causality for a particular patient is problematic. The difficulty of this process is not unique to the medical context; rather, courts are also challenged with the task of using risk estimates to infer evidence of cause in particular cases. Thus, an examination of how this is done in a legal context might provide insight into when and how it is appropriate to use risk information as evidence of cause in a medical context. A careful study of the case of Goodman v. Viljoen, a medical malpractice suit litigated in the Ontario Superior Court of Justice in 2011, reveals different approaches to how risk information is used as or might be considered a substitute for evidence of causation, and the pitfalls associated with these approaches. Achieving statistical thresholds, specifically minimizing the probability of falsely rejecting the null hypothesis, and exceeding a relative risk of 2, plays a significant role in establishing causality of the particular in the legal setting. However, providing a reasonable explanation or establishing "biological plausibility" of the causal association also seems important, and (to some) may even take precedent over statistical thresholds for a given context.

Introduction

The process of translating estimates of risk derived from clinical study to claims of causality for a particular patient poses a considerable problem in the practice of medicine. While it may be the case that claims of causality are unlikely to be confirmed with certainty, clinical application of a particular therapy is predicated on the belief that where it is used a specified health outcome can be achieved. That is, there is some expectation (by both the physician and patient) that the patient will (likely) avoid a disease or have his or her symptoms minimized or eliminated, making the perceived relationship between the therapy and outcome 'effectively causal'. However, this begs the question as to under what circumstances one can reasonably consider information derived from clinical study as functional evidence of causality in individual cases.

To determine where and how risk estimates might provide evidence of cause in a clinical context, it is prudent to determine how the claim that some factor is a cause of some health outcome is understood. Rothman and colleagues define cause as "an event, condition, or characteristic that preceded the disease onset and that, had the event, condition, or characteristic been different in a specified way, the disease either would not have occurred at all or would not have occurred until some later time" [1; p.6]. Epidemiology studies, including clinical trials, estimate the relative frequency of some outcome given the presence or absence of some risk factor (usually a clinical intervention, such as a medical therapy). Thus, epidemiology studies do not directly address the issue of causality (either at the population or individual level); rather, the relative difference in risk is used to imply that were the risk factor not present (i.e. event, condition, or characteristic were different), the outcome (i.e. disease) would not have occurred. As such, consumers of these studies are left to translate this relative difference in risk into evidence of cause for a particular context, in some cases at their own peril (as we propose to show in the following).

Epidemiology does not provide definitive instruction on how to best use risk information as evidence of cause, most notably in translating risk estimates at the population level to claims of cause for a particular patient. This situation is not unique to the medical context. For example, courts are often challenged with establishing cause in a particular case on the basis of risk information derived from scientific study. The burden of the plaintiff is to establish that were it not for the actions of the defendant, the outcome in question would not have occurred. This is known as the 'but for' test (i.e. *sine qua non*), and is the underlying principle of the law of torts [2,3]. This task parallels that of epidemiologists in establishing cause, as outlined in the definition by Rothman *et al.*, presented earlier. However, where causation cannot be definitely established, the 'but for' test is considered to be 'unworkable'. In such circumstances it is the task of the plaintiff to establish that the defendant materially contributed to the occurrence of the outcome

in question [2]. Again, one can see an intriguing (but perhaps spurious) parallel to the task of the physician in establishing the benefit of a specified therapy (i.e. risk information from the clinical trial as evidence that the therapy will materially contribute to attaining the desired health state in the patient). Thus, an examination of the process of how risk information is translated into evidence of cause in a legal context might provide insight into when and how it is appropriate to do so, and illustrate conceptual challenges regarding both what constitutes evidence of cause, and how to interpret findings from epidemiologic study. This will be achieved through the examination of the case in Goodman v. Viljoen [4], where it was successfully argued that 'but for' the actions of the physician of the plaintiff, a specified harm to the plaintiff would have been avoided or reduced in severity. The relevant details of this case are as follows:

In August 1995, Mrs Goodman was admitted to hospital for premature rupture of membranes and subsequent preterm twin birth. The children would later be diagnosed with cerebral palsy (CP). The plaintiff alleged that had she (i.e. Mrs Goodman) received proper advice and assessment from her physician 2 days earlier, she would have been admitted to hospital at that time and would have benefited from a full course of antenatal corticosteroids (ACS). Currently a standard of care in cases where there is risk premature birth or that such is imminent, ACS is administered in order to reduce the risk of neonatal death, respiratory distress syndrome, and intraventricular haemorrhage [5]. Some studies seem to indicate that neurodevelopment, specifically the avoidance of developmental delay and CP, might benefit from ACS [6]. The position of the defence was that ACS would not have prevented or reduced the severity of CP in Mrs Goodman's children, as the evidence from clinical science (available at the time of trial) did not support the claim of a causal association between ACS and the incidence/severity of CP. The ruling of the judge was in favour of the plaintiff.[1]

It is not our purpose here to settle the debate over the veracity of a causal association between ACS and CP, nor is it our intention to question if the defendant was indeed negligent in his actions or responsible for the outcome. These are interesting questions for further examination elsewhere. Rather, our goal is to critically appraise the clinical research with respect to what extent this information can be interpreted as evidence for or against a causal association between ACS and CP; at face value, does this evidence help to settle the matter of a causal association between ACS and CP? We will orient our examination around three distinct questions that were relevant to the assessment of causality in the *Goodman* trial, which we will address to in order:

1 Is ACS (or lack of its administration) a contributing cause of CP?

[1]The judge's ruling was as follows: 'Both Drs Barrett and Perlman, exercising their clinical judgement and expertise, considering the statistical evidence, answered unequivocally that it is more likely than not that had the twins received a full course of steroids, they would not have suffered the injuries they did, or the magnitude of those injuries would have been reduced. I agree with this conclusion. It is supported by the statistical and medical evidence and the common sense inferences that flow therefrom. I therefore find that the plaintiffs have established on a balance of probabilities that the defendant's negligence caused the CP from which they now suffer' ([4]; p.32, paragraph 207).

Table 1 Cases of cerebral palsy with respect to antenatal corticosteroid therapy

	CP	~CP	Total	Risk
~ACS	28	386	414	0.068
ACS	20	470	490	0.041
Total	48	856	904	0.053

Data adapted from Cochrane review of clinical trials by Roberts *et al.*, 2006 [8]. Relative risk = 1.66 (0.068/0.041).
ACS, antenatal corticosteroids; CP, cerebral palsy.

2 If ACS is administered[2] in an individual case, is it more probable than not that cerebral palsy will be avoided?

3 If cerebral palsy is not avoided after administration of ACS, does the administration of ACS reduce its severity?

Question 1: Is ACS (or lack of it administration) a contributing cause of cerebral palsy?

The plaintiff in *Goodman* argued that CP occurred because ACS was not administered (~ACS). Data acquired from a systematic review and meta-analysis of five randomized clinical trials demonstrates that ACS does not entail that CP would be avoided (~CP) [7,8]. In a combined sample of 490 women admitted to hospital for premature labour and administered ACS, 4.08% (20/490) of them gave birth to a child diagnosed with CP. This data is presented in Table 1. Assuming ACS was administered correctly, CP was correctly diagnosed, and the trial participants reflect the general population of interest, this data implies that CP has causes not always addressed by ACS. It was observed in the systematic review that 6.76% (28/414) of children born premature to women, who were not administered with ACS, were diagnosed with CP. Thus, the probability of CP given ~ACS is approximately 0.07. In other words, it is quite unlikely one will be born with CP even when ACS is not administered. This is a weak position from which to make claim causality.

However, this is not to say that failure to administer ACS does not contribute to the occurrence of CP. Assuming each study in the review was effective in eliminating systematic differences in confounding variables between groups (i.e. through randomization and blinding), there is an absolute difference of 2.68% (6.76%–4.08%) in the rate of CP that can reasonably be attributed to ACS. In other words, a fraction of the CP cases in the ~ACS group would likely have been avoided had ACS been administered, and thus, failure to provide ACS could be seen as contributing to the occurrence of those cases.

What has been suggested so far is a reduction in CP risk in the sample population that can be attributed to ACS.[3] This does not necessarily mean that we can expect the same results in the general population. It is possible that the randomization process failed to balance the groups on some important characteristics related to CP, and/or what was observed was due to a sampling error, such

[2]In the case of Goodman v. Viljoen [4], ACS was administered. What is at issue is the plaintiff's claim that the timeliness of injection was insufficient to gain the full benefit of the therapy. Therefore, I will assume that ACS was not administered for the purpose of this paper.

[3]See Broadbent (2013) [10] for further discussion on this interpretation of epidemiologic measures of causal association.

that the participants in these studies were not representative of the whole population of children born premature. Inferential statistics can provide us some guidance in these matters.

Let us start by assuming there is no true difference between those who receive ACS and those who do not with respect to the proportion of children born with CP. We will call this our 'null' hypothesis. We can calculate the probability that we would observe a given distribution or one more extreme, or find a difference in the absolute risk of ≥2.68% in our sample in the event that the null hypothesis is indeed true. If the probability was sufficiently low, it would give us grounds to reject the null hypothesis, implying that the difference we observed can be inferred to the whole population. Using statistical methods typical of clinical medicine research (i.e. chi-square test), the probability was calculated at 6.5% (i.e. $P = 0.065$).

What should we make of this probability? The convention in clinical medicine research is to reject the null hypothesis when the calculated probability is less than 5% (i.e. $P < 0.05$). Thus, from the perspective of clinical medicine research, there is no reason to believe that ACS has an effect on reducing the risk of CP. Despite the observed 2.68% absolute increase in risk, whether or not (and to what extent) ~ACS is a contributing cause of CP is inconclusive. There is a caveat to this interpretation. Because frequentist statistical methods, such as those described earlier, seeks to minimize error in the interpretation of study outcomes over the long run, it is possible that we may (1) incorrectly reject the null hypothesis when it is indeed true or (2) fail to reject the null hypothesis when it is indeed false. One cannot know for certain with any particular study if either of these errors has occurred, limiting our ability to conclude anything definitive with respect to the relationship between the variables in question from any single study. What we do know is that over the course of many studies this approach will rarely let us down as a general rule of interpretation.

Failing to meet the statistical burden of a causal association in epidemiologic study, which traditionally relies on a frequentist statistics, the plaintiff in *Goodman* took the unprecedented step in Canadian judicial context of adopting a Bayesian approach to interpreting trial data. Bayesian calculation seeks to determine the probability of some outcome (in this case, the relationship between CP and ACS) given some observation (in this case, the data from the systematic review). This approach invokes putative prior knowledge of the probability of the outcome, upgrading or downgrading this probability as new information is considered. Using the Bayesian approach, the plaintiff's expert witness calculated a 90% probability of a reduction in the risk of CP had ACS been administered, based on the clinical study data. In other words, this witness contended that it is highly likely that ACS contributes to reducing the occurrence of CP. This is not to say the ~ACS is a cause of CP, but rather that ~ACS is a contributing factor in its occurrence.

Both quantifying the prior probability and interpreting the assigned posterior probability are subjective, again limiting our ability to definitively identify ACS as a contributing cause of CP. Unlike in the frequentist method of hypothesis testing, Bayesian calculations do not provide guidance on how to interpret the posterior probability when it is short of 100% (or equal to 0%). It is noteworthy to mention that Bayesian analyses are uncommon in clinical trials or in the assessments of pharmaco-epidemiologic data. Furthermore, as the use of Bayesian analyses in court decisions is also rare, we believe that its use in the case of *Goodman* warrants careful consideration [9]. We will further discuss how differences in frequentist and Bayesian statistical inferences impact clinical decision-making in the context of *Goodman* (below).

Question 2: If ACS is administered in an individual case, is it more probable than not that cerebral palsy will be avoided?

Let us assume that ~ACS is a contributing cause of CP. That is, we will assume that the sample data presented in Table 1 is indeed representative of the whole population of children born premature. We are now challenged with demonstrating that it is more probable than not that CP would have been avoided had ACS been administered. In a legal context, specifically where more than one agent can cause the outcome in question, the standard of evidence is to establish that the 'balance of probabilities' favours one agent over another. That is, given 'state A' over 'state B', it is more probable than not that 'outcome X' would be avoided [10,11]. This equates to >50% of the combined risk attributed to the source (e.g. state A) in question, or in statistical terms, a relative risk (RR) of >2.

The risk of CP in the ~ACS sample population was estimated at 6.76%. As it is the case that for every 93 children born premature and without CP, there are only seven children born premature and with CP, one could argue that it is more probable than not that CP would have been avoided even if ACS had not been administered. That aside, the court's interest was ascertaining the probability that this particular case of CP was due to ~ACS rather than some other factor. In order to state that it is more probable that cerebral palsy would have been avoided, one must demonstrate that the risk attributable to ~ACS was greater than the risk from other sources.

The data acquired from the systematic review seems to imply that this is not the case. As we mentioned in the previous section, only a fraction of the cerebral palsy cases in the ~ACS sample population would likely have been avoided had ACS been administered – the 2.68% absolute risk attributed to ~ACS use is less than the 4.08% absolute risk that we are led to believe is independent of ACS therapy. In statistical terms, this equates to a RR of CP for ~ACS (compared with ACS) of 1.66 (i.e. we expect 2/3 more CP cases in the ~ACS population) – well below the RR threshold of 2 required by the 'balance of probabilities' criteria to claim someone is materially responsible in a legal context.[4]

Looking at the problem another way, consider a population of 10 000 children born premature. If none of their mothers were administered ACS, one could estimate that 676 of these children will develop CP. Assuming ACS is effective in reducing the risk of CP, our estimate is that 268 of the 676 CP cases would have been avoided had ACS been administered (Group A), leaving a

[4]The 66% increased risk is not statistically significant using the frequentist approach, as mentioned in the previous section. Thus, we cannot be sure that there is an increased risk, or that the magnitude of increase is as calculated (assuming the risk is in fact higher for ~ACS). The argument that it is not the case that it is more probable than not that cerebral palsy would have been avoided is made on the assumption that the estimated increased risk is factual. If we believe that there is no effect of ACS on cerebral palsy, then it is unnecessary to argue that it is more probable than not that cerebral palsy would have been avoided, as it would not be possible that those in the ACS group could have a reduced risk as a result of ACS administration.

balance of 408 cases where CP would not have been avoided had ACS been administered (Group B). For it to be more probable than not that CP would have been avoided had ACS been administered, the probability that a randomly chosen individual with CP in this population is from Group A must be greater than the probability he or she is from Group B. This is not the case here – an individual has a better chance of being from Group B than Group A.

Using the Bayesian approach, the plaintiff's expert witness calculated the risk of CP increases in the order of 70% where ACS is not administered. Again, considering our population of 10 000 children born premature, if we assume a base rate of 408 cases of CP had ACS been administered, the 70% increase in risk would mean an additional 281 cases had ACS been denied. From the Bayesian perspective, it is slightly more probable that that a randomly chosen individual with CP would have come from the group that could have benefited from ACS, when compared with that estimate using the frequentist approach. However, as was the case using the frequentist analysis, it is still more likely that such an individual would have come from the group where CP would not have been avoided. Thus, there is no statistical support for the claim that it is more probable than not that cerebral palsy would have been avoided had ACS been administered, based on the data provided in the systematic review of clinical trials and irrespective of the statistical methodology (frequentist or Bayesian) to which one subscribes.

Question 3: If cerebral palsy is not avoided after administration of ACS, does the administration of ACS reduce its severity?

It is not our purpose to challenge the validity of the plaintiff's diagnosis of CP. However, an examination of whether or not ACS would have reduced the severity of CP does require some understanding of what CP is, how it is diagnosed and what it is that could be reduced in severity. Historically, the term 'cerebral palsy' has been used to describe a wide range of permanent disorders related to neurodevelopment and neuromuscular control of movement and posture. Although attempts have been made to sharpen the clinical definition of this condition, the general agreement among experts is (1) CP represents a condition with significant heterogeneity of aetiology and types and severity of impairments and (2) CP is 'essentially a clinical formulation based on phenomenology' [12; p.9]. We will base our discussion on the definition provided by an international panel of experts on CP, as follows:

> Cerebral palsy (CP) describes a group of permanent disorders of the development of movement and posture, causing activity limitation, that are attributed to non-progressive disturbances that occurred in the developing fetal or infant brain. The motor disorders of cerebral palsy are often accompanied by disturbances of sensation, perception, cognition, communication, and behaviour, by epilepsy, and by secondary musculoskeletal problems. [12; p.9].

The virtue of this definition is comprehensive in the range of issues encompassed by cerebral palsy, but it is nevertheless quite vague in many respects. For example, it is not clear how one measures a disorder of movement and posture independent of its purported manifestations (i.e. activity limitations). It is conceptually difficult to attribute the cause of some effect to a particular entity if that entity cannot be measured independently of the effect. It is also not clear to what extent movement and posture abnormalities must limit activity in order to be classified as CP. The authors of this passage provide reference to the World Health Organization's definition of activity limitation as 'difficulties an individual may have in executing activities'. While this meaning of activity limitation is more precise, it is still not clear as to how one determines 'difficulty' in 'executing activities'. What constitutes 'difficulty?' How is 'difficulty' measured? Which activities are relevant? As no objective scale or criteria to determine if an individual has CP is available, the assessment of such requires judgement by a physician. Each physician will have some threshold by which he or she deems activity limitation as clinically relevant, and attributable to a congenital, permanent, and non-progressive movement and posture disorder. That is, there is a line of demarcation that separates a diagnosis of CP from a subclinical manifestation of symptoms. Those in the latter group would not be diagnosed with CP, and thus, would be classified as not having CP in clinical trials that use a dichotomous outcome (yes vs. no) for the disease, such as the trials described earlier. The same could be said about the disturbances of sensation, perception, cognition, communication and behaviour that may or may not accompany the primary activity limitation, although their occurrence is not required for a diagnosis of CP. Where these 'cofactors' are important to our examination will relate to how one determines a reduction in severity as a result of ACS.

The favoured method for evaluating the effect of ACS on reducing CP severity is a prospective clinical study, preferably a randomized controlled trial. One would compare the severity of CP in children born by mothers who did not receive ACS with that among those born by mothers who did receive ACS. One could measure disease severity using a scale, such as the Gross Motor Function Classification System [13,14]. However, it is not clear how one integrates changes in the cofactors when evaluating reduced severity, as these factors are not part of the scale. For example, one could imagine a scenario where the ACS group shows an overall improvement in activity limitation, but has a higher prevalence of each cofactor. Unfortunately, the medical literature does not offer empirical support for the position that CP severity would be reduced, as no such study has been reported in the context of ACS. The studies that examined the relationship between ACS and CP focused specifically on the prevalence of its diagnosis (as judged by the authors of each study), rather than its severity.

Despite the lack of empirical support for an observed direct effect on severity by ACS, it was nonetheless argued by the plaintiff's expert witnesses to be the case. The argument is made as follows: (1) CP is a condition that exists on a continuum, (2) steroids act to stabilize cell membranes, making them more resilient to damage, (3) it is accepted in the literature that ACS will reduce the severity of respiratory disease, chronic lung disease and other short term outcomes and (4) improvement on these outcomes, coupled with more resilient cell structures, reduces the risk of known risk factors for CP, such as intraventricular haemorrhage and periventricular leukomalacia,[5] therefore, (5) as infants born to mothers receiving ACS on average show reduced severity in those negative health outcomes that have been the target of previous clinical investigation, and as such infants

[5]It is noteworthy that neither of these conditions was present in the plaintiff's children.

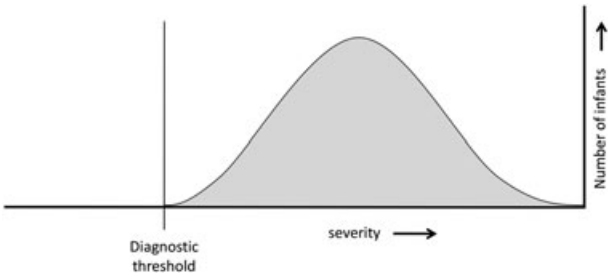

Figure 1 Theoretical distribution of cerebral palsy severity.

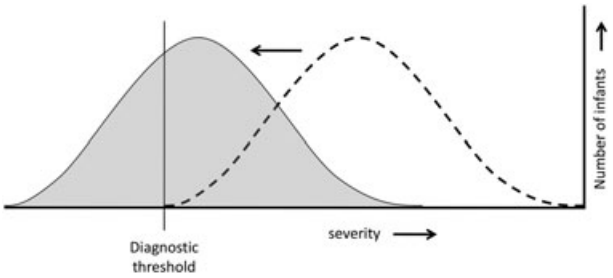

Figure 2 Shift in the distribution if cerebral palsy severity is reduced equally for all infants.

are generally in better health (relative to ~ACS infants), there is no reason to believe that CP symptoms would not improve as well. The presented position is strengthened by the fact that it also provides an explanation for why CP is shown to be less likely to occur when ACS is administered (i.e. it provides biological plausibility to the suggestion that the administration of ACS has efficacy with respect to diminishing the risk of CP [15]), but note that the inference is potentially biased, as there are no objective criteria to rule out some other, perhaps equally plausible explanation (e.g. the relationship between ACS and cerebral palsy is stochastic in nature – that is, risk is unrelated to disease severity[6]).

The model we choose to explain how ACS affects CP has significant implications on what support we have for the belief that severity of the disorder would be reduced in the event that it was not avoided. Let us assume that among those diagnosed with CP (had ACS not been administered), a measure of the disorder severity follows a normal distribution (Fig. 1). That is, there are few individuals showing low or high severity, and a large frequency of individuals showing moderate levels of severity. At the extreme low end of the distribution is the threshold for clinical significance (i.e. the point at which, above, one is diagnosed with CP and below, CP is deemed absent). If we believe that ACS reduces the severity of CP (the opinion of the plaintiff's expert witness), one would expect a shift in this distribution in the direction towards

[6]It is possible that the mechanism that promotes the risk of cerebral palsy is independent of the severity of the disease in the event that it occurs. That is, reduced risk does not entail reduced severity. For example, the probability that cancer may occur as a result of exposure to ionizing radiation is thought to increase as the radiation dose to the tissue increases, and yet, the severity of that cancer is not necessarily related to the radiation dose (e.g. [17]).

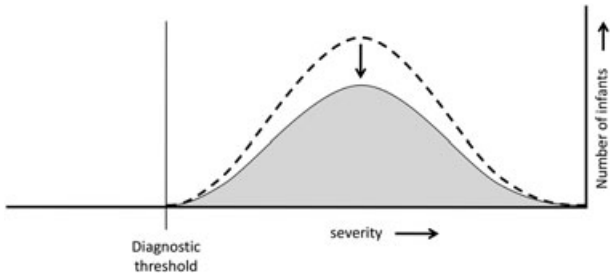

Figure 3 Potential change in the distribution if the probability of avoiding cerebral palsy is the same for all infants.

the threshold for clinical significance (Fig. 2). Using this model, those on the low end of the distribution of severity would be 'cured' of the disorder had ACS been administered (explaining the risk reduction due to ACS), and all other individuals would see their severity reduced in a linear fashion. An alternative model might posit that the mechanism that reduces the probability of CP in an infant is independent of the mechanism that determines the severity of the disorder. There is no expectation that ACS should influence both of these mechanisms. Thus, the probability that CP would be avoided given ACS is equal for all individuals on the severity distribution, causing a 'flattening of the curve' (Fig. 3). Using the alternative model, CP would be avoided in some individuals, whereas all the other individuals would not realize any changes in the severity of the disorder. In the absence of empirical data, any number of additional models is also plausible. For example, those on the extreme high end of the distribution might see a reduction in severity, with diminishing reductions as one moves towards the threshold (a non-linear relationship), or some will see a reduction in severity, whereas others will not, etc. What matters in the context of the legal case we are examining here is that the probability any individual would be from the group where CP would be avoided or severity reduced exceeds that where CP is not avoided or severity is not reduced. Only the model advanced by the plaintiff's expert witness guarantees this is the case.

What this case teaches us about when risk estimates are 'evidence' of causality

The *Goodman* case raises the issue of whether or not information of risk can substitute for evidence of causation [9]. Most notable is the extent to which the probabilistic association between an event (failure to administer ACS in a timely manner) and an outcome (CP) is evidence that had the event not occurred in the described scenario the outcome certainly would not have occurred. The debate in the case presented here primarily centred on statistical thresholds – minimizing the probability of falsely rejecting the null hypothesis, and achieving a $RR > 2$. Thus, it seems that achieving these statistical thresholds to some extent plays a role in establishing causality of the particular. Thresholds play an important part in determining evidence of causality in clinical medicine as well, specifically with respect to achieving a $P < 0.05$. However, the burden of a $RR > 2$ is not part of the determination of evidence (of effectiveness or causality) in a clinical setting, in part because the effects of many therapies or contributions of individual factors on some outcome are often small.

While much of the debate in *Goodman* focused on translating statistical risk from a trial population to an individual case, and establishing that statistical thresholds were met, it was noted by the judge that this is not the only important factor in determining causality:

> This statistical information is of course just one piece of evidence the court must consider in determining the issue of causation. Statistics represent generalizations and not the particular circumstances of a given case. At best, they are numbers which are dependent on the underlying data and the limits of scientific knowledge. Context, human experience and interpretation are required if statistics are to have any real meaning. ([4]; p.32, paragraph 206).

Recognizing that thresholds were not met, the plaintiff's expert witnesses invoked mechanistic arguments to show that the purported connection between the exposure and outcome was plausible – a strategy that seemed to convince some of the trial's key stakeholders. Again, we see some similarity with the clinical medicine community – what Hill (and others in the epidemiology community) would describe as 'biological plausibility' [1,15] is an important consideration in determining if risk estimates are evidence of causality in clinical medicine settings. Another important consideration in both the legal and clinical medicine contexts is that the data is acquired using a high quality methodology (e.g. randomized controlled trials).[7]

On the basis of the presented legal case, the process of translating risk information from a population to evidence of causation for an individual appears to require that (1) a statistical association between the exposure and outcome meets some pre-specified threshold(s) (e.g. RR $>$ 2 and the probability of a spurious result is sufficiently low), (2) data on population level risk is acquired using high quality research methods and (3) a plausible mechanism for why the exposure causes the outcome be established. Compare this with the standard model of prediction in medicine, as described by Fuller and Flores [16], which illustrates how physicians translate risk information from a population to individuals for purposes of therapeutic decisions and prognosis. The 'Risk Generalization-Particularization Model' is a two-step process, whereby the physician generalizes a study population derived risk measure to a target population containing the patient of interest, and then transforms this risk measure to yield a probability of benefit for a particular patient from the target population [16]. On the most basic level, both processes permit the assumption that what is established at the population level is true for the individual (a simple approach to translating risk). A notable difference is that the Risk Generalization-Particularization Model does not explicitly consider risk thresholds, such as the balance of probabilities principle (i.e. RR $>$ 2) in tort cases; rather risks are translated using the same process, irrespective of their observed magnitude. This is due to the fact that (1) therapeutic effects and the contribution of a particular exposure on outcomes are quite small, and thus, RR $>$ 2 is often not observed and (2) whether or not a RR $>$ 2 threshold is met is not informative in predicting the probability that the exposure will cause the outcome (for this one would use the risk difference, rather than the RR). With respect to the second issue, consider a circumstance where the base risk is extremely low (e.g. 1/million). Even where the risk would slightly more than double (i.e. RR $>$ 2), there would be no appreciable change in (perceived) risk to the individual (i.e. risk of 1/million $\rightarrow$ 2/million) such that management decisions with respect to care would likely change, especially considering other factors (e.g. side effects and financial burden). Whereas in the medical context, the RR $>$ 2 would likely have no bearing on the choice of management, any party or agent responsible for raising this risk (despite it size in absolute terms) would be liable in a legal context under the principle that the balance of probabilities was met, and thus, it would be established that the party or agent materially contributed to the occurrence of the outcome in question.

Even if it were shown that an increase in risk was established, one cannot definitely state that the individual is at the same risk, or that the occurrence of the outcome was even attributable to the established risk factor. Rothman and colleagues [1] recognize both that the cause of any negative health outcome is multifaceted, and that the occurrence of a particular risk factor is itself not sufficient; rather, there is a minimal set of conditions that must be met in order for a specified outcome to occur. For example, the fact that CP was present in a number of children whether or not ACS was administered suggests there are causes or conditions that must be met for CP to occur beyond ACS administration. Trial evidence does not provide information on this minimal set of conditions, and Bayesian calculations are no substitute for this lack of information. At best, trial evidence can provide us with information that some exposure is associated with the outcome, but this does not itself guarantee that where that exposure is present, the outcome will occur. The case presented here demonstrates some of the pitfalls that arise when one attempts to fill in knowledge 'gaps' related to the unknown conditions. Ultimately, personal values may be the arbiter of which information takes precedence in establishing causality in a given context, as was evident in how the plaintiff and defendant each gave primacy to different features of the presented data.

[7]Advocates of evidence-based medicine give primacy to data acquired using randomized trial methods and consider information using other methods as 'weaker' forms of evidence. Likewise, it was stated in the judge's summary of the ruling in *Goodman* that 'almost all of the expert witnesses who testified agreed that the Cochrane meta-analysis on steroids was the best available evidence and will remain so as no further randomized clinical trials on a single course of antenatal corticosteroids will ever be undertaken', suggesting that quality of methodology was considered in the trial debate ([4]; p.29, paragraph 179).

Acknowledgements

We would like to express our gratitude to both Nathan Schachtman and the anonymous reviewer for comments on an earlier draft that greatly improved this paper.

References

1. Rothman, K. J., Greenland, S., Poole, C. & Lash, T. L. (2008) Causation and Causal Inference. In Modern Epidemiology (eds K. J. Rothman, S. Greenland & T. L. Lash), 3rd edn, pp. 5–31. Philadelphia: Lippincott Williams & Wilkins.
2. Hillel, D. W., McCague, P. & Yaniszewski, P. F. (2005) Proving causation where the but for test is unworkable. *The Advocates' Quarterly*, 30, 216–238.

3. Snell v. Farrell (1990) 2 S.C.R. 311. Retrieved from: https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/634/index.do, on March 24, 2016.
4. Goodman v. Viljoen (2011) ONSC 821 (CanLII), aff'd, 2012 ONCA 896 (CanLII). Retrieved from: http://www.canlii.org/en/on/onsc/doc/2011/2011onsc821/2011onsc821.html, on March 24, 2015.
5. Miracle, X., Di Renzo, G. C., Stark, A., Fanaroff, A., Carbonell-Estrany, X. & Saling, E. (2008) Guideline for the use of antenatal corticosteroids for fetal maturation. *Journal of Perinatal Medicine*, 36, 191–196.
6. Wapner, R. J., Sorokin, Y., Mele, L., *et al.* (2007) Long-term outcomes after repeat doses of antenatal corticosteroids. *New England Journal of Medicine*, 357, 1190–1198.
7. Crowley, P. (2006) Prophylactic corticosteroids for preterm birth. *Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD000065. doi: 10.1002/14651858.CD000065.pub2
8. Roberts, D. & Dalziel, S. R. (2006) Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. *Cochrane Database of Systematic Reviews*, Issue 3. Art No.: CD004454. doi: 10.1002/14651858.CD004454.pub2
9. Schachtman, N. A. (2014 March 15) A black swan case – Bayesian analysis on medical causation. Retrieved from: http://schachtmanlaw.com/2014/03/ on March 24, 2016.
10. Broadbent, A. (2013) Philosophy of Epidemiology. New York: Palgrave MacMillan.
11. Federal Judicial Committee (2011) Federal Judicial Center Committee on the Development of the Third Edition of the Reference Manual on Scientific Evidence, and the Committee on Science, Technology, and Law Policy and Global Affairs. Reference Manual on Scientific Evidence (Third Edition). Washington, The National Academies Press.
12. Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M. & Bax, M. (2007) A report: the definition and classification of cerebral palsy: April 2006. *Developmental Medicine and Child Neurology*, 109, 8–14.
13. Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E. & Galuppi, B. (1997) Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine & Child Neurology*, 39, 214–223.
14. Rosenbaum, P. L., Walter, S. D., Hanna, S. E., *et al.* (2002) Prognosis for gross motor function in cerebral palsy: creation of motor development curves. *Journal of the American Medical Association*, 268 (11), 1357–1363.
15. Hill, A. B. (1965) The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, 58, 295–300.
16. Fuller, J. & Flores, L. J. (2015) The risk GP model: the standard model of prediction in medicine. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 54, 49–61.
17. Committee to Assess Health Risks From Exposure to Low Levels of Ionizing Radiation, Nuclear Radiation Studies Board, Division on Earth Life Studies, National Research Council of the National Academies. Health Risks From Exposure to Low Levels of Ionizing Radiation: BEIR VII Phase 2. Washington, DC, National Academies Press, 2006.

