

# UnHerd

## The danger of convicting with statistics [Courts have a bad history of using](#)



[Tom Chivers](#)

MAY 28, 2024 7 MINS

Sally Clark had two sons. Both died within weeks of birth, a year apart, apparently of sudden infant death syndrome (SIDS), sometimes called cot death. SIDS is — mercifully — rare; in England, at the time, it struck roughly one in 8,500 babies. That statistic led to Clark being prosecuted for double murder in 1998, despite there being little to no forensic evidence for her guilt.

A paediatrician, Roy Meadows, called as an expert witness for the prosecution, told the court that the probability of the two deaths happening by chance was [one in 73 million](#): that is, 8,500 times 8,500.

As it happens, that's not true. This calculation assumes that the deaths are entirely uncorrelated, but we know that SIDS can run in families and be affected by environmental conditions. If you have one case of SIDS in your household,

while incredibly rare, you are more likely to have a second; the 73 million figure is orders of magnitude too high. But that wasn't Meadows's big mistake.

His big mistake was the following: he assumed that if the probability of the two deaths happening by chance was one in 73 million, then the probability that Sally Clark was innocent was one in 73 million as well.

**“Courts, in the UK, US and elsewhere, have a bad history when it comes to the use of statistical evidence.”**

But this is wrong. Crucially, catastrophically wrong. As wrong as assuming that because only one human in eight billion is the President of the United States, there's only a one-in-eight-billion chance that the President of the United States is human.

Nonetheless, Meadows's testimony helped convict Clark in 1999. She spent three years in jail before her conviction was overturned on appeal. Her life was, obviously, ruined. It will not surprise you to learn that she drank herself to death four years later, alone. It's a haunting story.

The mistake made in Clark's case is a subcategory of a far wider failure of reasoning, a failure I discuss in my new book, [\*Everything is Predictable: How\*](#)

[\*Bayes' Remarkable Theorem Explains the World\*](#). But in legal circles it comes up again and again — often enough to have its own name: the “prosecutor’s fallacy”.

There was a recent article in the *New Yorker* about the nurse Lucy Letby, convicted of murdering seven babies in a neonatal ward. The online version is blocked in the UK, because of contempt-of-court laws, despite her appeal against the convictions being denied: she still faces a retrial on one count of attempted murder.

This piece is not about Letby. I do not know the facts of the Letby case and would not be allowed to write about them if I did; whether that is a strength or a weakness of British law, [I leave to others to discuss](#). But I do know that courts, in the UK, US and elsewhere, have a bad history when it comes to the use of statistical evidence. To understand why, we need to go back to the work of an 18th-century nonconformist minister.

The Reverend Thomas Bayes’s eponymous theorem was published in a paper called “[An Essay towards solving a Problem in the Doctrine of Chances](#)” in 1763, five years after his death. Previous work in probability theory had answered the question: how likely am I to see some event, given a hypothesis? For instance, if we assume that my dice are fair, and I roll three of them, I can expect to see three sixes one time in 216. That’s called sampling probability.

But most of the time, with statistics, we want to answer the opposite question: how likely is my hypothesis to be true, given some new event? That’s called inferential probability, and it’s a completely different thing.

Say I go to the doctor's and I get a cancer test. It's quite an accurate test: if I have cancer, it will correctly say so 99 times out of 100; if I don't have cancer, it will correctly say so 99 times out of 100. If I get a positive result, then, what's the chance that I have cancer? Is it 99%?

No. The answer is you don't know. At least, not with the information I've given you.

Imagine that this particular cancer is rare: only one person in 1,000 has it. You test 100,000 people at random. Of that 100,000, about 100 will have the cancer, and your test will pick up 99 of them. Of the remaining 99,900, your test will correctly say that 98,901 are cancer-free.

But that means that it will incorrectly say that 999 people do have cancer when they don't. So of your 100,000 tests, 1,098 will come back positive, and only 99 of them are true positives. If you are one of them, then there is just a 9% chance you have cancer.

You can't answer the question "How likely am I to have cancer, given this test?" without first answering the question "How likely did I think I was to have cancer in the first place?" That was Bayes's big insight. You need what is called a prior probability. If the cancer was less rare, then your positive test would be more worrying: if one person in every 100 had it, then a positive result would mean about a 50% chance you have the disease.

It's counterintuitive and weird. What do you mean, this 99% accurate test result is almost certainly wrong? But it is mathematically unavoidable.

You can probably see the bearing this has on statistical evidence used in court. Take DNA tests, for instance: you might do a DNA test on a sample from a crime scene. It matches a result on your database. There's only a one in 3 million chance that someone's DNA would match the sample by chance. Does that mean there's only a one-in-3-million chance that your suspect is innocent? As you'll realise by now, no it does not.

It depends on your prior probability. If your database is a random sample of the British population, then the prior probability that any given person is the culprit is one in 65 million. If you tested the whole population, you'd get about 20 matches just by chance.

But if you are a modern-day Hercule Poirot, and you're only testing 10 people trapped by a snowstorm in a country mansion, then your prior probability is one in 10, and the chance it's a false positive is about one in 300,000.

Real cases have turned on these details: a man called Andrew Deen was convicted of rape in 1990 on the basis of DNA evidence that an expert witness said there was only a one in 3 million probability of a chance match. [His conviction was overturned](#) — although he was found guilty again in a retrial — because a statistician pointed out that “How likely is it that a person's DNA would match the sample, if they are innocent?” and “How likely is it that someone is innocent, given that their DNA matches the sample?” are very different questions.

In Sally Clark's case, the problem was not testing, but clustering: two rare events happening simultaneously. But, again, the probability of seeing those events by chance is not the same as the probability that she was guilty.

Hers is far from the only case in which the use of statistics has raised suspicion. In 2022, the Royal Statistical Society wrote [a report on statistical evidence in criminal trials](#), and noted that one of the most common reasons that medical professionals are accused of murders is because “an apparently unusual number of deaths occurs among their patients”.

But these cases are doubly difficult to evaluate, the report noted, because “they involve at least two levels of uncertainty”. As well as the normal uncertainty over whether an individual committed a murder, there is uncertainty over whether any murders occurred at all.

The Dutch paediatric nurse Lucia de Berk was convicted, in two trials in 2003 and 2004, of seven murders and three attempted murders of children under her care. A criminologist told her trial that “the probability of so many deaths occurring while de Berk was on duty was only one in 342 million”.

But even if that were the case — and again, it wasn't; the RSS estimated that if you took into account all the relevant factors, the chance of seeing a cluster like that could be as high as one in 25 — that's not the same as the chance that de Berk was innocent. In order to establish that, you would need to take into account the prior probability that someone would be a multiple murderer — a mercifully tiny chance. De Berk's conviction was overturned in 2010, [thanks in part to the work of Bayes-savvy statisticians](#).

Bayesian reasoning doesn't only reveal the wrongly convicted — in some instances it could have led to the guilty being detected. During OJ Simpson's trial for the murder of his wife and her friend, for instance, the prosecution showed that Simpson had been physically abusive. The defence, though, argued that “an

infinitesimal percentage — certainly fewer than 1 in 2,500 — of men who slap or beat their wives go on to murder them” in a given year, so it wasn’t relevant to the case.

But this is simply the opposite mistake to the prosecutor’s fallacy. The probability that a man who beats his wife will go on to murder her in a given year might “only” be one in 2,500. But that’s not what we’re asking. What we want to know is, if a man beats his wife, and given that the wife is then murdered, what is the chance that he did it?

The scholar of risk Gerd Gigerenzer [had a go at answering that](#). The base rate for murders among American women is about five in 100,000. Assuming the one-in-2,500 probability is correct, then of 100,000 women with abusive husbands, about 99,955 will not be murdered. But of the remaining 45, 40 will be murdered by their husbands. The correct probability that the husband did it should be nearly 90%. Bayesian thinking might have helped convict Simpson.

And although he does not use full Bayesian reasoning but a slimmed-down version, the insight that we should update our existing probability estimates with new information — [cumulative monitoring, rather than one-off testing](#) — could, argued the statistician Professor Sir David Spiegelhalter, have spotted both the catastrophe at the Bristol Royal Infirmary and the murders of Britain’s worst serial killer, Harold Shipman, earlier, and saved many lives.

The Lucy Letby case does not turn solely on statistical evidence, and I make no arguments about whether she is guilty or innocent here. But people, including juries, prosecutors and judges, have misunderstood probability in the past; Sally

Clark and Lucia de Berk had their lives ruined by that misunderstanding.  
Thinking like a Bayesian might have helped prevent that.

---

Tom Chivers is a science writer. His second book, *How to Read Numbers*, is out now.

[TomChivers](#)