



## CONVICTION BY NUMBERS

Statistics have the power to trip everyone up — including judges and juries. Even when extra care is taken to get the numbers right in court, confusion often reigns. **Mark Buchanan** reports.

In March 2003, when nurse Lucia de Berk faced trial in a Dutch court for charges of murder and attempted murder, the statistical evidence against her seemed compelling. Investigators had identified a number of ‘suspicious’ deaths and near deaths in hospital wards in which de Berk had worked from 1999 to 2001, and records showed that she had been present when many of those events took place. The statistical expert testifying in the case, Henk Elffers, reported that the chance that her presence was mere coincidence was only 1 in 342 million.

On the basis of this number and on limited forensic evidence — traces of toxic substances found in two of the exhumed bodies — the court found de Berk guilty, and sentenced her to life in prison. But some Dutch mathematicians now say that the figure cited was incorrect, and that the case is a classic example of how statistical reasoning can go horribly wrong. “The magical power of the big number led everyone at an early stage to be totally convinced of Lucia’s guilt,” says mathematician Richard Gill of the University of Leiden in the Netherlands. “Then they went to work to confirm their theories.”

A court of appeal later upheld de Berk’s conviction, but an advisory judicial committee has now been appointed by the central office of public prosecutors to reassess the case. The committee’s decision is expected later this year, and it could recommend that the case be reopened. But whatever the result, the case illustrates the ongoing difficulty of ensuring that courts use statistical reasoning properly. “This is a serious problem,” says statistics professor Philip Dawid at University College London, “but there is no easy solution.”

### Learning from experience

The case also raises questions over whether courts have learned anything from past misuse of statistics. In a high-profile British case in 1999, Sally Clark was convicted of murdering her two small children based, at least in part, on faulty statistical reasoning by an expert witness — the paediatrician Roy Meadow. Dawid was invited to submit written evidence on the statistical arguments during Clark’s first appeal. He was not, however, allowed to present oral testimony. “The lawyers and the judge argued that it was not

rocket science, so a statistical expert was not needed.”

The appeals court confirmed Clark’s original conviction, but Dawid says that the written judgement showed that the jury had failed to grasp the relevant statistical issues. A second appeal in 2003 freed Clark, concluding that the jury might have been misled in part by the statistics presented in the original trial. Without a change in legal attitudes and procedures, mathematicians worry that statistical arguments will continue to be misinterpreted by the courts.

In de Berk’s case, Gill and another Dutch mathematician, Peter Grünwald of the National Research Institute for Mathematics and Computer Science in Amsterdam, have submitted letters to the judicial committee arguing that the original figure of 1 in 342 million was incorrect — or, at best, irrelevant to the proceedings. De Berk first became a suspect when working at the Juliana Children’s Hospital in The Hague. But because observations from this ward were the source of the initial suspicion, Gill and Grünwald say, those observations should not have been used in a calculation that tested the validity of the suspicion.



Elffers, of the Netherlands Institute for the Study of Crime and Law Enforcement in Leiden, combined the Juliana data with data from another hospital in which de Berk had previously worked to get his figure. Gill and Grünwald insist that the analysis was misleading. "It makes little sense to do formal hypothesis testing when the data themselves have suggested the hypothesis," says Gill. "The only safe thing is to go out and collect new independent data."

Gill's own calculation estimates that the probability that the correlation arose by chance is not 1 in 342 million, but a much smaller 1 in 48, or even as low as 1 in 5 — figures that are unlikely to meet the 'beyond reasonable doubt' needed for a criminal conviction. But Elffers defends his original calculations, arguing that he applied a factor that corrected for his using some of the data twice. "Everyone is aware that I applied a correction," he says.

**Fact or fallacy?**

Aside from this debate, equally important is how the court interpreted the number. Philosopher of science Ton Derksen of the University of Nijmegen, who has written a book that criticizes de Berk's conviction, argues that the court made an elementary statistical error known as the prosecutor's fallacy.

The court needs to weigh up two different explanations: murder or coincidence. The argument that the deaths were unlikely to have occurred by chance (whether 1 in 48 or 1 in 342 million) is not that meaningful on its own — for instance, the probability that ten murders would occur in the same hospital might be even more unlikely. What matters is the relative likelihood of the two explanations. However, the court was given an estimate for only the first scenario. Without additional information, says Derksen, Elffer's number is meaningless — and could easily be misinterpreted as a very small chance that de Berk is innocent.

To raise further doubt, other important statistics were neglected by the court. When de Berk worked at Juliana between 1999 and 2001, there were six unexplained deaths in her unit. The

same unit, in a similar period before de Berk started working there, had seven unexplained deaths. "It seems very strange," says Grünwald, "that fewer people die when there is a serial killer around." Derksen says that the statistics comparing deaths before and after de Berk started work at the hospital were mentioned by her defence lawyers, but were not sufficiently emphasized to have any influence on the court.

**Due process**

This neglect illustrates a difference between legal and scientific processes. Although science aims to bring together all relevant evidence, this is not necessarily true with the law. David Kaye, an expert in statistics and the law at Arizona State University in Tempe, notes that lawyers have an incentive, and even a duty, to select the evidence that makes their case stronger. "What the judge ends up hearing often comes from the two extreme ends of the distribution," he says.

Procedures to correct such distortions are also lacking, even after a trial has reached a verdict. In the United States, written statistical arguments are often protected by court orders, and so are not available for review or correction. "The data pertaining to an individual deserve some protection," says statistical expert Joseph Gastwirth of George Washington University in Washington DC, but a summary of the expert reports should be made publicly available, he suggests.

Independent scientific comment of this kind occurred during the Clark case, but to unknown effect. In the 1999 Clark trial, Meadow testified that the chance of two infants from the same mother dying of Sudden Infant Death Syndrome (SIDS) was only 1 in 73 million. Two years later, after the first appeal, the Royal Statistical Society in London, condemned both this figure and its interpretation. The figure would be valid only if SIDS cases arise independently within families, the

statement said, whereas there may be unknown genetic or environmental factors that predispose families to SIDS.

After the Clark case, the society established a working group, chaired by statistician Colin Aitken of the University of Edinburgh, to examine how courtroom use of statistics might be improved. In the immediate future, the society hopes to help provide continuing education to practising lawyers, judges and other legal practitioners, so that they can at least recognize the potential hazards in statistical reasoning. Ultimately, Aitken suggests, including statistics in the core curricula of law degrees will be more effective. "Things will not improve overnight," says Aitken, "but we are in it for the long haul."

**"The magical power of the big number led everyone at an early stage to be totally convinced of de Berk's guilt."**  
— Richard Gill

**A matter of opinion**

In the United States, the Federal Judicial Center, an organization created by Congress to improve federal courts, has published a reference manual on the use of scientific evidence, with one chapter devoted to statistics.

"But education is only palliative," says Kaye, who helped to write the statistics chapter. "I don't think there is any single way to ensure that statistics and other scientific evidence gets used accurately."

Although courts expect one simple answer, statisticians know that the result depends on how questions are framed and on assumptions tucked into the analysis, giving tremendous room for legal argument. Kaye recalls giving lawyers in one case what he thought was a crystal-clear explanation of a statistical argument. "Their response was 'Let's just put him on the stand, he will confuse everyone.'"

Indeed, the biggest practical challenge, some argue, lies in the unusually subtle nature of statistical reasoning, which research shows confuses experienced professionals, such as physicians, just as easily as the general public. In de Berk's case, even Elffers now suggests that with so much uncertainty swirling around the statistics, they should play no further role in the considerations. "The review committee should concentrate on non-statistical arguments," he says.

Those arguments are also a matter of dispute. Until the committee decides, de Berk languishes in prison, quite possibly because of mathematical errors. "I am so convinced that the statistics are wrong," says Gill, "that I am more inclined to believe in the incompetence of the entire process than in the existence of a serial killer." ■

Mark Buchanan is a science writer based in Europe.

K. WIGGLESWORTH/EMPICS



Sally Clark was released on appeal after questions were raised about statistical arguments related to her conviction for murder.



Lucia de Berk is still waiting to hear whether a judicial review of her conviction for murder will recommend that her case be reopened.

DE BERK'S FAMILY/DE TELEGRAAF