The UK civil law approach to epidemiology and statistical evidence

Dr Amy L Wilson, Professor Jane Hutton, Leigh-Ann Mulcahy QC and Dr Claire McIvor

1. Introduction

This document summarises discussions held at a colloquium on the UK civil law approach to epidemiology and statistical evidence. This colloquium was organised by the Statistics and Law section of the Royal Statistical Society and held at Fountain Court Chambers on the afternoon of 11 June 2018. There was a wide range of attendees from academia, the legal profession and industry with expertise in epidemiology, statistics and legal epidemiology.

The event was split into five main presentations covering the views of a legal practitioner, a judge, a legal academic, a statistician and an epidemiologist and was followed by a discussion. The aim of the event was to focus on finding what the main issues and questions are that face the legal profession regarding epidemiology and statistical evidence. Further events can then be focussed on finding solutions to these issues.

The authors wish to thank all those who attended and contributed to the discussion as well as the Royal Statistical Society and Fountain Court Chambers for assisting in the organisation and hosting of the event.

2. Lies, damned lies and statistics: a legal practitioner's perspective
Leigh-Ann Mulcahy QC

Ms Mulcahy QC summarised some key issues that have arisen in cases concerning statistical and epidemiological evidence. She spoke about:

- **Loveday v Renton [1990]**. This case questioned whether a pertussis vaccination can cause permanent brain damage or death in young children. The claim in this case failed on the basis causation was not established - the absence of epidemiological evidence in support of the case was regarded as "of great moment" and it was noted by the judge that there was no generally accepted standard of scientific proof.
- **Hope and Reay v British Nuclear Fuels [1994].** This case focussed on whether epidemiological evidence could show that paternal pre-conception irradiation could cause leukaemia (ALL and NHL). The claim failed on the balance of probabilities, with no evidence of an association shown. The judge applied the Bradford-Hill criteria and said that these criteria were an attempt to systemise common sense.
- **Gregg v Scott [2005].** This case concerned medical negligence in a diagnosis of NHL. The negligence was purported to have reduced the patient's chance of survival from 42% to 25%. The case was rejected because 42% is less than 50% and so the original chance of survival was less than half (and hence rounded to zero). Questions arose as to whether you can have damages for the loss of a chance of recovery, but the House of Lords rejected this approach in medical negligence cases
- **XYZ v Schering [2002].** This case looked at whether harm had come from use of a third generation combined oral contraceptive (COC) following a warning from the Committee on Safety of Medicines (CSM) that this COC caused a higher risk of venous thrombo-embolism.

Various epidemiological studies were presented with mixed conclusions. The judge decided that a Cox regression analysis presented the most compelling evidence, with a relative risk of 1.7. With this relative risk being less than 2 ("doubling the risk"), the claim was rejected. This result led to questions about the use of Cox regression analysis in this context and whether doubling the risk is an acceptable standard to use.

- **McTear v Imperial Tobacco [2005].** The claimant sued because of her husband's death from lung cancer but the claim failed both because epidemiological evidence on its own could not be used to prove that tobacco *did* in fact cause her husband's lung cancer but also because the judge was not satisfied that lung cancer could generally be caused by tobacco,, which was surprising because it is generally regarded by the UK and US Governments and the WHO that there is an overwhelming link between the two.
- **Sienkiewicz v Greif [2011].** This case concerned whether a case of mesothelioma had arisen from occupational or background exposure to asbestos. The judge derived their own estimate that the relative risk was 18% (occupational vs background) and rejected the claim on the basis that the occupational exposure had not doubled the risk. The Supreme Court later applied the "Fairchild exception" and said that the exposure had materially increased the risk and so the company was liable. This case resulted in an *obiter* discussion and general negative view of epidemiological evidence from the justices but in this case no epidemiological evidence had been called and it appeared that statistics was being conflated with epidemiology.
- **Heneghan v Manchester Dry Docks [2016]**. This case concerned causation of asbestos related lung-cancer. The Court approved an approach that causation is proved where a tortfeasor has more than doubled the risk of injury where science does not permit determination with certainty as to how injury is caused.
- **Gee v Depuy [2018].** This case concerned whether a hip replacement had caused soft tissue damage. Statistical evidence was presented deriving from the National Joint Registry. The court accepted Depuy's arguments that the statistics were unreliable and subject to limitations and confounding factors.

Ms Mulcahy QC highlighted the schizophrenic approach between the caution expressed about use of epidemiological evidence in *Sienkiewicz* compared with the fact that now the doubling of the risk test is becoming a means for proving factual causation without more (*Heneghan*). The Bradford-Hill criteria (BHC) seem to have been forgotten. It is not clear what the legal position on the relevance or significance of the BHC is now.

She noted that there are differences between law and science in relation to the standard of proof. It is perceived by lawyers that the scientific standard of proof is much higher than that found in the law and that the scientific standard of proof is "95% certainty" i.e. equating it with the 95% confidence limit or p-value of >0.05. In civil law, the standard of proof is the balance of probability, i.e. 50%+. There are also differences in terminology between the two disciplines. For example, in legal circles the term "risk" means a chance of an adverse outcome whereas statisticians equate it to the probability of an event, good or bad and "significance" means important whereas statisticians would attach a very specific scientific meaning to this term. The Royal Statistical Society guide [1] is a first step in dealing with differences in terminology.

To conclude, Ms Mulcahy QC listed the following key issues:

- Lawyers have different views on the meaning of risk. Important with causation tests like "materially increase in risk". Do different groups (e.g. lawyers, epidemiologists, judges) mean different things?
- The current approach towards epidemiology is unclear. There was negativity after Sienkiewicz v Greif but on the other hand, it is said that doubling the risk is enough.
- Can the balance of probabilities in the legal sense be mapped onto statistical evidence?
- The relative risk is currently used in court. What relevance does absolute risk have?
- Is there a defined scientific standard of proof for causation?
- What evidence is required to show causation and who should be presenting evidence in these cases – a statistician, a clinician or an epidemiologist? Courts appear happy to accept the evidence whoever it is from although cost is a factor and so more guidance is needed about when expert evidence is necessary. Should the admissibility of statistical evidence be considered pre-trial (cf. the US *Daubert* procedure) ?
- An interdisciplinary approach is needed to solve these issues. How can lawyers and judges become better educated in statistical and epidemiological matters?

3. A judge's view of causation
Mr Justice Jeremy Stuart-Smith

Mr Justice Stuart-Smith started by outlining the role of a judge. He said that judges are not specialists. They are presented with a range of different views (e.g. from a lawyer, statistician or epidemiologist) and are required to get the right answer (unlike litigators). A judge is concerned with the allocation of responsibility, after something has gone wrong. It is important that judges recognise that they can only have a tenuous grip on complicated topics such as statistics and risk. He added that a judge should write judgements for the loser so that they can understand why they have lost.

Mr Justice Stuart-Smith then discussed several key questions and points that arise for a judge when confronted with epidemiological evidence:

- Sienkiewicz exemplifies the risk of tampering with the "but for" test.
- When can conclusions be based "on the basis of epidemiological evidence alone"? When is epidemiological evidence sufficient to show causation?
- There is a risk of false precision when working with statistical evidence. For example, a 51% balance of probability does not make sense – the probability cannot be specified with that much accuracy.
- When is it acceptable to move away from proof on the balance of probabilities? If the defendant's conduct cannot be shown on the balance of probability to have made a difference,

---

[1] http://www.rss.org.uk/RSS/Influencing_Change/Statistics_and_the_law/Advocates_guide/RSS/Influencing_Change/Current_projects_sub/Advocates_guide.aspx?hkey=883603a7-fc93-4921-a2cc-36ac14e1cf82

why should they be liable? Two possibilities were discussed: 1) where proof of the balance of probability is impossible, and 2) when there is over-determination on the basis of multiple insufficient causes. An example of this latter possibility would be if five people had participated in an act where just three would be sufficient for damage. Are all five people acting tortiuously or are three people acting tortiuously and two lawfully? (See Stapleton below for a discussion of the same point)

- Can you have the loss of a chance? Chaplin v Hicks [1911] concerned a beauty contest in which acting jobs would be given to 12 winning entrants. The claimant was one of 50 contestants to make it through to the final interview stage but missed out because the defendant wrongfully failed to inform her of the date of the interview. The claimant successfully argued that she had been denied the chance of winning and was awarded £100 damages. It is important to note that her claim was in contract and that as one of fifty candidates for twelve jobs, she had stood an objectively strong chance of winning. In tort, the case of Gregg v Scott (see discussion above by Mulcahy) makes it clear that there is no recovery for a loss of a chance of avoiding personal injury.

Mr Justice Stuart-Smith concluded that to win on causation a lawyer should stick to the "but for" test. Epidemiology should not be overplayed, and the lawyer should be careful to check any possible counterfactuals.

4. A legal academic's view
Professor Jane Stapleton

Professor Stapleton began by highlighting the need to adjust the "but for" test to account for the problem of multiple insufficient causes. She noted the example of several people pushing a car over a cliff. Each person alone would not be strong enough, but there might have been more people than necessary. Allocating responsibility is a challenge[2].

Professor Stapleton then discussed the case of Heneghan v Manchester Dry Docks [2016]. In this case the claimant had died from lung cancer and had been exposed to asbestos by multiple employers. There were four possible mechanisms that might have caused the cancer: background sources, tobacco, asbestos, or tobacco and asbestos acting synergistically. Professor Stapleton focussed on two key epidemiological questions asked in Heneghan:

1) Was an asbestos mechanism likely involved in the claimant's lung cancer? (general causation)
2) Were the asbestos exposures linked to the individual defendant likely involved in the victim's lung cancer? (individual causation)

She considered whether these questions could be validly answered by statistical comparisons (validity), whether the available data were reliable and relevant (probity) and whether the law should accept statistical comparisons based on the data (judicial policy).

---

2 Stapleton, J., Unnecessary Causes, Law Quarterly Review, Jan 2013, 39-65

She argued that the statistical comparisons could be used to answer Q1 in the affirmative because estimates of the effects of the four mechanisms were available[3]. The asbestos and tobacco smoke would have interacted synergistically, thus making the causative potential of the asbestos even greater.

By contrast, the statistical comparisons could not validly be used to answer Q2. This is because it is a much more specific question than Q1 and requires far more detailed information about the aetiology of lung cancer than currently exists. In Heneghan, the Court of Appeal used the policy-based Fairchild test to answer Q2 in favour of the claimant.

Professor Stapleton stressed that statistical evidence should be presented only by statisticians and that the Sally Clark case is a clear example of why this is the case. As well as presenting numerical results, it is crucial that statisticians give evidence on the validity and probity of the evidence.

In the discussion that followed, it was said that the law is wrong to label ordinary arithmetic as statistical or epidemiological evidence. In Heneghan the statistical evidence was just arithmetic. The right experts are needed so that evidence is presented on a more defensible basis.

5. A statistician's view

Professor Jane Hutton

Professor Hutton began by describing the difference between arithmetic and statistical analysis. A statistical analysis is concerned with the understanding of uncertainty. Arithmetic uses numbers presented with no consideration of uncertainty.

Professor Hutton emphasised the importance of using the right experts to present evidence. She then discussed several key issues that arise for statistical expert witnesses:

- There are gaps in understanding on both sides so the first thing that must be done is to get a consensus on what the question is. It is important to ask the right questions and clarify all points.
- It is important to know what data are available and whether this data will be available directly to the expert or through reports or academic articles.
- How much data is available? Big data is not always the best sort of data. A small amount of biased data can be slightly misleading whereas a lot of biased data can be very misleading.
- Are the data good quality? Are there missing data? If so, this must be examined and explained.
- There is often a lot of material provided in cases. The expert will not be able to read it all so it is important that he/she is able to isolate material that is most relevant.
- Are there any alternative explanations? For example, if the claimant has cancer and is suing an employer for asbestos exposure, what other possible causes are there?
- How should the data be summarised? For example, the lawyer might require the relative risk to show that the risk has been doubled. Is there agreement on this summary statistic? An example linking blood pressure with risk of stroke was shown to demonstrate that the relative risk might not always be the most appropriate statistic. For young people, the relative risk is doubled with increasing blood pressure more quickly than for older people. But younger people have a much smaller absolute risk of stroke than older people.

---

[3] Contrast *Sienkiewicz* v *Greif Ltd* [2011] UKSC 10 see Stapleton, J., Factual Causation Mesothelioma and Statistical Validity, Law Quarterly Review, April 2012, 221-231

- Terminology differs between different groups. For example, the terms “independence”, mutually exclusive” and “risk” all have different meanings in the legal profession when compared to the statistical definition.
- Legal expertise is needed to identify the most interesting questions. Statisticians will not be familiar with different cases, especially in the lower courts.

Professor Hutton concluded with a quotation by Ronald Fisher to highlight the importance of consulting statisticians well in advance:

“To consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of.”

6. Challenges in proving causation - an epidemiological perspective
Professor Alan Silman

Professor Silman outlined the difficulties faced by an epidemiologist when asked about causation. He noted that there are various possibilities for causation in linking a specific hazard to a specific health outcome:

- **Necessary and sufficient.** The hazard is both necessary and sufficient for the outcome. For example, radiation exposure above a certain level will always result in radiation damage and radiation damage must have arisen from radiation exposure. In this case, there is no need for epidemiology.
- **Necessary but not sufficient.** The outcome must have arisen from exposure to the hazard, but such exposure does not necessarily lead to the outcome. An example is women who took thalidomide in pregnancy. Only women who took thalidomide gave birth to babies with the very specific limb damage, but many women took the drug with no adverse effects.
- **Sufficient but not necessary.** The hazard will always lead to the outcome, but the outcome might not have arisen from exposure to the hazard. For example, overwhelming radiation makes the risk of leukaemia in those who survive is extremely high, but the majority of patients who get leukaemia were not exposed to such radiation.
- **Neither necessary nor sufficient.** The hazard only infrequently leads to the outcome and the outcome only infrequently may have arisen from the mechanism. This does not mean that in those who are exposed, that exposure was a risk factor. This is the most common situation faced by epidemiologists, justifying the complex studies that need to be done. An example is the link between the vaccine Pandemrix and narcolepsy. Most people who received the vaccine do not develop narcolepsy (a rare disorder) and most patients with narcolepsy have not been vaccinated with Pandemrix. Thus even in those patients with narcolepsy that have been vaccinated, on its own that observation does not imply a relationship even in that patient

Professor Silman highlighted the need to understand the difference between a cause and a risk factor. It may be possible to show an association between some mechanism and an outcome to demonstrate that the mechanism is a risk factor but to show that the mechanism causes the outcome requires an independent experiment. Normally it is not possible to perform such an experiment in the time available.

Four key challenges in epidemiology were then discussed in relation to showing a relationship between a possible hazard and an adverse disease outcome:

- **Variation/ complexity.** Medical issues can be very complex, with variation between patients in several factors, including the severity of the disorder, the therapy and health care received for the disorder of interest and other disorders as well as many background demographic and lifestyle factors. It can be very difficult to identify the distinct influence of one specific risk factor
- **Noise/ random error.** If the disorder, and/or the intervention in question is rare then the inference from small samples will lead to imprecise estimates of an effect. Similarly, if the methods used to measure the outcome are subject to error, it is difficult to prove an effect. The importance is that the data may not be biased in the way they were collected, but that the random error makes it much harder to demonstrate a true risk if it existed
- **Bias, or systematic deviation from the truth.** This is different from the above. For example, the chance that a person participates in a study might increase if the person has the outcome being studied. Or the retention rate of participants in a study might be higher if they are benefitting from the study. Bias does not necessarily have to be a problem if it is well understood. For example, if a bias makes it more likely to see a difference between groups in the study and the results still show that there is no difference between the groups.
- **Confounding, or results that can be explained by intermediate variables.** For example, in a study on the effect that asbestos has on lung cancer, whether the participants are smokers might be a confounding variable. This can be one of the most difficult aspects of epidemiological evidence. In any study it is unlikely that all the potential confounders will be understood. The question is to what extent the results of the study can be explained by confounding. There can also be complex interactions between confounders, e.g. smoking in males might have a different effect to smoking in females.

To conclude, Professor Silman urged that we should not expect a perfect epidemiological study - it is often a case of doing the best job possible with the data that are available. In this situation, it is important to understand the imperfections in the study and how these imperfections might have impacted the results.

7. Discussion

Following the formal presentations there was an open discussion session. The following key points were made during this discussion:

- **Doubling the risk.** There seems to be a misconception amongst lawyers that epidemiologists treat “doubling the risk” as proof of causation. In practice, this is not the case. For example, in some large genetic studies a relative risk of 1.1 will be treated as statistically significant (though not in itself determinative of causation). Eliminating one gene might eliminate the illness, but this gene might be one of many that are required for the disease to be present. There is no logic to using doubling the risk as a test for causality and relative risk assessments will not necessarily be relevant to every tort claim involving causal uncertainty. The problem is that doubling the risk is now in vogue. The court is trying to use it as a proxy for certainty.
  If “doubling the risk” is not appropriate, what is the alternative? It was suggested that a clearly defined standard is needed so that claimants know in advance of paying for legal representation what the requirement is. However there is no rationale in law or in science for using set numerical benchmarks to establish causation in tort law. Lawyers should instead try to make the best argument possible for their client with the data available.
- **Uncertainty.** From a legal perspective, it can be difficult to use statistical and epidemiological experts as they tend to highlight all the weaknesses in any analysis (this is generally only good for

the defendant). It can also be difficult to know at the start of a case that a statistical or epidemiological expert will be needed.
From a statistical perspective, the problem is that it is usually not possible to find "a relative risk". There will be uncertainty in estimates and it may be necessary to present a range of figures. If information is plausible and reliable (although uncertain), why should this not be used to make a decision? Does the court need to become more comfortable with uncertainty?

- **Pooling of evidence.** In science it is common to pool evidence across different studies to reduce uncertainty (and get a more precise estimate of the relative risk). This is useful for studies of e.g. a rare drug. The problem is that different studies might be addressing different questions or have different types of data. How can studies be pooled if they are of different qualities, or if studies are flawed in different ways? In answer to this, any good epidemiologist carrying out such a meta-analysis (or systematic review) should be able to apply recognised methodologies addressing such issues.
- **Objectivity.** Experts can become wedded to a cause and lose objectivity. This can lead to the court distrusting evidence in a particular area.

The biggest issues arise when the court is dealing with factual causation, making this a good place to begin to identify what needs to change. To work towards identifying the relevant questions and solutions requires an interdisciplinary approach. The problems discussed have been very distinct in nature and therefore need input from those with a background in law, statistics and epidemiology. Going forward, the statistics and law group will identify the most pressing individual problems, draw up specific projects and allocate to separate working groups.