Was It Wrong to Use Statistics in *R v Clark*? A Case Study of the Use of Statistical Evidence in Criminal Courts

Amit Pundik

Abstract This chapter discusses the use of statistical evidence to prove the material fact of causation in criminal courts. It focuses on *R v Clark*, in which a mother was wrongfully convicted of murdering both her babies. In order to disprove a potential defence claim that the babies died of SIDS (aka cot death), the prosecution adduced statistics that allegedly showed that the probability of two SIDS deaths in a family similar to the Clarks was 1 in 73 million. This chapter considers the question of whether it was wrong to use such statistical evidence in *Clark*. Four common explanations of why it was wrong, each of which attributes the wrongful convictions to the use or misuse of the statistical evidence, are scrutinised and rejected. However, drawing on the theory of contrastive explanation, it is argued that it was still wrong *in principle* to use the SIDS statistics in Clark, because using them properly would require another piece of evidence which is clearly objectionable: statistical evidence on the rate of smothering among mothers who are similar to Clark. Regardless of whether the exercise of comparing probabilities of SIDS and smothering is feasible, such an exercise should not be conducted as part of criminal proceedings. This chapter thus concludes that *Clark* should serve as a warning against any attempt to prove the fact of causation using statistical evidence about the rate of potential exonerating causes.

I would like to thank Mike Redmayne for his helpful and constructive comments on an early version of this chapter. I would also like to thank Asaf Wiener and Maya Ben-Meir, who have provided me with excellent research assistance and helpful comments. The research leading to these results received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement n° 299653.

A. Pundik (✉)
Buchmann Faculty of Law, Tel Aviv University, Tel Aviv, Israel
e-mail: amitp@post.tau.ac.il

F. Zenker (ed.), *Bayesian Argumentation: The practical side of probability*,
Synthese Library 362, DOI 10.1007/978-94-007-5357-0_5,

1 Introduction

Sally Clark's two baby boys, Christopher and Harry, were found dead on separate occasions.[1] Christopher's death was at first treated as sudden infant death syndrome ('SIDS', also known as cot death).[2] However, when Harry died two years later, the autopsy revealed suspicious injuries and the findings from Christopher's autopsy were re-evaluated. Clark was then charged with and convicted of the murder of both babies. At the trial, the prosecution called Professor Meadow, an expert paediatrician, to counter a potential defence claim that both deaths were natural as due to SIDS. In his testimony, Professor Meadow said that the probability of two occurrences of SIDS in a family similar to the Clarks was 1 in 73 million. This calculation was found to be flawed, but the Court of Appeal upheld the two murder convictions after finding the case against Clark still 'overwhelming' on each account (*Clark-I*: [254], [272]). A few years later, Clark's husband found evidence in the hospital archives of microbiological results indicating that Harry had died from natural causes. A second appeal was allowed, and Clark was set free on 29 January 2003 after serving more than three years in prison.[3] Sadly, four years later, she was found dead in her home as a result of alcohol poisoning (BBC 2007).

While the wrongful convictions[4] of Clark raise various issues,[5] much of the public and scholarly attention given to this case focused on the use of the SIDS statistics (the probability of two cases of SIDS in the same family).[6] The SIDS statistics were commonly regarded in the public media as somehow responsible, at least in part, for the wrongful convictions of Clark.[7] One newspaper even went

[1] *R v Clark* (Crown Court Chester 9 November 1999), R v Clark (No 1) [2000] EWCA Crim 54 (hereafter '*Clark-I*'), *R v Clark (No 2)* [2003] EWCA Crim 1020 (hereafter '*Clark-II*')

[2] SIDS (sudden infant death syndrome) is defined as 'the sudden death of a baby that is unexpected by history and in whom a thorough necropsy examination fails to demonstrate an adequate cause of death' (*Clark-I*: [104]).

[3] The convictions were quashed and Clark was released after the Crown decided not to seek a retrial (*Clark-II*: [5]). The Court of Appeal commended this decision (*ibid*.: [181]).

[4] The term 'wrongful convictions' is used because it is assumed that the admissible evidence available today, including the microbiological results found later by the husband, is insufficient to prove beyond reasonable doubt that Clark murdered Christopher and Harry (this assumption is made explicitly in the second appeal, *Clark-II*: [179]). It should be noted that this assumption is weaker than the claim that Clark is in fact innocent (though this is also probably true), but none of the points made in this chapter requires the stronger claim of actual innocence.

[5] For example, the duties of expert witnesses repeatedly received scholarly attention. See Wilson (2005), Dwyer (2003), Blom-Cooper (2006).

[6] For the centrality of the SIDS statistics in the discussion of *Clark*, see, for example, Editorial (2000) and Nobles and Schiff (2005).

[7] 'The statistic was quoted in every headline and is widely believed to have led to Sally Clark's conviction' (Barraclough 2004); '[t]he jury at Sally Clark's trial, however, was apparently persuaded by the evidence of a leading expert called by the prosecution, Professor Sir Roy Meadow, who maintained that the probability of two cot deaths occurring in a single affluent family was "one in 73 million"' (Telegraph 2007).

as far as to assert that the statistical evidence was almost the only evidence against Clark.[8] The existence of some causal connection between the (mis)use of the SIDS statistics and the wrongful convictions not only appeared in the public media but was also hinted at by expert statisticians, such as Professor Donnelly[9] and Professor Dawid.[10]

This chapter examines whether it was wrong to use the SIDS statistics in *Clark*, and if so, why. After providing a more detailed factual description of *Clark* in Section 2, in Section 3 the chapter scrutinises and rejects four common explanations, each of which draws a connection, either explicitly or implicitly, between the SIDS statistics and the wrongful convictions. The first explanation concentrates on the flaws in the calculation made by Meadow that led to the 1 in 73 million figure. The second holds that the SIDS statistics had an overwhelming psychological effect which overshadowed other evidence more favourable to Clark. The third explanation focuses on a logical mistake called the 'prosecutor's fallacy' (explained below). The last explanation considers the courts' refusal to use Bayes' Theorem (also explained below). This chapter seeks to show that each of these explanations suffers from its own specific weaknesses. More generally, it is shown that these explanations are unpersuasive, taken either separately or jointly, since any connection between the SIDS statistics and the wrongful convictions is hard to establish. This is because it is difficult to accept that the wrongful convictions could have been prevented if the SIDS statistics had been used correctly, or even if they had not been used at all.

Section 4 suggests an alternative explanation of why it was wrong to use the SIDS statistics in *Clark*. Drawing on the theory of contrastive explanation from the philosophy of science (van Fraassen 1980: 97–157; Lipton 1990), this chapter shows that the use of SIDS statistics in *Clark* was wrong *in principle*. This is because the only way to make the SIDS statistics meaningful is to compare them, even implicitly, to another piece of statistical evidence: in this case, the rate of mothers who smothered to death both of their babies among the population of mothers who are similar to Clark. Using this second piece of statistical evidence is clearly objectionable. Hence, even if all the difficulties highlighted by the existing explanations were rectified, it would still have been wrong to use the SIDS statistics in *Clark*.

[8] 'You are incarcerated for their killing – *for almost no other reason* than that a leading paediatrician, Sir Roy Meadow, was permitted to tell the jury that the likelihood of there being two infant deaths in the same family was one in 73 million' (Wansell 2007), my emphasis. See also Shaikh (2007).

[9] In a public presentation, Professor Donnelly stated that '[the mistaken conviction] happened in large part here because the expert [Professor Meadow] got the statistics horribly wrong'. He concluded his presentation of the case with the statement that 'there is a situation where errors in statistics had really profound and really unfortunate consequences' (Donnelly 2005).

[10] 'Although we cannot know how the jury regarded the statistical evidence, it is reasonable to speculate that it was strongly influenced by the extremely small probability value of 1 in 73 million that both deaths could have arisen from SIDS, regarding this as ruling out the possibility of death by natural causes' (Dawid 2002: 75).

This explanation provides a principled reason that does not hinge on the actual consequences of the use of statistics in the specific case of *Clark*. It is thus not limited to *Clark* and can be applied to any similar case in which such statistics are adduced to eliminate potential natural, accidental, or other types of non-culpatory causes. It is argued here that even when the statistical evidence is gathered, analysed, and presented in the utmost professional manner by experienced expert statisticians and interpreted correctly by the judges and jurors, it is still wrong in principle to use it in cases such as *Clark*. This chapter thus concludes that such statistics should not be used in criminal proceedings to prove the material fact of causation (i.e. the fact that the harm was caused by the accused's misconduct).

This chapter focuses mainly on the decision of the first appeal to uphold the convictions, rather than on the initial trial or the second appeal, for several reasons. Firstly, trained and experienced judges, such as those who presided over the first appeal, are said to be less vulnerable to counsel's trial tactics, logical fallacies and so on, than lay juries, such as those who convicted *Clark* in the actual trial and who, most likely, were participating in a trial setting for the first time in their lives. Focusing on the first appeal thus reduces the importance of factors external to the available evidence and to the legal reasoning. Secondly, while the jury in the actual trial did not provide any detailed account of its reasoning,[11] the decision of the first appeal contains immense detail about the judges' reasons for upholding the convictions. Lastly, the convictions were upheld by the first appeal *despite* the immense attention to detail and argument.[12] It is the thoroughness of the judgment which makes its erroneous outcome so troubling and thus also makes it a more suitable source for analysis than the second appeal or the actual trial.

2 Factual Background

Before considering possible explanations, it is first necessary to provide a more detailed description of the case.[13] Clark, who was 35 years old at the time of the trial, was a solicitor with no previous criminal record. Christopher, her first child, was born healthy but died at the age of 11 weeks while her husband was out at an office party. Dr Williams, a pathologist, initially treated the death as a case of SIDS and considered the cause of death to be lower respiratory tract infection. He found bruises on the body and a small split in the frenulum, but he thought at the time that these were caused by resuscitation attempts. Before the body was cremated,

[11] In general, the jury is not required to provide reasoning for its 'guilty' or 'not-guilty' judgment (Roberts and Zuckerman 2010: 65–72).

[12] By way of illustration, this decision consists of over 35,000 words of detailed description of the facts, the witnesses' testimonies and the arguments of the parties, together with the judges' reasoning.

[13] The following background is based on the description in the first judgment of the Court of Appeal, in order to remain as close as possible to the standpoint of the judges of the first appeal.

Dr Williams took photographs and preserved slides of the lungs. Two years later, Harry, the second child, was also born healthy but was found dead at the age of 8 weeks by his mother when the husband was not in the room. The findings at the autopsy were indicative of nonaccidental injury, and Dr Williams determined the cause of death as shaking. He also revisited Christopher's case, conducted further tests, and altered his opinion, concluding that Christopher's death was also unnatural and that the evidence was suggestive of smothering (*Clark-I*: [2]–[3]).

The prosecution case was that Clark had murdered both babies (*Clark-I*: [6]–[7]). According to the prosecution, neither could be considered SIDS because of the existence of recent and old injuries that had been found in each case. There were several similarities between the cases: both babies were of similar age and found unconscious shortly after having been fed, in the same room, by Ms Clark, when she was alone with them, and when Mr Clark was absent or about to leave. Most importantly, in each case there was evidence of previous abuse and of deliberate injury recently inflicted. With regard to Christopher,[14] three pieces of medical evidence were adduced. Firstly, there was bleeding in his lungs, both old and fresh. The old bleeding is a marker for asphyxia and cannot be explained by the nosebleed Christopher had during a family trip to London because so much blood going into the lungs would have required urgent hospital treatment, but in fact Christopher recovered spontaneously. The nosebleed was also independently consistent with a prior attempt at smothering. Secondly, in the autopsy, Christopher's frenulum was found to be torn. The prosecution alleged that this was a result of deliberately inflicted injury rather than of resuscitation efforts and thus consistent with smothering. Lastly, Dr Williams had no doubt that he saw bruises on the body.

The defence case was that the evidence available was insufficient to prove beyond reasonable doubt that the babies' cause of death was unnatural (*Clark-I*: [10]–[15]). The defence therefore suggested that the babies must have died of natural causes, though it did not commit to any specific natural cause (including SIDS).[15] The defence emphasised that two of the prosecution's expert pathologists (Professor Green and Dr Keeling) gave the cause of death as unascertained and that therefore the prosecution's entire case hinged on the reliability of Dr Williams, who performed the autopsies. With regard to the medical evidence, the defence claimed that the marks on the body, which were interpreted by Dr Williams as bruises, were not examined under a microscope and were not seen at the hospital by other personnel who saw the baby. The injured frenulum could have been caused during insertion of the laryngoscope. As for the fresh bleeding, the defence claimed that it was only a marker for smothering and was often found both in suspicious cases and

[14] The description focuses on the prosecution case regarding Christopher, because the microbiological results that led to the second appeal and to Clark's release were relevant only to Harry, and the available medical and pathological evidence concerning Christopher's death remained more or less the same following Clark's conviction.

[15] In order to secure an acquittal, the defence did not need to provide a (natural) explanation for the babies' deaths but only to establish that there was a reasonable doubt in the (unnatural) explanation provided by the prosecution.

in cases of cot death. The old bleeding could have been a result of the nosebleed, the occurrence of which was not disputed. The defence argued that it was not caused by an attempted smothering because it was unlikely that Clark would have attempted to smother Christopher on the day she had brought him to London to show him to her friends.

3 Existing Explanations

3.1 The Flaws in Meadow's Calculation

A common explanation of why it was wrong to use the SIDS statistics in *Clark* focuses on the way the statistical evidence was used or, perhaps more precisely, misused. The statistical calculation used by Professor Meadow, which aimed to establish that the probability of two cases of SIDS occurring in the same family was 1 in 73 million, drew substantial criticism from the statistician community[16] and was even singled out as the cause of the wrongful convictions.[17] Meadow reached this figure by first calculating the probability of a SIDS death in a family similar to the Clarks (professional, non-smokers, and mother aged over 26) to be 1 in 8,543. He then multiplied this figure by itself to reach the probability of two cases of SIDS in the same family. This calculation was challenged on two grounds. Firstly, the probability of one case of SIDS (1 in 8,543) was contested (*Clark-I*: [138]), also with reference being made to another study (the CONI study, *ibid.*). Secondly, and more importantly, Meadow's calculation was flawed because it assumed independence between the two events of death.[18] In other words, Meadow assumed the probability of a second SIDS death to be equal to the probability of a first SIDS death.[19] However, there are numerous potential genetic and environmental reasons why a family which has already experienced SIDS is at higher risk of experiencing (another) SIDS death than a family which has never previously experienced SIDS (e.g. if the parents have certain genes which increase the risk of SIDS).[20] Had the

[16] See, for example, the official statement of the Royal Statistical Society, concluding that '[t]he well-publicised figure of 1 in 73 million thus has no statistical basis' (RSS 2001). See also Hill (2004).

[17] See Professor Donnelly's remarks, cited at footnote 9.

[18] Dawid states that 'this calculation is extremely dubious, being based on unrealistic assumptions of independence' (Dawid 2002: 75).

[19] In formal notation, if S1 is the first SIDS death and S2 is the second SIDS death, Meadow assumed that $p(S1) = p(S2 \mid S1) = p(S2 \mid \neg S1)$.

[20] Indeed, research has found a correlation between certain genes and SIDS. See Summers et al. (2000) and, more recently, Dashash et al. (2006).

calculation been accurate, the correct figure would have been less dramatic than 1 in 73 million.[21]

However, it remains questionable how the flaws in Meadow's calculation, serious as they may be, can explain why it was wrong to use the SIDS statistics in *Clark*. On their own, the most that these flaws can establish is that the calculation should have been done more carefully. However, that leaves open both the question of whether the wrongful convictions would have been prevented and the question of whether such statistics, even if calculated correctly, should have been used in *Clark*.

Furthermore, the flaws in Meadow's calculation were known and highlighted not only during the first appeal but also during the actual trial, before the involvement of the expert statisticians. During the trial, the defence referred to Professor Emery's study which showed cases of a second cot death to be more frequent than was argued by Sir Meadow (*Clark-I*: [116]). Professor Berry, one of the defence experts, emphasised the possibility of unknown factors, which further undermined Meadow's assumption of independence (*Clark-I*: [122]), as the court also noted (*Clark-I*: [155]). The jury was also warned by the trial judge that the 73 million figure should be treated with caution (*Clark-I*: [144]) and was reminded that the risks of SIDS were inherently greater in a family which had already experienced SIDS (*Clark-I*: [145]), a warning which undermined the assumption of independence made by Meadow in his calculation. Given that Meadow's calculation was strongly disputed from the outset, it is difficult to accept that the flaws in the calculation were responsible for the wrongful convictions. Therefore, while such calculations should be made more carefully, it is hard to see how the flaws in Meadow's calculation can offer an explanation of why it was wrong to use the SIDS statistics in *Clark*.

3.2 The Psychological Effect of the Statistical Evidence

Another related common explanation of why it was wrong to use the SIDS statistics in *Clark* refers to the psychological effect that the impressive figure of 1 in 73 million might have had on the jurors in Clark's trial or on the judges in the first appeal. According to this explanation, the figure of 1 in 73 million had such a strong psychological impact that it caused the suppression and underappreciation of the nonstatistical evidence that was more favourable to Clark. Stephen Clark, her husband, for instance, commented on the statistics as being 'an arrow through the fog' that gave the jury a compelling case against his wife.[22] According to this explanation, it was wrong to use the SIDS statistics because had they not been used,

[21] It was reported in the media that 'the odds are closer to 200 to one' (BBC 2005). See also Wansell (2007). But note that the grounds for this estimation are unclear.

[22] Sweeney and Law (2001). A similar comment was made by Dawid (2002), as cited in footnote 10.

other evidence would have received adequate attention and the wrongful convictions could have been prevented.[23]

However, this explanation relies on the empirical assumption that, when considering statistical evidence, people tend to discount other pieces of nonstatistical evidence which are available. By contrast, several commentators have pointed out that empirical research reveals the exact opposite (Shaviro 1989: 545). For example, the influential psychological experiments conducted by Kahneman and Tversky show that people tend to disregard statistical evidence ('background information') when other specific evidence is also available.[24] This empirical research is particularly relevant to *Clark*, since there, the fact-finder had immensely detailed *specific* evidence to consider, mostly about the pathological findings found in the autopsies.[25] The empirical research may thus suggest that it is unlikely that this specific evidence was suppressed by the general base-rate statistics presented by Meadow. Furthermore, if a hypothetical fact-finder in a similar case to *Clark* were presented with specific evidence which clearly showed a *natural* cause of death, it is difficult to accept that she would be so overwhelmed by the 1 in 73 million figure that she would convict in spite of such specific evidence. If anything, she would be more likely to fail to give this base-rate statistic the weight it deserved.[26] Hence, if there is a concern about the evaluation of the SIDS statistics, it is more likely to involve the unjustified *underweighting* of the statistics than any overwhelming effect.

[23] This explanation resonates with a more general point made by Tribe over 35 years ago: '[t]he problem of the overpowering number, that one hard piece of information, is that it may dwarf all efforts to put it into perspective with more impressionistic sorts of evidence' (Tribe 1971a: 1360).

[24] Kahneman and Tversky (1982). For similar findings about various types of circumstantial evidence when used in criminal proceedings, see Heller (2006: 250–252).

[25] For a detailed description of the prosecution's medical evidence, see *Clark-I*: [25]–[33] for Christopher and [50]–[63] for Harry.

[26] To illustrate this, Schoeman suggests the following example (Schoeman 1987: 180–181). A pellet dish was overturned in a rabbit pen. Among its 500 rabbits, only one was brown and the rest were white. A person who was watching the pen says it was the brown rabbit which overturned the pellet dish. Her colour identification ability is tested and found accurate in 95% of cases. On this evidence, what was the colour of the rabbit which overturned the pellet dish? Most people would believe the eyewitness and accept that the colour of the rabbit in question was brown rather than white. However, when the two pieces of evidence are properly combined, the probability of the rabbit's being brown is *only 4%*. If the eyewitness is presented with each of the 500 rabbits, she is very likely to identify the brown rabbit correctly (95%) but also would misidentify 25 white rabbits as brown (because her colour identification has a 5% error rate). Among the 500 rabbits in the pen, about 26 would be identified as brown, though only one of them is really brown. The probability that the rabbit which was identified as brown is indeed brown is therefore around 4% (1/26 = ~0.04). The probability that the rabbit which overturned the dish is white *despite* the eyewitness testimony is 96%.

3.3 The Prosecutor's Fallacy

Yet another common attempt to explain why it was wrong to use the SIDS statistics in Clark, one which also refers to the psychological effects of the SIDS statistics, concerns the 'prosecutor's fallacy'.[27] The term 'prosecutor's fallacy' identifies the logical mistake of treating the probability of the occurrence of the available evidence given the innocence of the accused as if it were the probability of innocence given the available evidence.[28] In the first appeal, the defence argued that the probability of two deaths given Clark's innocence was confused with the probability of Clark's innocence given the occurrence of two deaths. The defence alleged that the jury fell prey to the prosecutor's fallacy by understanding the SIDS statistics as showing that the probability of Clark's innocence given these two deaths was 1 in 73 million, instead of understanding them as showing that the probability of two SIDS deaths given Clark's innocence was 1 in 73 million.[29]

However, to the extent that this explanation has strength, it is mainly with regard to the decision of the jury rather than of the judges of the first appeal.[30] In general, judges have more experience in dealing with complex and scientific evidence because such evidence appears in many cases and evaluating it is a frequent task in their day-to-day routine. More importantly, the judges of the first appeal were equipped with the expert opinions of two distinguished statisticians (Professor Phil Dawid and Dr Ian Evett).[31] Therefore, it is hard to accept that the judges of the first

[27] Nobles and Schiff (2005). For more about the prosecutor's fallacy, see Balding and Donnelly (1994). Concise explanations may be found in R v Doheny and Adams [1997] 1 Cr App R 369, 372–375 and in the Forensic Science Service's Guide to DNA, 27, accessible at http://www.cps.gov.uk/legal/assets/uploads/files/lawyers'%20dna%20guide%20kswilliams%20190208%20(i).pdf, accessed 22 June 2012.

[28] In a formal notation, where $p(x)$ stands for the probability that the proposition x is true, $p(x \mid y)$ for the probability of x given that proposition y is true, G for the proposition that the accused is guilty (and ¬G for the proposition that he is innocent) and E for the occurrence of the evidence, the prosecutor's fallacy means the confusion of $p(E \mid \neg G)$ with $p(\neg G \mid E)$. See also Balding and Donnelly (1994: 718–720).

[29] *Clark-I*: [162]. Interestingly, the court also refers (at [177]) to another known difficulty in using statistical evidence, which relates to the application of base-rate frequencies (indefinite probabilities) to an individual case (definite probabilities). For this difficulty, see Pollock and Cruz (1999: 92–111). They argue that none of the existing theories of probability can support a move from indefinite to definite probabilities.

[30] Yet even with regard to the jury, it is difficult to substantiate the concern that the jury might have fallen prey to the prosecutor's fallacy. See the detailed discussion, including extracts from the actual trial, in the judgment of the first appeal, *Clark-I*: [162]–[184].

[31] The fact that the statisticians were not called to give oral testimony does not mean that the court ignored their expert opinions or was unwilling to engage in the statistical issues at stake. The legal procedure allows expert opinion to be given in either oral or written testimony (Criminal Justice Act 1988 c. 33s. 30(1)). The court in the first appeal was satisfied with the experts' written reports and accepted the defence's point that in the initial trial 'the judge appeared to endorse the prosecution's erroneous approach' (*Clark-I*: [184]). The court therefore defined '[t]he ultimate question' to be 'whether the error of approach rendered the conviction unsafe' (*ibid.*), which is a question of law rather than of statistics, and hence there was no need for oral testimony.

appeal fell prey to the prosecutor's fallacy, especially since they had been warned about it specifically by the expert statisticians (even assuming such a warning was in any way necessary).

3.4 Bayes' Theorem

A further explanation for why it was wrong to use the SIDS statistics in *Clark* considers the courts' refusal to use Bayes' Theorem. Bayes' Theorem is a mathematical formula that aims to instruct an agent on how to rationally alter his or her initial (prior) probability in light of new evidence.[32] Several eminent statisticians support using Bayes' Theorem in situations where the jury faces both statistical and nonstatistical evidence.[33] It could be argued that it was wrong to use the SIDS statistics in *Clark* because the figure was adduced on its own. Instead, Bayes' Theorem should have been used to combine the statistical evidence with the nonstatistical evidence and to assess accurately the probability of Clark's guilt given all the available evidence.[34] It should be emphasised that such an explanation does not imply that using the SIDS statistics in *Clark* was wrong in principle. On the contrary, such critics tend to be sympathetic to the use of statistical evidence in court, as long as it is used correctly,[35] which means, *inter alia*, using Bayes' Theorem.[36]

---

[32] A good introduction to Bayes' Theorem can be found in Dawid (2002: 72–78), Fairley and Finkelstein (1970: 498–501), and Roberts and Zuckerman (2010: 153–159).

[33] In an unprecedented attempt, Professor Donnelly was allowed to take a jury through the application of Bayes' Theorem to determine the accused's guilt (R v Adams (No 1) [1996] 2 Cr App R 467, hereafter *Adams-I*). The Court of Appeal responded to this attempt by noting that 'to introduce Bayes' Theorem, or any similar method, into a criminal trial plunges the jury into inappropriate and unnecessary realms of theory and complexity deflecting them from their proper task' (*Adams-I*: 482). This unprecedented attempt proliferated scholarly debate on the issue, a summary of which can be found in Roberts and Zuckerman (2010: 153–163).

[34] Dawid, for example, holds that in order to incorporate the SIDS statistics into the body of other available evidence, '[i]t is necessary to make an assessment (formal or informal) of the probability of observing the medical evidence, under each of the two causes under consideration' (Dawid 2001: §21). He then concludes that '[e]ven though assessment of the relevant probabilities may be difficult, there is a clear and well established statistical logic for combining them and making appropriate inferences from them, which was not appreciated by the court' (Dawid 2001: Conclusion).

[35] Dawid even compares the current legal approach to statistics to the state of science before Galileo: '[t]he current state of legal analysis of evidence seems to me similar to that of science before Galileo, in thrall to the authority of Aristotle and loth to concede the need to break away from old habits of thought. Galileo had the revolutionary idea that scientists should actually look at how the world behaves. It may be equally revolutionary to suggest that lawyers might look at how others have approached the problem of interpretation of evidence, and that they might even have something to learn from them' (Dawid 2002: 71–72).

[36] See the sources at footnotes 33 and 34.

However, any explanation that refers to Bayes' Theorem faces various difficulties, most evidently the need to quantify the numerous pieces of complex medical evidence into precise probabilities. Furthermore, the legal literature has noted several general difficulties in applying Bayes' Theorem in criminal courts, and the question of whether this statistical method should be used in court is probably one of the most debated issues in the theory of evidence law.[37] It is worth mentioning some of the main difficulties in simplified form. Firstly, Bayes' Theorem requires an assignment of prior probability of guilt, before any evidence is introduced (p(G)). However, it is questionable how assigning prior probability of guilt could be consistent with the presumption of innocence.[38] Secondly, it is questionable whether jurors, lawyers, and judges, all of whom usually lack any statistical training, would be able to deploy this method accurately.[39] Thirdly, once the calculation has produced a figure, it is contentious as to whether and how this figure could or should be translated into a guilty/not-guilty verdict (Nesson 1985).

Given these difficulties in quantifying the complex medical evidence and in applying Bayes' Theorem in court, it is far from clear that using Bayes' Theorem would have assisted the administration of justice to avoid the wrongful convictions. While the applicability of Bayes' Theorem in court remains an important theoretical question, it is hard to see how the courts' refusal to use Bayes' Theorem can explain why it was wrong to use the SIDS statistics in *Clark*.

3.5 *The Insignificance of the SIDS Statistics*

All four explanations seem to assume, either explicitly or implicitly, that the use of the SIDS statistics in *Clark* was wrong because it was responsible, at least in part, for the wrongful convictions.[40] They mainly diverge on the details of exactly how the SIDS statistics led to the wrongful convictions: by exaggerating the rarity of two SIDS deaths in one family, by creating a psychological effect that overshadowed other exonerating evidence, by inducing the prosecutor's fallacy or by requiring the use of Bayes' Theorem, a method which the courts refused to adopt. However, all these common explanations seem to share the assumption that had the SIDS statistics been used correctly, or had they not been used at all, the wrongful convictions could have been prevented.

[37] For a good summary of this intensive and extensive debate, see Roberts and Zuckerman (2010: 153–159). In particular, in a controversial and challenging book, Cohen has provided six paradoxes which challenge the applicability of the mathematical theory of probability to the legal context (Cohen 1977).

[38] Tribe (1971a). Cohen also argues that giving the presumption of innocence its true meaning by assigning $p(\text{G}) = 0$ will render the formula useless (Cohen 1977: 107–109).

[39] This concern receives empirical support from the work of Kahneman and Tversky (1980).

[40] Remarks in that vein can be found both in the public media (see e.g. footnotes 7 and 8) and among the expert statisticians (see e.g. footnotes 9 and 10).

Contrary to this shared assumption, it is argued here that the role of the SIDS statistics in *Clark* is much overrated. Firstly, the prosecution had a strong case even without the SIDS statistics. In the *weaker* case of Christopher, the prosecution pointed to bruises, the torn frenulum and fresh bleeding in the lungs. The defence challenged the existence of each of these pieces of evidence, and they received thorough and repeated scrutiny both during the long trial and in the lengthy decision of the first appeal. Yet it is worthwhile to note that each of the defence medical experts 'agreed that if there was bruising, the injury to the frenulum and bleeding in the lungs, it suggested asphyxia' (*Clark-I*: [40]). The prosecution case about Christopher did not hinge on a single piece of medical evidence, while the evidence relating to Harry was even more worrying.[41] Even if the defence were successful in establishing a reasonable doubt about one of the pieces of evidence, the cumulative weight of these pieces of evidence was probably (and should have been) higher than the sum of its parts. It is therefore questionable whether using the SIDS statistics correctly, or not using them at all, would have, or should have, changed either the jury's decision to convict or the first appeal's decision to uphold the convictions. Given that the microbiological results which led to Clark's release were not yet known at that stage, the convictions would probably have been reached by the jury and upheld by the court in the first appeal, with or without the SIDS statistics.

Secondly, the SIDS statistics were inessential to the prosecution case, since the defence experts accepted the very fact that this evidence was adduced to prove: that SIDS was not the cause of death. This was most evident in Harry's case. Dr Whitwell, for the defence, testified that '[s]he would not classify this a SIDS death because a true SIDS death should be completely negative and would not normally occur at this time in the evening, after a feed, with the child in a bouncy chair' (*Clark-I*: [77]). Dr Rushton, also for the defence, went even further and 'agreed that there were features in both deaths that gave rise to very great concern and for that reason he would not class them as SIDS deaths' (*ibid.*). Little wonder that, when considering the fallacies of Meadow's statistics, the court concluded that:

> [The statistical evidence] was very much a side-show at trial. The experts were debating the incidence of genuine SIDS (unexplained deaths with no suspicious circumstances) in a case where both sides agreed that neither Christopher's death nor Harry's death qualified as such.[42]

[41] See a summary in *Clark-I*: [8] and a detailed description in *Clark-I*: [50]–[63].

[42] *Clark-I*: [142]. The defence experts' surprising concession that the deaths were not SIDS may be explained by a subtle yet crucial difference between how the experts used the term 'SIDS' and how the court used it. SIDS is defined as 'the sudden death of a baby that is unexpected by history and in whom a thorough necropsy examination fails to demonstrate an adequate cause of death' (*Clark-I*: [104]), and the court rightly noted that '[c]learly the accuracy of that definition depends on the pathologists' thoroughness in autopsy, and on his or her interpretation of the findings' (*Clark-I*: [105]). However, the first appeal also referred to SIDS as a basket classification for all unexplained natural deaths (including cases where the autopsy was insufficiently thorough). For example, when discussing the SIDS statistics, the court referred to SIDS as 'unexplained deaths with no suspicious circumstances' (*Clark-I*: [142]; another example appears at [170]). Perhaps this subtle difference in definition caused the court to misinterpret the defence experts as accepting that the babies' deaths could not be classified as 'unexplained deaths with no suspicious circumstances', while the defence experts probably agreed that the babies' deaths could not be classified as SIDS because of their concerns about the thoroughness of the autopsies.

4 The Contrastive Explanation

Having considered and rejected the existing common explanations of why it was wrong to use the SIDS statistics in *Clark*, this part proposes an alternative explanation. It is argued here that it was wrong in principle to use the SIDS statistics, regardless of their actual share in the responsibility for the wrongful convictions. The explanation here is also not connected to concerns regarding the correct use and presentation of statistics, which have been raised in relation to *Clark* and have been discussed more generally in the academic literature.[43] While objections to the use of statistical evidence which are based on such concerns are important, they are nevertheless practical in nature and may be overcome by better education and training of the legal profession[44] and/or by more assistance from expert statisticians.[45] Such concerns should be distinguished from the position that the use of the SIDS statistics in *Clark* was wrong in principle, which is what the following section seeks to establish.

The most appropriate point to begin is the way in which the expert statisticians in *Clark* thought the SIDS statistics *should have been used*. Consider the testimony of Professor Dawid, one of the two expert statisticians for the defence, who states that:

> The laws of probability now focus attention on, not the absolute values of these probabilities of the two deaths in one family arising from the different causes considered, but on their **relative** values.[46]

He then concludes that:

> [The probability of two SIDS deaths in the same family] could only be useful if compared with a similar figure calculated under the alternative hypothesis that both babies were murdered.[47]

---

[43] See, for example, the exchange between Tribe and Fairley and Finkelstein about the correct statistical analysis of an example brought by Fairley and Finkelstein (1970), Tribe (1971a, b), and Fairley and Finkelstein (1971).

[44] Various scholars have rightly called for such training. See, for example, Koehler (1992: 148–149).

[45] Dawid, for example, states that 'statisticians ... have much to contribute towards identifying and clarifying many delicate issues in the interpretation of legal evidence' (Dawid 2002: 71–72). Dawid refers to both statistical and nonstatistical evidence. These remarks resonate with similar suggestions made decades ago in the United States. See, for example, Good (1950: 66–67) and Fairley and Finkelstein (1970: 502, 516–517). See also the official statement of the Royal Statistical Society (footnote 16).

[46] Dawid (2001: §18), emphasis original. This view is repeated in Dawid (2002: 76).

[47] Dawid (2001: Conclusion). It should be noted that Dawid's reference to 'murder' is mistaken and a more accurate category should have been used instead, such as 'causing the death'. 'Murder' is inaccurate because it assumes that all the material facts of the offence were proven beyond reasonable doubt and that no criminal defence was applicable. However, the SIDS statistics were brought to prove a specific material fact in the actus reus, namely, the fact of causation, and therefore, they should have been compared with the probability that it was the accused's conduct which caused the babies' deaths. Many cases which would not fit into the category of 'murder' would still fit into the category of 'causing the death' (e.g. cases in which the accused was insane or did not have the intent to cause the death).

To some extent, Dawid's conclusion already includes an objection to the use of the SIDS statistics. According to this objection, the absolute probability that the deaths occurred for innocent reasons (the first hypothesis) is meaningless on its own. It does not matter how likely or unlikely it is that the deaths occurred for innocent reasons. All that matters is how low (or high) this probability is *relative to the probability that the events occurred for culpable reasons* (the second hypothesis). Even if the probability of two SIDS deaths is strikingly small, all that matters is how it stands relative to the probability that a mother would murder both her babies, which was estimated by Dawid to be about 1 in 2 billion.[48] As this probability is much lower than the probability of two SIDS deaths, the SIDS statistics actually support Clark's *innocence*.[49] In addition to his criticism of Meadow's amateur calculation, Dawid's main point is that discussing the probability of two SIDS deaths in one family without referring to the alternative hypothesis is meaningless and misleading.[50]

In the first appeal, the court dismissed the exercise of comparing the probabilities of the two hypotheses as 'not realistic' and rejected any suggestion of estimating the probability of the second hypothesis (*Clark-I*: [176]). This dismissal was too quick. The probability of the first hypothesis (two SIDS deaths), as given by Meadow, was accepted by the court in the first appeal, after it had concluded unequivocally that this evidence was 'clearly relevant and admissible' (*Clark-I*: [166]). So why was it unrealistic to consider the probability of the second hypothesis (two murders)? The difference between the two pieces of evidence cannot lie in a lack of information, because Dawid's expert report offered preliminary statistical evidence for the second hypothesis.[51] Nor can the difference lie in a lack of statistical expertise, because the court in the first appeal was equipped with the expert opinions of two experienced statisticians, Professor Dawid and Dr Evett. Lastly, the difference

[48] Dawid, in his expert opinion, states the figure of 2,152,224,291 (Dawid 2001: §16), though he cautions that this is merely illustrative figure and 'its realistic estimation would be subject to all the caveats and cautions that have already been sounded above for the case of estimating the probability of two deaths from SIDS' (Dawid 2001: §17).

[49] Dawid (2001: §21). However, Dawid's preliminary calculation does not take into account at least two important factors. Since not all perpetrators are caught, indicted, and convicted, the recorded number of murders, on which Dawid's preliminary calculation is based, might be significantly smaller than its actual number. As a result, the actual probability of murder might be significantly higher than what Dawid's preliminary calculation shows. However, there is another neglected factor in his calculation, one which pushes the probability in the opposite direction. The prosecution is required to prove a specific unnatural cause of death (e.g. smothering, shaking, knifing). In contrast, Dawid's tentative suggestion that the probability of a mother murdering her two babies is about 1 in 2 billion refers to the *generic* unnatural cause of murder (i.e. it lumps together cases of different unnatural causes). If the court were to compare probabilities, the SIDS probability should be compared with the probability of a mother murdering her two babies *by smothering them*. This probability is probably much lower than the probability of murder to which Dawid refers.

[50] See also the warning given by the Royal Statistical Society, according to which '[a]side from its invalidity, figures such as the 1 in 73 million are very easily misinterpreted' (RSS 2001).

[51] Dawid (2001: §16), but see the reservations at footnotes 47 and 49.

cannot lie in a general antagonism to statistics.[52] Such an antagonism would apply equally to the SIDS statistics, which the court readily admitted. The approach of the court in the first appeal is therefore hard to understand or defend.

In the second appeal, the court reached a different conclusion, according to which the SIDS statistics probably should not have been admitted.[53] However, the court did not offer any explanation of why the first appeal's ruling regarding the admissibility of the SIDS statistics should be reversed. Nor did the court in the second appeal give its reasons for concluding that the SIDS statistics should not have been admitted. No guidance was provided for distinguishing this piece of statistical evidence from other kinds of statistical evidence which are regularly used in court (e.g. DNA evidence). The inconsistency between the two decisions and the lack of judicial reasoning in the second appeal demonstrate the need for a principled approach to the use of statistical evidence in court. Developing such a general approach requires a separate and extensive research project,[54] and thus it is beyond the scope of this chapter. However, the following explanation for why it was wrong in principle to use the SIDS statistics in *Clark* can be applied to other similar cases, and thus this case study of *Clark* is concluded with a more general recommendation.

The starting point of the proposed explanation is the issue of competing hypotheses, to which the expert statisticians drew attention. Competing hypotheses are central to the theory of legal fact-finding.[55] One context in which competing hypotheses are particularly important is that of determining the role of causal explanations in legal fact-finding. The importance of explanatory aspects of legal fact-finding has recently been highlighted by Allen and Pardo in their discussion of the theory of inference to the best explanation and its application to evidence law (Allen and Pardo 2008). A related theory, one which can shed light on the way the SIDS statistics were used to prove the material fact of causation, is the theory of *contrastive explanation* (van Fraassen 1980: 97–157; Lipton 1990).

The theory of contrastive explanation can be described as follows. The idea, in outline, is that most, if not all, requests for explanation ('why' questions) contain a contrast: instead of understanding the request as 'why P', the theory of contrastive explanation suggests understanding such questions as 'why P *and not Q*'. To borrow Lipton's terminology, 'why' questions contrast between facts and foils.[56] Sometimes the contrast is explicit: 'why did you order cheese quiche rather than

[52] See also Posner's criticism of the legal profession for its 'prevalent (and disgraceful) math-block' (Posner 1987: 778).

[53] 'If there had been a challenge to the admissibility of the evidence we would have thought that the wisest course would have been to exclude it altogether' (*Clark-II*: [177]).

[54] I began outlining such a general approach in Pundik (2009), and I intend to develop it further as part of my current research project, titled 'Generalizations in the Law'.

[55] See, in general, Kaye (1992), Robertson and Vignaux (1993) and Allen (1986). The issue of competing hypotheses also arises in the context of the story model. See Pennington and Hastie (1992) and Twining (2006).

[56] Lipton (1990: 249–252). For a more elaborated version of his fact/foil distinction, see Lipton (2004: 30–37).

beef lasagne?' Here the foil (beef lasagne) is stated explicitly as part of the question. However, the theory of contrastive explanation argues that often (if not always)[57] the contrast is implicit. For example, when someone asks 'why did the mercury in the barometer go up before the storm?', the implicit contrast is 'why did it go up *rather than stay where it is*', and not 'why did it go up rather than break the glass'.[58] If someone answered the question with a detailed explanation of the rigidity of the glass,[59] we would consider this explanation to be irrelevant and inadequate.[60]

Exploring the full application of the theory of contrastive explanation to legal fact-finding in general, and to the proof of causation in particular,[61] would probably require yet another separate research project. Nonetheless, in the limited context of analysing the use of the SIDS statistics in *Clark*, this theory can be employed to substantiate the expert statisticians' assertion that the SIDS statistics make sense

[57] For the issue of whether all why questions are contrastive, see Lipton (1990: 252–254).

[58] This example is taken from Lipton (1990: 252).

[59] The theory of contrastive explanation thus highlights the importance of the *context* in which the explanation is sought. Whether something counts as a *good* explanation depends on the context in which the question is asked (*ibid.*). If the question about the mercury rising is asked in a chemistry class about the qualities of glass, then the rigidity of the glass might be a good explanation. Thus to determine whether the answer constitutes a *good* explanation, we first need to know the context in which the question was put.

[60] Van Fraassen further suggests that a why question, a request for explanation, has three elements (van Fraassen 1980: 141–142). Firstly, it has a *topic*, which is the subject of the why question, the element that requires an explanation. For example, the topic of the question 'why did the mercury in the barometer go up' is the fact that the mercury went up. This is what the question assumes to be true. Secondly, the question has a *contrast class*, a group of propositions about the topic. For example, 'if the question is "why does this material burn yellow" the contrast-class could be the set of propositions: this material burned (with a flame of) colour x' (van Fraassen 1980: 142). The third element is *explanatory relevance*, which is determined by what is already known and what further information is required. For example, if it is already known that mercury in barometers goes up before storms (for instance, in the context of a discussion between weather forecasting experts), then reiterating what is already known will be irrelevant. In such a context, only answers that add new information will bear explanatory relevance (for instance, answers about the cause of the storm). Van Fraassen also suggests that an answer to a why question takes the form of *the preferred contrast* and not *the other contrasts* because of *the answer* (van Fraassen 1980: 144–145). For example, the mercury in the barometer went up (the preferred contrast) and did not stay where it was (the other contrast) because of the low air pressure that precedes storms (the answer). The answer includes what van Fraassen terms *the central presupposition*, which is that only the preferred contrast is true and the other contrasts are false. The question arises only when the background knowledge implies this central presupposition.

[61] Another related version of the theory of contrastive explanation which may yield interesting applications in the legal context is Schaffer's theory of *contrastive causation* (Schaffer 2005). According to Schaffer, it is misleading to construct questions of causation in terms of 'was it C which caused E'. Instead, the causal questions should be understood as having *four* elements: 'was it C rather than C^* which caused E rather than E^*'. Schaffer's work also includes an application of his theory to the law (Schaffer 2010). At this stage, it suffices merely to note the issue, as engaging with the metaphysics of causation will distract this chapter from its central question.

only if they are compared with the probability that a mother would smother her two babies. The fact-finder in *Clark* sought to answer the question of why the babies died. The contrast class of this why question was constructed from possible causes of death. The prosecution needed to establish that the preferred contrast was unnatural (smothering in this case)[62] rather than being any other contrast of natural causes of death such as SIDS. On its own, the very low probability of SIDS does not provide support for preferring one specific contrast to another. The low probability of SIDS supports the prosecution's preferred contrast only if this probability is low*er* than the probability of the prosecution's preferred contrast (smothering). But this comparative claim can be established only with evidence about the probability of the prosecution's preferred contrast. To support the prosecution's answer to the question of why the two babies died, the SIDS statistics must be accompanied by evidence detailing the rate of mothers smothering both their babies in the population of mothers who are similar to Clark.

The probability of smothering may be referred to either explicitly or implicitly. According to the expert statisticians, this reference should be made explicit: one should use empirical data to calculate the rate of mothers who smother both their babies in the wider population of non-smoking professional mothers aged over 26.[63] One might reject this idea and hold that the fact-finder should not be provided with any evidence of that sort. Yet ample empirical research shows that fact-finders consider explanations of innocence by contrasting them with explanations of guilt, and vice versa, *as a matter of course*.[64] In the absence of empirical data, the fact-finder might resort to general knowledge and common sense at best, or to prejudice and arbitrary guesswork at worst. Therefore, as supported by empirical research, emphasised by the expert statisticians and substantiated by the theory of contrastive explanation, using the SIDS statistics requires the fact-finder to compare, either explicitly or implicitly, the probability of SIDS with the probability that a professional non-smoking mother aged over 26 would smother both her babies on two separate occasions.

Once the need for this comparison is recognised, the reason why using the SIDS statistics in *Clark* was wrong in principle becomes clear. Developing a comprehensive and general account to justify why courts should not use statistical evidence about the rate of similar misconduct in a population similar to the accused is outside the scope of this chapter. Yet it is difficult to underplay the fact that justice systems around the world hardly ever, if at all, use such statistical evidence to convict.[65] While explaining why this is the case is a challenging theoretical question, the

[62] Interestingly, it was claimed that the prosecution had changed the allegation from shaking to smothering just before the trial commenced. See Batt (2004: 140).

[63] These three variables (professional, non-smoking and age over 26) were used by Meadow to extract the number of 1 in 73 million. See *Clark-I*: [118], and the empirical data in [121].

[64] See the various empirical studies surveyed in Heller (2006: 261–262).

[65] For descriptive research on the use of statistical evidence in United States courts, together with an attempt to identify patterns of when statistical evidence is used or rejected, see Koehler (2002).

probability of smothering seems a paradigmatic example of statistical evidence of a kind which should *not* be used in court.

Bearing in mind the scope of this chapter, a few preliminary remarks can nevertheless be made as to why evidence about the probability of smothering seems so intuitively objectionable. Using evidence about the rate of smothering among other similar mothers seems objectionable because it is inconsistent with regarding Clark as unique and morally autonomous individual. Using such evidence assumes that she would exhibit the typical behaviour of her peers (professional non-smoker mothers aged over 26). It also presupposes that these characteristics determine, either fully or partly, the individual's behaviour. Without presupposing that such a common determining property is shared between these mothers, it is unclear why this is the relevant group from which an inference to Clark's individual case should be made. Why not refer to mothers in general, parents of both genders, people whose last name starts with C and so on? The choice of this particular group as relevant to Clark's individual case thus implies that the members of this group share a property which affects their behaviour.[66] The guilt of an individual should be proved based on the particular facts of her case rather than on the rate of similar misconduct among other people with similar characteristics to hers.[67] These remarks are necessarily preliminary. But the main point is that evidence regarding the probability of smothering is clearly objectionable, even if it is difficult to account for why this is so.

5 Conclusion

This chapter has questioned whether it was wrong to use the SIDS statistics in *Clark*, and if so, why. Four common explanations were discussed and rejected. The first explanation refers to the flaws in Meadow's calculation of the 1 in 73 million figure. Yet, however serious these flaws are, it is difficult to accept that they can explain why it was wrong to use the SIDS statistics in *Clark*. The second explanation raises the suspicion that the SIDS statistics had an overwhelming psychological

[66] This point is linked to the problem of the reference class. Some legal scholars argue that relying on reference classes in legal fact-finding is problematic (e.g. Colyvan et al. 2001; Allen and Pardo 2007). However, these objections are unpersuasive, mainly because the use of *any* generalisation requires reliance on a reference class, so if there were something inherently wrong with relying on reference classes, the entire enterprise of legal fact-finding would become impossible. The point made in the text above is based on a different type of reference class argument: it assumes that the *choice* of one reference class over other alternative reference classes must be made for a reason. In the context of determining the individual's behaviour, it suggests that this reason is related to *causation*, namely, that the individual members of the reference class share something that causes them to behave in a similar way.

[67] It could be argued that this objection applies to any evidence rather than to statistical evidence alone because such generalisations are used in inferences from any type of evidence. This is an important objection, which is discussed and rejected in Pundik (2008: 312–315).

effect, which overshadowed other evidence more favourable to Clark. However, if there is a concern about the weight given to the SIDS statistics, it is more likely to be a concern of underweighting rather than overweighting. The third explanation focuses on the 'prosecutor's fallacy', yet it is hard to substantiate the concern that the court fell prey to this fallacy, especially given it was specifically warned about it by the expert statisticians. The last explanation considers the courts' refusal to use Bayes' Theorem. However, the difficulties with quantifying the complex medical evidence and applying this statistical method in court make explanations which hinge on it difficult to sustain.

A more general difficulty with all four explanations is that the connection between the SIDS statistics and the wrongful convictions is hard to sustain. The role of this evidence in *Clark* is much overrated, since the prosecution case was strong without the SIDS statistics. Furthermore, even the defence experts agreed that the deaths should not be categorised as SIDS, making the SIDS statistics inessential for the prosecution case. It is therefore difficult to accept that the wrongful convictions could have been prevented had the SIDS statistics been used correctly, or even had they not been used at all.

However, the question still remains whether it was wrong in principle to use the SIDS statistics in *Clark*, and if so, why. This chapter has argued that even if the issues identified by the existing explanations were properly addressed, it would still have been wrong *in principle* to use the SIDS statistics in *Clark*. This position is based on the theory of contrastive explanation. Its application to *Clark* begins with a similar point to that of the expert statisticians, namely, that the SIDS probability makes sense only when it is compared with the smothering probability. However, unlike the expert statisticians' approach, the theory of contrastive explanation does not require any commitment to Bayesian methods in order to reach this point. This is an advantage because Bayesian methods were explicitly rejected by the English Court of Appeals for '[plunging] the jury into inappropriate and unnecessary realms of theory and complexity deflecting them from their proper tasks' (*Adams-I*: 482).

Yet the main difference between the approach of this chapter and that of the expert statisticians lies in the conclusion. The expert statisticians' approach would require the application of Bayes' Theorem in order to make sense of the SIDS statistics, because using this method is 'the only logically sound and consistent approach to considering situations such as this'.[68] Their position thus implies that in order to make sense of the SIDS statistics, it would be necessary to use Bayes' Theorem and thus also necessary to adduce further statistical evidence to assess the probability that a mother similar to Clark would smother both her babies.

This chapter, by contrast, has reached a different conclusion. The problem with the use of statistical evidence in *Clark* was more fundamental than merely a flaw in the statistical analysis. The SIDS statistics should not have been used, for

[68] The phrase is taken from Professor Donnelly's response in *Adams*, after he was asked whether both the statistical and nonstatistical evidence could be evaluated in 'statistical terms' (namely, using Bayes' Theorem) (*Adams-I*: 471).

they would have required the use of another piece of evidence which is clearly objectionable: statistical evidence on the rate of smothering. If one piece of evidence cannot be used without another piece of objectionable evidence, then neither of them should be used. Regardless of whether the exercise of comparing probabilities of SIDS and smothering is 'realistic' or not, it should not be conducted as part of a criminal proceeding.

This chapter has focused on only a single case study, and hence any attempt to generalise from *Clark* to the wider issue of the use of statistical evidence in court should be made with caution. However, perhaps there is one lesson from *Clark* that can be phrased in more general terms. It is sometimes tempting to ask how likely it is that a given unusual event (e.g. two baby deaths in one family) will happen randomly, by accident, naturally, or for any other non-culpatory cause. If the event in question happens so rarely in the course of nature, this fact alone seems to provide evidential support for the hypothesis that the event must have happened as a result of deliberate human intervention, namely, the accused's misconduct. However, *Clark* alerts us to the fact that this type of inference is problematic, because the natural/random/non-culpatory causal explanation must be contrasted with a culpatory one. Using statistical evidence on the rate of a certain non-culpatory cause would thus require also using statistical evidence on the rate of the type of criminal behaviour attributed to the accused among other people similar to him or her.[69]

There may be contexts in which it would be appropriate to refer to the probability of an unusual event happening naturally and to use statistical evidence and methods to calculate this probability. For example, in the decision whether to equip parents with an apnoea monitor which detects pauses in breathing in young babies,[70] it may be wise to consider the probability of SIDS among families with similar characteristics. However, when the same question arises in the context of determining whether an individual is guilty of an alleged crime, using this kind of evidence becomes problematic. Since statistical evidence regarding the rate of misconduct among other people similar to the accused should not be used to determine the accused's guilt, the same goes for statistical evidence regarding the probability of non-culpatory causes. No doubt the use of statistical evidence in court is a complex issue, and the same logic would not necessarily apply

[69] Another case which illustrates the temptation to prove the fact of causation with statistical evidence is *Veysey*, in which the defendant was charged with arson and insurance fraud after his house was burnt down for the fourth time. See *United States v Veysey* 334F 3d 600 (7th Cir 2003). The conviction was based in part on an actuary's testimony according to which 'the probability of four residential fires occurring by chance during the 106 months between April 1989 (when Veysey bought the first house that he is known to have set fire to) and January 1998 (when he set fire to his last house) was only one in 1.773 trillion' (*ibid*.: [8]). It seems that the same objection against the use of the SIDS statistics in *Clark* would apply to the use of statistical evidence in *Veysey*.

[70] After the death of Christopher, the Care of Next Infant programme provided the Clarks with an apnoea monitor (*Clark-I*: [43]).

to statistical evidence used to prove other facts (such as identification proven with DNA evidence) or to statistical evidence used in noncriminal proceedings. However, this case study of the use of statistics in *Clark* exemplifies the following general point. Statistical evidence on the probability of non-culpatory causes should not be used in criminal courts to prove the material fact of causation. Before delving into the technical complexities of how to gather, analyse and present statistical evidence which shows the exact value of such probabilities, one should first ensure that raising the question of probabilities (and using statistical evidence to answer it) is appropriate in the given context. At least when determining an individual's guilt, it seems rather that it is not.

References

Allen, R. J. (1986). A reconceptualization of civil trials. *Boston University Law Review, 66*, 401–437.

Allen, R. J., & Pardo, M. S. (2007). The problematic value of mathematical models of evidence. *The Journal of Legal Studies, 36*(1), 107–140.

Allen, R. J., & Pardo, M. S. (2008). Juridical proof and the best explanation. *Law and Philosophy, 27*(3), 223–268.

Balding, D., & Donnelly, P. (1994). The prosecutor's fallacy and DNA evidence. *The Criminal Law Review, 10*, 711–721.

Barraclough, K. (2004). Stolen innocence: A mother's fight for justice—The story of Sally Clark (book review). *British Medical Journal, 329*, 177.

Batt, J. (2004). *Stolen innocence*. London: Random House.

BBC. (2005). *Cot death expert defends evidence*. http://news.bbc.co.uk/1/hi/health/4641587.stm. Accessed 22 June 2012.

BBC. (2007). *Alcohol killed mother Sally Clark*. http://news.bbc.co.uk/2/hi/uk_news/england/essex/7082411.stm. Accessed 22 June 2012.

Blom-Cooper, L. (2006). *Disciplining expert witnesses by regulatory bodies*. *Public Law, Spring*, 3–5.

Cohen, L. J. (1977). *The probable and the provable* (Clarendon library of logic and philosophy). Oxford: Clarendon.

Colyvan, M., Regan, H. M., & Ferson, S. (2001). Is it a crime to belong to a reference class? *The Journal of Political Philosophy, 9*(2), 168–181.

Dashash, M., Pravica, V., Hutchinson, I., Barson, A., & Drucker, D. (2006). Association of sudden infant death syndrome with VEGF and IL-6 gene polymorphisms. *Human Immunology, 67*(8), 627–633.

Dawid, A. P. (2001). *Expert report for Sally Clark appeal*. http://www.statslab.cam.ac.uk/~apd/SallyClark_report.doc. Accessed 16 June 2012.

Dawid, A. P. (2002). Bayes's theorem and the weighing evidence by juries. *Proceedings of the British Academy, 113*, 71–90.

Donnelly, P. (2005). *How juries are fooled by statistics*. http://www.ted.com/index.php/talks/view/id/67. Accessed 22 June 2012.

Dwyer, D. (2003). The duties of expert witnesses of fact and opinion: *R v Clark (Sally)* (Case Note). *International Journal of Evidence and Proof, 7*, 264–269.

Editorial. (2000). Conviction by mathematical error? Doctors and lawyers should get probability theory right. *British Medical Journal, 320*, 2–3.

Fairley, W. B., & Finkelstein, M. O. (1970). A Bayesian approach to identification evidence. *Harvard Law Review, 83*(3), 489–517.

Fairley, W. B., & Finkelstein, M. O. (1971). The continuing debate over mathematics in the law of evidence: A comment on "trial by mathematics". *Harvard Law Review, 84*(8), 1801–1809.
Good, I. J. (1950). *Probability and the weighing of evidence*. London: C Griffin.
Heller, K. J. (2006). The cognitive psychology of circumstantial evidence. *Michigan Law Review, 105*, 241–305.
Hill, R. (2004). Multiple sudden infant deaths – Coincidence or beyond coincidence? *Paediatric and Perinatal Epidemiology, 18*(5), 320–326.
Kahneman, D., & Tversky, A. (1980). Causal schemas in judgment under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology* (Vol. 1, pp. 49–72). Hillsdale: Erlbaum.
Kahneman, D., & Tversky, A. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristic and biases* (pp. 153–160). Cambridge: Cambridge University Press.
Kaye, D. H. (1992). Proof in law and science. *Jurimetrics Journal, 32*, 313–322.
Koehler, J. J. (1992). The probity/policy distinction in the statistical evidence debate. *Tulane Law Review, 66*, 141–150.
Koehler, J. J. (2002). When do courts think base rate statistics are relevant? *Jurimetrics Journal, 42*, 373–402.
Lipton, P. (1990). Contrastive explanation. In D. Knowles (Ed.), *Explanation and its limits* (pp. 247–266). Cambridge: Cambridge University Press.
Lipton, P. (2004). *Inference to the best explanation* (International library of philosophy 2nd edn.). London: Routledge.
Nesson, C. (1985). The evidence or the event? On judicial proof and the acceptability of verdicts. *Harvard Law Review, 98*(7), 1357–1392.
Nobles, R., & Schiff, D. (2005). Misleading statistics within criminal trials – The Sally Clark case. *Significance, 2*(1), 17–19.
Pennington, N., & Hastie, R. (1992). A cognitive theory of juror decision making: The story model. *Cardozo Law Review, 13*, 519–557.
Pollock, J. L., & Cruz, J. (1999). *Contemporary theories of knowledge* (2nd edn.). Oxford: Rowman & Littlefield.
Posner, R. A. (1987). The decline of law as an autonomous discipline: 1962–1987. *Harvard Law Review, 100*(4), 761–780.
Pundik, A. (2008). Statistical evidence and individual litigants: A reconsideration of Wasserman's argument from autonomy. *International Journal of Evidence and Proof, 12*, 303–324.
Pundik, A. (2009). *Statistical evidence: In a search of a principle*. DPhil thesis, University of Oxford.
R v Adams (No 1) [1996] 2 Cr App R 467.
R v Clark (No 1) [2000] EWCA Crim 54.
R v Clark (Crown Court Chester 9 November 1999).
R v Clark (No 2) [2003] EWCA Crim 1020.
R v Doheny and Adams [1997] 1 Cr App R 369.
Roberts, P., & Zuckerman, A. A. S. (2010). *Criminal evidence* (2nd edn.). Oxford: Oxford University Press.
Robertson, B., & Vignaux, G. A. (1993). Probability—The logic of the law. *Oxford Journal of Legal Studies, 13*, 457–478.
RSS. (2001). *Press release on the Sally Clark case, Oct 2001*. http://www.rss.org.uk/uploadedfiles/documentlibrary/348.doc. Accessed 22 June 2012.
Schaffer, J. (2005). Contrastive causation. *Philosophical Review, 114*(3), 297–328.
Schaffer, J. (2010). Contrastive causation in the law. *Legal Theory, 16*, 259–297.
Schoeman, F. (1987). Statistical vs. direct evidence. *Nous, 21*(2), 179–198.
Shaikh, T. (2007). *Sally Clark, mother wrongly convicted of killing her sons, found dead at home*. http://www.guardian.co.uk/society/2007/mar/17/childrensservices.uknews. Accessed 22 June 2012.

Shaviro, D. (1989). Statistical-probability evidence and the appearance of justice. *Harvard Law Review, 103*(2), 530–554.

Summers, A. M., Summers, C. W., Drucker, D. B., Hajeer, A. H., Barson, A., & Hutchinson, I. V. (2000). Association of IL-10 genotype with sudden infant death syndrome. *Human Immunology, 61*(12), 1270–1273.

Sweeney, J., & Law, B. (2001). *Gene find casts doubt on double "cot death" murders*. http://www.guardian.co.uk/uk/2001/jul/15/johnsweeney.theobserver. Accessed 22 June 2012.

Telegraph. (2007). *Obituary: Sally Clark*. http://www.telegraph.co.uk/news/obituaries/1545933/Sally-Clark.html. Accessed 22 June 2012.

Tribe, L. H. (1971a). A further critique of mathematical proof. *Harvard Law Review, 84*(8), 1810–1820.

Tribe, L. H. (1971b). Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review, 84*(6), 1329–1393.

Twining, W. (2006). Narrative and generalizations in argumentation about questions of fact. In *Rethinking evidence: Exploratory essays* (2nd edn.). Cambridge: Cambridge University Press.

United States v Veysey 334F 3d 600 (7th Cir 2003).

van Fraassen, B. C. (1980). *The scientific image* (Clarendon library of logic and philosophy). Oxford: Clarendon.

Wansell, G. (2007, 18 March). Whatever the coroner may say, Sally Clark died of a broken heart. *The Independent*.

Wilson, A. (2005). Expert testimony in the dock. *Journal of Criminal Law, 69*(4), 330–348.