

The Fabrication of Facts: the Lure of the Incredible Coincidence

Ton Derksen (6 juli 2007)

Content:

PART I

- (1) Introduction: The charge, the conviction and three reasoning instincts
- (2) The smoke and the fire
- (3) A chance that was too small
- (4) The two little engines that could not
- (5) The Incredible Coincidence that did
- (6) The Incredible Coincidence and the Fabrication of Facts
 - (a) the nurse that complained about Lucia on the 4th of September 2001
 - (b) on the 4th of September 2001 a list of Lucia-related incidents circulated in the hospital
 - (c) the coordinating paediatrician Van M.
 - (d) the police
 - (e) the managing-director of the Juliana Children's Hospital mr. S goes statistics
 - (f) the independent research team of the Julian Children Hospital
 - (g) medical experts before the courts
 - (i) paediatrician professor V.
 - (ii) the surgeon L. who treated mrs. S.
 - (iii) paediatrician S. at the Juliana Children's Hospital
 - (iv) medical-director S. of the Free University Medical Center
 - (v) paediatrician D. at the Juliana Children's Hospital
 - (h) the prosecution before the court of first instance
 - (i) the court of first instance
 - (j) the prosecution before the court of appeal
 - (k) the court of appeal
- (7) Conclusion of Part I

PART II

- (8) The Incredible Coincidence Argument in its statistical version
- (9) Problems with the data
 - (i) the selection and collection of the data was biased
 - (ii) the data were restricted to three wards and 1 ¼ years
 - (iii) the data were used twice, both in the formulation of the hypothesis and in the testing of the hypothesis
- (10) Problems within the method as used
 - (iv) multiplication of the chances with respect to the different wards is wrong
- (11) Recalculation with Fisher's Exact test
Three ignored incidents outside Lucia's shift. Chance of 1 in 44.
- (12) Problems with the method
 - (v) the wrong question is being asked
- (13) Other Relevant Data: Total Collapse of the Case
A different statistics: fewer deaths while Lucia worked in the Juliana Children's Hospital than during the years before her presence
- (14) Conclusion

PART I

(1) Introduction: The charge, the conviction and three reasoning instincts

The nurse Lucia de B. was convicted by the Court of Appeal¹ in The Hague for seven murders and three attempted murders with the penalty of life imprisonment. She had been charged by the prosecution with thirteen murders and four murder attempts. The police examined a total of some thirty allegedly suspicious cases. Yet there was little evidence apart from the fact that at the hospital where it all started, quite a number of reanimations happened during the shifts of the nurse. Then, after one more reanimation it was felt that that all this could not be just a coincidence.

Very early in the process a quantification was given to the uneasy feeling that proportionally there were too many reanimations during the nurse's shifts. The probability that such a coincidence could have happened by mere accident, was calculated as 1 in 7 billion,² and somewhat later as 1 in 342.000.000. With that number out in the open the general notion was that what had happened could not be just an accident. The nurse definitely had to be a serial killer, even if no one had any further evidence.

This notion was in fact the driving force in the whole judicial process. It coloured the perception of the hospital, the medical experts, the prosecution and both the court of first instance and the court of appeal. And it fabricated a whole series of incriminating facts which inexorably led to the conclusion.

In the first part of this paper we will document the driving force of the Incredible Coincidence, the disbelief of people that the spectacular coincidence was *just* a coincidence. We also examine the fabrication of the facts that resulted from this disbelief. In the second part we expose the statistical elaboration of this idea as it is presented before the court. We shall detect that also in the statistical elaboration the facts have been manhandled. Careful attention to the facts would have yielded a very different outcome. Actually, when all the facts are in, the coincidence shrieks drastically. Nothing particularly shocking remains. So there is every reason to think that the remaining coincidence is just that — a coincidence.

Throughout the process we see three reasoning instincts at work:³ (1) the Small Chance Instinct which makes us neglect small chances that p and automatically turn to the belief that non-p; (2) the Smoke and Fire Instinct, which caused colleagues of the nurse to get worried about those many incidents during her shifts, and which later on caused the general public to be content with the verdicts of the courts in spite of the shaky evidence (someone convicted of so many murders must be a murderer); and (3) our instinct to neglect base rates (in not-everyday situations). These instincts are natural in the sense that they guide our cognitive housekeeping. They yield quick conclusions which usually are evolutionarily good enough. Unfortunately, in the case of Lucia de B. their influence was disastrous.

¹ In the Dutch legal system, courts of first instance determine both the facts and the extent of the punishment. There is no jury. After a conviction appeal is possible at a court of appeal. This court has a second look at the (alleged) facts and determines the degree of punishment all over again. A further, higher appeal at the supreme court is — in essence — only about the proper legal procedure. The prosecution presents a *requisitoir* (final argument) to the court and after the reaction of the defense lawyers (*pleitaantekeningen*, pleadings) they present their reply (*repliek-aantekeningen*, rebuttal of the pleading) to which the lawyers reply in their turn with their *Dupliek*.

² This is the American billion, to wit 7.000.000.000.

³ They are related to the heuristics of Nisbett and Ross (1980, *Human inference: strategies and shortcomings of social judgment*, Prentice-Hall: Englewood Cliffs, New Jersey I prefer the term 'instinct' as the term 'heuristic' suggests too much conscious awareness of the operation.

(2) The smoke and the fire

The nurse Lucia de B. was convicted for seven murders and three attempted murders. So what was the evidence? No one saw the nurse do anything suspicious, there were no needles or bruised skin or any other incriminating evidence, there were just incidents during the shifts of the nurse. But there were lots of them, too many — or so it was felt.

In spite of the public awareness that there was little evidence, the reaction was: with so many alleged murders and so many suspicious incidents, the nurse must have committed at least some murders. The prior concern was that society was protected. So it did not matter that in many cases the proof was somewhat shaky and not beyond reasonable doubt. There was so much smoke, there must be a fire.

Here we see a human reasoning instinct in full operation: Where there is smoke, there must be a fire. A careful judicial examination may make a mistake or two but it cannot have been mistaken in all ten (attempted) murders convictions. Moreover the court had dropped seven (attempted) murder charges. So it must have looked carefully into the cases. Let us rejoice that a murderer is behind bars and let us not worry about some small flaws in the process.

Earlier the same Smoke & Fire Instinct had done its work in the hospital: Lucia was present at so many incidents (reanimations) within such a short period of time (two years and especially during the last year as a qualified children's nurse) that people became suspicious. Actually, Lucia herself had become distressed by all those reanimations during her shifts and she had talked about it with her superior and the hospital's social worker. The official reaction had been of a practical nature: such a concentration of incidents happens often and it happens to almost every nurse once in a while, so don't worry, it will pass. The problem for Lucia was that the series of incidents did not stop in time. When on the 4th of September 2001 during Lucia's night shift yet another child had to be reanimated and died, a colleague told her superior in the morning that she was worried because 'during the two years that Lucia had worked at the Juliana Children's Hospital she had been present at nine reanimations'⁴ and she also remembered an incident from a previous hospital in which Lucia had been involved. Though there was no incriminating evidence whatsoever against Lucia — and the incident mentioned had nothing to do with Lucia — , the train took off at that moment. At the end of the day the police had been informed and the next day the interrogations started. There was too much smoke, there had been too many incidents during her shifts.

(3) A chance that was too small

Apart from this Smoke & Fire Instinct there was an idea that soon became the driving force of the process, namely the idea that the chance that the concurrence of all those reanimations during one nurse's shifts was just a coincidence, was so extremely small that it was incredible that the coincidence was just that — a coincidence.

This chance-argument became explicit in the thinking of the managing-director of the hospital, mr. S. He made some statistical calculation [an

⁴ As reported by Tony B., head of intensive care PV 5 September 2001. B. herself told the police: 'I can tell you that speaking from my experience nine reanimations in medium care are much for one nurse' and 'I can tell you that in our hospital it was generally known that Lucia had been present at many reanimations'.

amateurish calculation of sorts, as he himself called it] and came up with an extremely small chance. After that everybody knew for sure: the coincidence is not just a coincidence, Lucia must have had a hand in the incidents.

Let us remind ourselves here that it is quite common that people die in hospitals and that, as the social worker knew and all nurses know, sometimes there will be an accumulation of deaths during somebody's shifts.⁵ Should we conclude that all of them are serial killers? So what was the specific argument in Lucia's case, or rather what made the minds lean so quickly and so inexorably in one direction?

The answer is — we think — the *Small Chance Instinct*, a basic human intuition based on the Small Chance Principle:

when the chance is very small that p (where p stands for some statement), then it is reasonable to believe that not- p .

It is a principle that works in everyday situations. For example, when we go to the beach and we see written in the wet sand a message "I'll be back in a moment", we do not consider the possibility that this is a freak-cooperation between the sand and the forces of wind and water. The chance that such a message would be formed, is so small that we may reasonably neglect it.

It is not just at the normative level that this Small Chance Principle operates. It also operates in the domain of actual belief. Doubt in such a case does not occur — unless the risks involved get very high. The winner takes all. This is a basic principle of actual belief which parallels the normative Small Chance Principle.

As long as there is no specific reason for distrust, we automatically use the Small Chance Principle, and we feel justified in applying it. We will see that the principle needs to be supplemented. The Small Chance Instinct which makes us neglect the small chances (that is, makes us apply the Small Chance Principle automatically) is not, however, "aware" of these complexities and conditions for proper use. The Small Chance Instinct may well lead us astray. We will see that the Lucia-situation is one in which this instinct ran amok.

During Lucia's trial the Small Chance Instinct ran the show. The chance that all those incidents happened during one person's shifts just by accident, was too small for comfort. That is what the nurses, the doctors, the manager-director of the Juliana Children's Hospital (The Hague) believed and this is what in the end the prosecution, the court of first instance and the court of appeal believed.

The hospital installed its own reaserch team that reexamined all cases of death since Lucia's entrance at the JKZ. The team found ten suspect incidents (reanimations, deaths, intoxications) during one year, and the nurse was present during all of them. The extremely small chance that this would happen by accident, while the nurse worked only 142 of all 1029 shifts, made it incredible that what had happened was *just* an accident. On the 17th of September, thirteen days after the last incident, the media were informed by the hospital and excuses were offered to the parents, leaving no doubt in the public's eye that the hospital had spotted a monstrous murderer.

⁵ For instance, a nurse wrote in the Dutch newspaper NRC (20 March 2004) that during her two-year training period she was present at 30 cases of death, while her friend did not see any death.

We do not know what actual number occupied the minds of the people in the hospital, but we do know that the police soon talked about a chance of 1 in 7.000.000.000. The police told Lucia's family: "Look, the chance that Lucia has not committed these crimes is 1 in 7.000.000.000, and there are only 6.000.000.000 people on this world. So she must be the murderer".

In the meantime three other hospitals in The Hague, where Lucia had worked previously, had been warned. They were asked to investigate whether during Lucia's shifts there had been inexplicable incidents.⁶ And before long these hospitals complied. The Red Cross Hospital found that seven deaths during Lucia's shifts turned out to be suspicious once their status was reconsidered. The Leyenburg Hospital discovered two cases serious enough to be prosecuted, while twelve (old) deaths raised enough doubt for the police to look into these cases. The Penitentiary Hospital brought in one more death.

Such a coincidence had never been shown before. The assistance of an official statistician was called in. The data of three wards were sufficiently reliable for statistical treatment (or so it was said). Then the expert calculated the chance that Lucia would — by accident — have been present at "her" incidents at these three wards. He found the shockingly small chance of 1 in 342.000.000. Then everybody knew for certain that Lucia must be a dangerous serial killer. The idea that the concurrence was just an accident, was preposterous. Lucia herself must have caused this abundance of incidents.

The press described her as the Angel of Death. Her name was even mentioned in international newspapers. And what about the hospital that committed itself to five murders and five attempted murders, what about the parents whose wounds had been reopened, and what about the newspapers that had been notified about a serial murderer in Dutch hospitals? One had to come up with a killer and because of the incredible coincidence everybody knew before the trial even started that Lucia was the killer.

There was one problem, and a serious problem at that: there was very little evidence, if any, beyond the spectacular coincidence. So the prosecution set out to collect independent evidence, yet all the time it was the Incredible Coincidence (1 in 342.000.000) that bewitched the minds and was the driving force behind the process

(4) The two little engines that could not

There were loads of suspicions but little evidence beyond the concurrence of all those incidents during the nurse's shifts. The prosecution had pinned its hope on two cases that stood out: only in those two cases the prosecution managed to specify the (alleged) poison of the (alleged) intoxication (digoxin, chloral hydrate) and only in those cases they could give some evidence for their charge of intoxication — or so it was claimed. The other cases were merely medically inexplicable incidents, which either had been classified as natural deaths at the moment of death, or as a not-suspicious incident. The prosecution called the two allegedly strong cases their two locomotives which had to pull the other cases (the wagons) to the station called "proof beyond reasonable doubt".

⁶ Mr. T., board member of the Leyenburg Hospital, testimony 10 February 2004: 'We started our search on request of the police and the prosecution. Questions to be asked were: notification of all patients who died during the shift of the suspect or just after her shift, reanimations during the shifts of the suspect or just after her shift'.

In *Lucia de B. : a reconstruction of a judicial error*⁷ Ton Derksen argued in great detail that these two little engines could not. The allegation of the digoxin intoxication is fraught with problems. During the time that the nurse is supposed to have administered the poison two medical doctors were examining the baby.⁸ And the charge of an *acute* digoxin intoxication is refuted or at least made severely questionable by the non-contracted state of the heart of the baby at the obduction⁹ and by the absence of any digoxin in the liver while digoxin was found in the kidneys.¹⁰ This is considered an impossible (extremely unlikely) distribution of digoxin after one hour. So the locomotive driven by digoxin intoxication cannot even get going, let alone pull some eight wagons.

The second engine used an alleged chloral hydrat intoxication as fuel. But since the hospital prescribed the maximum amount (625 mg) for the boy in question, and allowed that two extra doses of that amount might be added — in case of restlessness — and the nurses complained about the restlessness of the boy, there is a real possibility of overmedication. Actually, there is evidence that at two consecutive days, shortly before the incident, the boy received a double dose. So we need much more than Lucia's presence to convict her of an attempted murder.¹¹

So the two engines that should, could not. It was the Incredible Coincidence that in fact carried all the weight.

⁷ Ton Derksen, *Lucia de B. : een reconstructie van een gerechtelijke dwaling*, Veen Magazines, Diemen, 2006.

⁸ The court of appeal noticed that the monitors of the baby-patient Amber had not worked during the very 30 minutes during which — in their reconstruction — the poison (digoxin) must have been administered, and that Lucia was the nurse responsible for the baby patient. So the court concluded that Lucia must have switched off the monitors during those thirty minutes to make sure that she could safely commit her hideous crime. This coincidence (monitor not working during the crucial period and Lucia being in control of these monitors) could not have been an accident! But this argument is based on readings of the trend tables of the monitor. Since the trend table indicates the readings only every fifteen minutes, there is an apparent time-opening of thirty minutes when one such reading drops out. The court uses this time-opening to settle the time of the examination of the baby by two medical doctors. But the continuous trend graphs tell us that the actual time-opening is only six minutes (time for a diaper change), much too short for the examination which lasted some twenty minutes. So we have to find another time-frame for the medical examination. According to the trend graphs there is just one such period, but that is the very period that the alleged poison administration had taken place according to the reconstruction of the court. So we find that at the very moment of the alleged murder two medical doctors are quietly examining the baby girl, and that the charge against Lucia is based on wilfully restricting the attention to the trend tables and by neglecting the trend graphs.

⁹ In the case of acute digoxin intoxication the heart is contracted. At the obduction, some nine hours after the demise this should still be the case. But the heart was not-contracted, according to the coroner.

¹⁰ In June 2004 the remaining tissue and blood was examined by the 'Institut de Médecine Légale et de Médecine Sociale' (Stasbourg). It used a highly digoxin-specific HPLC-MS method. It found a digoxin level of 10,2 ng/ml in the kidney tissue, but repeated examination did not find any digoxin in the liver. Assuming that a fatal dose had been administered 60 minutes before the baby's death, this is an extremely unlikely distribution of digoxin in the organs. Thirty minutes after an intravenous administration still 50 % of the digoxin is in the blood. Gradually most of it disappears to the organs, especially the heart, the kidney and the liver. So after an injection with digoxin, after 30 minutes, one should expect a fairly high concentration in both the blood, the kidney and the liver. The fact that nothing was detected in the liver suggests that the digoxin found is the remainder of therapeutic use. (The girl had been treated with digoxin therapeutically). (The digoxin concentration in the kidney is way too small as well).

¹¹ This case is especially shocking because Lucia was the only one who was worried all day long, trying to convince the doctors to come and look at the patient. During the day two assistant doctors appeared and did not do much. When in the afternoon the medical specialist appeared she was really annoyed that she was not called earlier. (However, Lucia, being just a nurse, was not allowed to call the specialist).

(5) The Incredible Coincidence that did

If anything is responsible for the conviction it is the Incredible Coincidence, the coincidence of all those incidents during the shifts of one nurse. The chance that this would happen by accident, is extremely small. In that sense the coincidence is spectacular or *incredible*. How small is the Incredible Coincidence? It does not matter much, it is extremely small anyway, whether it is 1 in 7.000.000.000, 1 in 342.000.000 or 1 in 9.000.000.¹²

The coincidence is incredible in another sense as well: the chance that this spectacular/incredible coincidence would happen by accident, is thought to be so small that that it is *incredible* that the coincidence is *just* an accident (just a coincidence).

We have already seen the underlying argument which depends on the Small Chance Principle:

the chance that the coincidence of all those incidents (reanimations, deaths) during Lucia's shifts is just an accident, is so small that, by the Small Chance Principle, we may reasonably neglect that chance and reasonably believe that the coincidence is not just an accident (just a coincidence).

According to this principle we may then reasonably believe — and due to the Small Chance Instinct we will actually believe — that the coincidence is not just a coincidence. There must be some specific cause.

The statistical expert-witness reminded the courts that this does not imply that Lucia must be the cause of the incidents. There may be another cause. For instance, (1) someone who worked the same shifts as Lucia, may be responsible, or (2) Lucia worked many night shifts and more deaths occur during the night. But these alternative options were excluded quickly by the court — and by almost anybody else. So the practical and actual conclusion of the Coincidence Argument is: the coincidence is too incredible, Lucia must be the murderer.

(6) The Incredible Coincidence and the Fabrication of Facts

This Incredible Coincidence (combined with the Small Chance Instinct) was the driving force behind the process which led to Lucia's conviction for seven murders and three attempted murders. We will illustrate this by giving a brief history of the fatal influence of the Incredible Coincidence. It led to the fabrication of many incriminating facts.

(a) the nurse that complained about Lucia on the 4th of September 2001

We already wrote about the nurse for whom the death of that night was one too many. Her superior to whom she confided her suspicions, summarized her worries as follows:

she really had the feeling that this could not just be a coincidence any longer^{13 14}

(b) on the 4th of September 2001 a list of Lucia-related incidents circulated in the hospital

¹² All these ratios have been mentioned and taken seriously.

¹³ Tony B., head of intensive care, PV 5 September 2001.

¹⁴ PV Moniek 10 September 2001: "If I consider how many reanimations I had during all those years during my work and if I compare this with the number of reanimations that Lucia had, then that just does not add up".

The baby Amber died in the early morning of the 4th of September 2001. Already that very morning a list with nine (suspect) reanimations during Lucia's shifts circulated through the wards.¹⁵ And this influenced many of the nurses. On the 6th of September 2001 nurse Martine told the police that 'she had an uneasy feeling about it. It was all too accidental'.¹⁶ On the 11th of September 2001 nurse Ingrid declared:

If I am honest about it, I think it is too much of a coincidence that within a few days time two of Lucia's patients had to be reanimated.
17

She also reported that she overheard other nurses saying: "again Lucia".¹⁸

So we may assume that it is the coincidence (expressed by this list of reanimations — and already known by many) which caused the general feeling of uneasiness about Lucia. For there were no concrete facts concerning the Juliana Children's Hospital which could be referred to.¹⁹

Some nurse claimed to remember an incident in a previous hospital. This incident was examined, and it was just a mix up. Someone else remembered some death in yet another hospital due to a not properly working infusion. Further examination led to Lucia's conviction of murder by the court of first instance. However, during the process before the court of appeal it turned out that during the alleged incident Lucia was not in the hospital, actually she had been sick for three days. The murder charge was dropped, and the qualification "murder" was dropped as well. The alleged murder became a case of natural death again, as it was before Lucia was accused of murder. No Lucia, no murder.

(c) the coordinating paediatrician Van M.

Already before 16.00 hour on the same 4th of September the coordinating paediatrician Van M. notified the police about (attempted) murders. Apart from the death of Amber, whose natural death qualification had been changed into a non-natural death during the day, he informed the police about 'five similar [inexplicable] deaths'.²⁰

(d) the police

That very evening the police visited the hospital. On the 5th of September the interrogations of the nurses started. Soon the police used the chance of 1 in 7 billion that Lucia has not committed the murders. There being only 6 billion people, Lucia had to be the murderer. We need not comment on the crooked reasoning, but it is intriguing to see how quickly also the police caught on: the coincidence was too large to be just a coincidence.

(e) the managing-director of the Juliana Children's Hospital mr. S goes statistics

¹⁵ Tony, PV 5 September 2001: "I can tell you that in our hospital it was generally known that Lucia had been present at many reanimations. I do not know whether the number nine is correct. Marianne has told me that number and it is this number that circulated in the hospital".

¹⁶ Martine, PV 6 September 2001.

¹⁷ Ingrid, PV 11 September 2001.

¹⁸ For instance K.B. team leader of Medium Care Unit-1, PV 5 September 2001: "I had already witnessed a number of other incidents of Lucia. It went through my head once or twice "again Lucia, those incidents happen really often around her".

¹⁹ Although there is a coincidence of Lucia's shifts and incidents, Lucia was present at only one death.

²⁰ B., PV 4 September 2004..

During the evening of the 4th of September 2001 manager-director mr. S. made some calculations.

Using a computer program I have combined the number of deaths with a specific period and the shifts of Lucia and have done a statistical calculation. The result was that Lucia was involved in an extremely unlikely high number of these incidents.²¹

To the police he stressed:

With this calculation nothing further happened. It only contributed to my own conviction to notify the police.²²

Well, that is not nothing! The statistical calculation contributed to the notification of the police of five possible murders and five possible attempted murders. The Incredible Coincidence made a huge difference.²³

(f) the independent research team of the Juliana Children's Hospital

The managing-director S. did do more. He set up a committee to 'separate the rumours from the facts'.²⁴ But also in this committee the Incredible Coincidence played a crucial role. The story is that the committee first examined each death and reanimation since Lucia worked in the Juliana Children's Hospital (without knowing if Lucia was on duty) to see whether the death (reanimation) was non-natural after all, in spite of the fact that immediately after the death, all deaths had been declared natural. The committee came up with ten cases, five non-natural deaths (possible murders) and five non-natural incidents (reanimations, intoxications) (attempted murders). Subsequently the committee checked which nurses were on duty. It found out that Lucia was always present. After finishing the research, the committee asked an outside expert to evaluate their findings.

There may well be some doubt about the alleged anonymity with which the deaths and reanimations were examined. Remember that at the very start of the crisis in the hospital the paediatrician Van M. informed the police that beside Amber there were already five other suspect cases, and that at that time a list of nine Lucia-related incidents circulated in the hospital. So it is hardly credible that the research team consisting of doctors of the ward did not know whether Lucia was involved or not, when they examined a specific incident. And indeed, there is further indirect evidence that they knew and that their judgement, however sincere, was influenced by the Incredible Coincidence

One would think that the issue here was the possible non-natural status of the deaths and reanimations. Yet the outside expert who was consulted,²⁵ professor V., a paediatrician himself, tells us differently and very illuminatingly.

²¹ PV 17 September 2001.

²² PV 17 September 2001:

²³ Incidentally, we do not know why this general manager incriminated Lucia in the case of Ahmad (death) by telling the police that 'the evening shift ends at 23:30 hour... [and that] Lucia, having evening shift that day, was presumably present at the death, in view of the moment of the death (23:30 hour)'. (PV 17 September 2001). As all the nurses knew the evening shift ended at 22:45 hour and there was a transitional period till 23:15 hour. Her colleague told the police that that very evening Lucia left early (around 23.00 hour).

²⁴ PV 17 September 2001.

²⁵ He was not quite such an independent expert as he knew the leader of the research team very well indeed. This questions the independence, not the integrity of his judgement.

*The first question that I was asked was: is it more than a coincidence that five children died and that four children had to be reanimated?*²⁶

The next worry is that, with the Incredible Coincidence apparently at the back of their mind, the committee moved more lightly to a non-natural death (and reanimation) when Lucia was involved and that, being very worried about a possible serial killer in the hospital, it thus fabricated some new facts. There is some indication that this happened.

For example, the committee claimed that the death of baby Amber must have been non-natural because her heart was in excellent shape. This is a "fact" that continued to play a role in the discussion. But at the moment that the committee took the excellent condition of Amber's heart for granted, the coroner had already written that the heart was dilated (which indicates heart weakness) and that the death might have been caused by heart problems.²⁷ He declared that for him "it was hard to understand that the declaration of a natural death was changed into a declaration of non-natural death some hours later".²⁸

Another fact was fabricated when the committee decided that the death of baby Ka was non-natural because Lucia had shown non-professional behaviour. She combined two actions, namely measuring blood pressure and spraying. None of the nurses present objected to this behaviour; actually it was fairly common. So the judgement that Lucia was involved in a non-natural death because of allegedly non-professional behaviour, is based on a fact that was manufactured by the committee itself.

There is also some doubt about the consistent judgement about incidents, the outcome depending on Lucia's presence or absence. For instance, the committee decided that the reanimations of Kemal on the 20th December 2000 and on the 1st March 2001 were suspect. But Kemal's reanimation on the 10th October 2000, which was very similar to that of 20 December 2000 (during the reanimation Kemal's heart never stopped beating and Kemal did not need any hart massage) remained a natural incident. Why? The only difference is that Lucia was not present. The reanimation is never mentioned and it was not counted among the statistical data of incidents outside Lucia's shifts, a crucial discriminating category in Fisher's Exact Test. Whether a reanimation is suspect seems — at least — in part to depend on the absence or presence of Lucia.

Another case: why is Ahmad's comatose state on the 25th of January 2001 due to a choral hydrate intoxication suspect, while Kemal's comatose state on the 6th June 2000 due to chloral hydrate is not? The concentration of chloral hydrate was not measured in Kemal's case but the visiting doctor expressed his worries in the medical files. Again, does Lucia's absence make the difference?

One more case: why did the committee call the alleged incident of Sadia on the 18th January 2001 suspect when she only stopped breathing for less than a minute and started breathing after some physical assistance? Was it because Lucia was present? No one else saw anything suspicious here. Yet the hospital notified the police of a possible non-natural reanimation.

²⁶ Testimony 12 February 2004.

²⁷ Testimony 19 February 2004, p. 11: "My conclusion is ...that it is not wholly excluded that the serious congenital defects [of the heart] are the cause of the death".

²⁸ Testimony 19 February 2004, p. 11.

Here we run into another general cause for concern: as regards each of the allegedly suspect incidents (suspect according to the hospital) there were always medical experts who have declared before the court that the death (reanimation) was natural rather than non-natural. Actually, in most cases most medical experts believed in a non-natural death or reanimation. So the judgement that the incidents were suspect is not undisputed and hence not obvious and far from objective. For the same reason the allegations against Lucia lack an objective base.

In its testimony before the prosecutor the research team emphasized that the examination was not suspect-driven. We do not question the good faith of the committee, but when we look at the actual situation it is hard to ignore a strong influence of the Incredible Coincidence.

(g) medical experts before the courts

The court's judgement was guided by the judgements of the medical experts. But at least some of the crucial medical judgements turned out to be co-determined by the Incredible Coincidence, they were not just medically motivated. Thus the Incredible Coincidence haunts even the arguments of the medical experts. We will give five examples in which the coincidence is explicitly mentioned.

(i) paediatrician professor V.

Professor V. was asked to examine the results of the Juliana Children's Hospital research committee, but he also was an expert-witness before the court of first instance and the court of appeal. Speaking about all of the incidents in the Juliana Children's Hospital he testified:

The point is here that it happened so often with so many children. ... Each incident viewed individually, it is possible and cannot be excluded that a child with such a medical history dies suddenly without a clear cause for the death.²⁹

That is to say, looked at individually the deaths and reanimations would not have raised enough questions to notify the police. Individually each case could be understood as natural. Only in combination the deaths and reanimations turn suspect.

Professor V. explicitly recognizes the importance of the Incredible Coincidence:

it is because of the medical situation of the children and the extremely unlikely statistical chance that one specific nurse was involved with every incident, that I deem it improbable that all these children have died a natural death.³⁰

That is, from a purely medical viewpoint there was not sufficient reason to regard the deaths as non-natural. When the court relied on professor V.'s judgement, as it did in many cases, the court did not then in fact rely on his *medical* expert knowledge. It relied on the Incredible Coincidence. Professor V. judged that six children in the Juliana Children's Hospital died a non-natural death. Since we have to remove the effect of the Incredible Coincidence, we have to discount six judgements of non-natural deaths.

(ii) the surgeon L. who treated Mrs. S.

²⁹ Testimony 12 February 2004, p. 14.

³⁰ It is an interesting question how Visser could be so sure of the extremely unlikely statistical chance so early on in the process. Testimony 12 February 2004, p. 14. (Requisitoir before court of first instance, p. 4)

Also surgeon L. relied in his medical judgement on the Incredible Coincidence. This is of crucial importance as his allegedly medical judgment that Mrs. S. died a non-natural death, is used by the court of appeal to back *its* judgement that the death of Mrs. S. was non-natural.

The court noted that on the 27th of August 1997 Lucia wrote in her diary: "Today I gave in to my obsession". That very day Mrs. S. died during Lucia's shift. Surgeon L. testifies that she died a non-natural death. So the court reasoned: the obsession is an obsession to kill very ill patients. Lucia denied this. She said the obsession was an obsession to lay tarot cards for patients. The psychiatrist and psychologist who examined her for six months, concluded that her explanation is plausible. The court has chosen not to believe Lucia. It was desperately in need of an obsession to kill as background support for all the convictions. But apart from the psychiatric report the court faces another major problem: six out of the seven medical experts in this case and two nurses state very clearly and without any hesitation that the death of Mrs. S. is a natural death. They may have been surprised at the precise time of death, but they did not have any qualms about the death itself.

The exception is the surgeon L. At the time of her death (she was his patient) he was surprised about *the exact moment* of her death. He declared before the court that he had expected her to live some more hours or days. Yet immediately after the death of Mrs. S. he gave a natural cause of death in his letter to the family doctor and told the head of the nurses Van B. that he had no doubts about the case.³¹ Only four years later, in 2001, he changed his opinion. He started to believe that the death was non-natural after all, when, as he clarified before the court of appeal in 2004,

*in the media attention was given to the inexplicable deaths in different hospitals in The Hague.*³²

However, note that this second and revised judgement is not a medically informed judgement. It crucially depends on the Incredible Coincidence.

That is, only after the incidents and inexplicable deaths had been in the newspapers, surgeon L. changed his mind. His later judgement that Mrs. S. died a non-natural death, is then determined by the force of the Incredible Coincidence. The medical facts that he knew were not sufficient for that judgment. Thus the alleged fact of Mrs. S. non-natural death is fabricated, and it is fabricated due to the Incredible Coincidence.

(iii) paediatrician S. at the Juliana Children's Hospital

When paediatrician S. ordered an extensive blood test while Ahmad was in coma, she made her reliance on the Incredible Coincidence quite clear:

*You ask why I had my suspicions. It was not so much the death of Ahmad in itself but the piling up of crises, which occurred in the presence of the nurse Lucia.*³³

*I ordered this blood test because I did not understand the situation and because I had my suspicion against the nurse present, namely Lucia, who according to my knowledge had been present at many inexplicable death cases of children at Medium care Unit 1.*³⁴

³¹ Pleading of the defense lawyers before the court of appeal (pleitaantekeningen), p. 92.

³² Testimony 11 May 2004.

³³ Testimony 24 September 2004.

³⁴ Testimony 24 September 2004.

We do not suggest that the blood test should not have been ordered, but we note the reliance on the Incredible Coincidence in the decision making of this doctor.

(iv) medical-director S. of the Free University Medical Center

The court of first instance had asked medical-director S. to comment on the medical condition of the elderly patients. Both the court of first instance and the court of appeal have judged that three of those patients have been murdered by Lucia. The problem is that almost all medical experts have declared before the court that these death were natural. In two cases medical-director S. agreed with the court that the deaths are (somewhat) suspect but in his argumentation he relied on the Incredible Coincidence:

Although it is not generally uncommon that patients from this (elderly) group will be found dead in their beds, there is — in my judgement — an (accidental?) concentration of this version of decease within a relatively brief period at one surgical ward.³⁵

(v) paediatrician D. at the Juliana Children's Hospital

Paediatrician D. was chairman of the research committee set up by the manager-director mr. S. She too was struck by the coincidence of the many incidents (reanimations in her case) and Lucia's shifts.

In the period that Lucia worked in our hospital, there was a striking number of reanimations in the ward Medium Care Unit-1.³⁶

And though – again - we do not cast doubt on the integrity of the medical doctor, we are surprised at the ease with which the conclusion of a non-natural death is reached. In the case of patient Jaouad doctor D. complains that

it is peculiar that the death has not been described by Lucia in the medical files.³⁷

But (1) we do not know whether it has not been described by Lucia as the relevant page is missing, and (2) another nurse claimed to have reported the death and complains to the police that the page has apparently been lost.³⁸ So, apart from the Incredible Coincidence, there is no reason whatsoever to hold Lucia responsible and imply that Lucia has something to hide. Yet it became an incriminating fact that was brought forward by the court of appeal in its argument for the conviction of Lucia.

(h) the prosecution before the court of first instance

The prosecution before the court of first instance stressed the importance of the Incredible Coincidence.

The picture painted by the witnesses was ever the same: the suspect was excessively often involved with cases of death and reanimations.³⁹

They quote the famous ratio of 1 in 342.000.000 purportedly describing the chance that the coincidence happened by accident. They add that though

³⁵ Report 12 March 2002.

³⁶ PV 7 August 2002

³⁷ PV 7 August 2002

³⁸ Compare Jenny PV 27 September 2001 (*Lucia de B*, p. 227)

³⁹ Requisitoir, p. 2.

it cannot be proved [on this base] that the suspect has in fact caused the incidents, this calculation can show that it cannot be just an accident that the suspect was present at all these incidents.⁴⁰

They quote with approval expert-witness De M.:

In view of the fact that the incidents occurred in four different hospitals and that no other cause can be inferred from the medical record and the testimony of the suspect, it has to be concluded that there is a causal relation between the occurrence of inexplicable cases of death and life-threatening incidents involving patients on the one side and the presence of the suspect in the four hospitals.⁴¹

The prosecution also fabricates a new fact. It states:

From the declaration of the manager-director mr. S. of the JKZ it can be inferred that in the period from 1996 till the present only five patients died at the ward MCU-1. At all five deaths the suspect was present. That is to say, during the last five years there have not been cases of death at which the suspect was not present. That cannot be just an accident.⁴²

But the five deaths from 1996 till the present is not a fact!! The ward changed its name on 9 November 1999 from IN1 (Internal-1) to MCU-1. And it *only* changed its name. If with this knowledge we look at the period from 1996 till 2001, we see a very different picture: in the three years 1996, 1997 and 1998, the years that Lucia did not work in MCU-1, there were 7 cases of death in that ward. In the years 1999, 2000 and 2001, years that Lucia did work in MCU-1, there were 6 cases of death in that ward.⁴³

What are we to make of the fact that in a period that a serial killer is supposed to be active in the ward, the number of deaths drops rather than increases sharply? The most plausible explanation is: there was no serial killer.

Note that *technically* the prosecutor does not lie. In the ward *with the name* MCU-1 there were just five cases of death.⁴⁴ But in the ward, which got the name MCU-1 in November 1999, there were 13 cases of death, 7 death in the years that Lucia did not work there and 6 while she did work there.

(i) the court of first instance

The court of first instance is explicit about its dependence of the Incredible Coincidence and the Small Chance Instinct on which it is based:

11. The court judges that it follows from the probability calculation of the expert E. ... that it should be deemed extremely improbable that the suspect was accidentally present at all the incidents in the JKZ and the RKZ which have been charged at her account. This calculation indicates therefore that it is highly likely that there is a

⁴⁰ Requisitoir, p. 72. For the causal link between the presence of the suspect and the incidents the prosecution claims to have independent proof (Requisitoir, p. 73).

⁴¹ Testimony before the court of appeal p. 73.

⁴² Prosecution before the court of first instance, Requisitoir, p. 73.

⁴³ The reader may be confused about these six deaths and the five deaths of the prosecution. The answer is: there were five deaths from November 1999 till 9 September 2001. The sixth death in 1999, 2000 and 2001 was in March 1999, when the ward was still called Internal-1.

⁴⁴ By the way, the prosecution charged Lucia with only four of those. Perhaps it could have mentioned that too, for clarity's sake.

connection between the doings of the suspect and the occurrence of the incident meant.

(j) the prosecution before the court of appeal

In their indictment the prosecutors appeal to the Big Number (1 in 324.000.000) to convince the court of Lucia's guilt:

In the mean time it is generally known that the suspicion which has led to the prosecution of Mrs. De B., depends on the fact that in the last hospital she worked in, she was uncommonly often involved with suspect deaths and reanimations. One has tried to give a better quantification to this "uncommonly often". For this purpose Mr. E. has been approached. ... From his calculations his general conclusion emerged that the fact that so many incidents occurred during the suspect's shift, is not compatible with mere accident.⁴⁵

After the exclusion of some alternative explanations the Incredible Coincidence settles the guilt question for the prosecution.

The two prosecutors also create a coincidence of their own. They do not feel embarrassed when they refer to their own shoddy, amateurish statistics of sorts:

At last [we] note that with respect to the intoxications the probability that there is an external factor in all these cases, or that we have to search for a cause outside the suspect, is so small that we may reasonably speak of an impossibility. Granted, this is a shoddy, amateurish statistics of sorts.⁴⁶

Apparently, shoddy amateur statistics of sorts, even if the unreliable status is recognized publicly, is acceptable in the context of a murder charge with a possible sentence of lifelong imprisonment. The prosecutors do hope that the "professional and expert statistical reports" make good their claims.⁴⁷ So they recommend the court of appeal to take the statistical considerations 'as their starting point and guiding principle in their deliberation'.⁴⁸ The Incredible Coincidence could not have been clearer in their mind. And it apparently exculpates the use of their own shoddy and incompetent statistics, as long as this is self-acknowledged.

(k) the court of appeal

The court of appeal claims that statistical considerations have not played a role in their considerations, but they in fact have as the consideration in the court's *arrest* demonstrates. For the court the coincidence of seven incidents during the shifts of one nurse is deemed too improbable to be believed as merely accidental.⁴⁹

11.13 There is no plausible explanation for the fact that the suspect was involved in so many cases of death and life-threatening incidents.

⁴⁵ Requisitoir of C.J.M.G. Strack and G.C. Haverkate, prosecution before the court of appeal, pp. 2-3.

⁴⁶ Prosecution before the court of appeal, Repliekaantekeningen, June 2004 p. 24.

⁴⁷ Prosecution before the court of appeal, Repliekaantekeningen, June 2004 p. 24.

⁴⁸ Prosecution before the court of appeal, Repliekaantekeningen, June 2004 p. 24.

⁴⁹ Compare consideration 11.24 E in the construction of the proof. The court reminds the reader "that seven cases of death and life-threatening incidents that occurred in the JKZ, happened in a relatively brief period (18 September 2000 - 4 September 2001)".

This is not just a remark. It is part of their construction of the proof

(7) Conclusion of Part I

Many people who played a crucial part in the trial of Lucia, were deeply impressed by the (purportedly) extremely small chance that all those incidents happened during her shifts purely by accident. With the Small Chance Instinct — and the Smoke & Fire Instinct — operating in the background the belief became commonly accepted that the concurrence could not be just a coincidence.

When the Small Chance Instinct operates, the precise number does not matter. The number of Elffers — 1 in 324.000.000 — is often mentioned, but, so the intuition goes, 1 in 1.000.000 or 1 in 100.000 are also sufficient. All these chances are too small to be taken seriously. The idea that the coincidence of incidents and Lucia's shifts is just an accident, has to be rejected.

PART II

(8) The Incredible Coincidence Argument in its statistical version

We have seen that both the hospital and the judicial system shared the intuition that the coincidence of seven incidents during Lucia's shifts in the JKZ could not just be a matter of chance. Lucia must have done it! But how do you determine the reliability of this intuition? The police and the prosecution asked Elffers, professor in psychology of law, erstwhile statistician, to work out the mathematics of the case.

Elffers started from the hypothesis that the coincidence of seven incidents and Lucia's shifts was a matter of chance. Call this the Chance Hypothesis. He then calculated how probable it is that such a coincidence of incidents [7 out of 7] within Lucia's shifts [142 out of 1029] would have occurred by chance. If that probability does not reach a certain limit (the significance level), then the Chance Hypothesis is rejected. In other words, you deem the probability that the coincidence is just a matter of chance, too small and you accept that what happened is *not* a matter of chance.

Elffers compares the situation with the drawing of balls (shifts) from a box. In the box there are 1029 balls, 7 back balls (shifts with an incident) and 1022 white balls (1022 incident-free shifts). The probability that Lucia will draw a specific ball (shift) is the same for all balls (shifts), namely 1 in 1029. This is Elffers' translation of the hypothesis that the seven incidents all occurred within Lucia's shift by chance. Elffers used Fisher's Exact Test to calculate this probability precisely.

JKZ MCU-1: 11 October 2000 – 9 September 2001	no incident	incident	total
shifts of Lucia	135	7	142
shifts without Lucia	887	0	887
tota number of shifts	1022	7	1029

Fisher's Exact Test yields the result: in those circumstances the probability⁵⁰ that by pure chance someone would be present at seven out of seven

⁵⁰ Statisticians will tell you that it is a p-value. That is relevant because these p-values may not be multiplied.

incidents is 0,000008370726, or 1 in 1.194.640.⁵¹ This probability is way below the significance level that Elffers thought absolutely safe, viz. 1 in 10.000. Actually he could not imagine that someone would disagree with him.

In this case the coincidence is so improbable that I cannot imagine that someone could think that the coincidence is compatible with chance.⁵²

So the conclusion is: the Chance Hypothesis should be rejected. The coincidence is not a matter of a chance. Elffers taught the court that this conclusion does not imply that Lucia is guilty. There are some alternative explanations which have to be ruled out first. For example, it may be that Lucia had many night shifts and that during the night more people die, or someone else may have been with her at all those seven shifts. The court had no trouble in getting rid of these alternatives quickly.⁵³ It then concluded that the coincidence could not plausibly be explained in any other way than by assuming that Lucia is guilty of four murders and three attempted murders.

(9) Problems with the data

This statistical argument has many problems. Apart from serious problems with the data used, there are problems with the way Elffers used the method, and problems with the method itself. I start with the problems with the data.

(i) the selection and collection of the data was biased

Three of the four hospitals involved have been asked to search for incidents during Lucia's shifts. However, there was no determined search for incidents *outside* her shifts. And it is precisely these incidents outside her shifts which crucially determine how remarkable it is that Lucia "drew" seven shifts-with-incidents. For example, with 40 shifts-with-incidents it is much less improbable that seven incidents occurred during Lucia's shifts.

The prosecution did not only neglect to search for incidents outside her shifts, it actually removed incidents which fell outside her shifts, from the lists of incidents. This may be acceptable from a prosecutions' viewpoint, it is not from a statistical viewpoint as it only leaves the incriminating evidence.

We have examined the incomplete available records and found three new incidents outside Lucia's shifts, and five more in which it was not clear whether the incidents were within or outside Lucia's shift. We also discovered that only five of the seven incidents of the JKZ fell within Lucia's shifts. If one does statistics based on shifts, one should do it properly.

All this brings out that the data as used are quite unreliable. The least we can do is to recalculate Fisher's Exact test on the basis of the new (less biased) data.

(ii) data were restricted to three wards and 1 ¼ years

There is another bias in the data which we cannot even begin to straighten out: the data used for the calculation concern only three of the five wards

⁵¹ Fisher's Exact test calculates the probability that by chance the person will be present at *at least* seven incidents. With seven incidents that coincides with just seven. In the text I will not always mention the *at least*.

⁵² Elffers, Testimony 29 January 2007, p. 11.

⁵³ Arrest, consideration 10.13 and following considerations.

where Lucia worked, and only 1 year and 3,5 months out of the 11,75 years that she worked in those hospitals. The data of the other hospitals and other periods were not available, or so it was said. But in view of a possible conviction for life one should have been worried about this limitation. There is now a risk that the alleged extremely small improbability that the coincidence is a matter of chance, is the result of a highly specific collection of the data, namely of those data that stand out. Those periods are selected in which there are proportionally more incidents. Small wonder that we find more incidents than on average.

(iii) the data were used twice, both in the formulation of the hypothesis and in the testing of the hypothesis

The situation is just as bad as it looks. The seven incidents in the JKZ triggered the investigation. They led to the formulation of the hypothesis that Lucia might well have murdered patients in the JKZ. But subsequently these data were used a second time to test that very hypothesis. This is very tricky statistics. To encounter seven incidents during one's shifts is very improbable, but it may happen by chance. So we have to examine whether the coincidence is just a chance event *or* an indication of questionable behaviour. We may not use the old coincidence again. We want to know about the coincidence whether it was due to bad luck or due to bad behaviour. So we need to find *other* evidence.

Let me use an example to illustrate this. Suppose we find that mr. A. has won the lottery. Now a priori there is only a small probability of winning a lottery. Yet he won. That happens. But it is possible that he cheated. To demonstrate that he cheated we cannot again point to his winning. We need other evidence, for example that he has won many other lotteries recently or that he had special contacts with the lawyer who did the drawing. You cannot use his winning itself as an argument that his winning was achieved with illegal means. In that case we should have to arrest all winners of lotteries.

Of course, this is ridiculous. But the point is that a similar mistake is made in Lucia's case when we first pick her out as a possible murderer because of the seven incidents during her shifts and then convict her because of these very seven incidents.⁵⁴

(10) Problems within the method used

Before we recalculate the relevant probability using Elffers' method we need to make one adjustment. Elffers calculated the probabilities related to three different wards, and then multiplied those probabilities. In his 'Lying Statistics Damn Nurse Lucia de B' Richard Gill, professor of statistics at Leiden University, calls this a 'technical blunder', which 'biases the result against Lucia'.⁵⁵

⁵⁴ To compensate for the double use Elffers made a *post hoc* correction. He multiplied the outcome of Fisher's Exact Test by 27, because in the period used by Elffers 27 nurses worked at Lucia's ward. We will not argue here why this is a wrong theoretical move. It should suffice that the number Elffers reached after this post hoc correction ($27 \times 0,000008370726 = 0,000226009602$, or 1 in 44.246) is still way off compared to the 1 in 44 which emerges from a recalculation with the proper data.

⁵⁵ See his personal website (<http://www.math.leidenuniv.nl/~gill/lucia.html>), where he gives references to support this claim: 'A technical blunder further biases the conclusion: combination of p-values by multiplication instead of by the easy "last resort" Fisher's method or the more appropriate Cochran-Mantel-Haenszel test; see the standard textbook Agresti (2002), website Categorical Data Analysis, and the accompanying manual Thompson (2006). ... Expert for the prosecution Henk Elffers, Professor at the University of Antwerp, senior-researcher at the Netherlands Institute for the Study of Crime and Law Enforcement, (a friend, former colleague and co-author of mine from our Mathematical Centre days thirty years ago, before he moved to geography, economics, psychology and law) apparently does not know the

The inappropriateness of multiplying can be shown in a simple way. Suppose that the probabilities related to different hospitals should be multiplied (quod non). Now look what happens when a nurse changes hospitals. For her first hospital we calculate the probability that by chance she would encounter the number of incidents during her shifts. Suppose this probability is around 1/3. Nothing remarkable happened.⁵⁶ In the next hospital this history repeats itself, and so and so. So far still nothing remarkable happened. But then the prosecution starts to multiply all the probabilities of 1/3. It will take some time but in the end the resulting probability will get below the significance level, and some prosecutor can improve his/her career prospects with another brilliant catch: here is another nurse during whose shifts something so improbable happened that it cannot be a matter of chance. Bingo a new serial killer, although her colleague who stayed in the first hospital, may have encountered exactly as many incidents during her shifts!⁵⁷

This is of course a travesty of reasoning, but it demonstrates that multiplying the probabilities of the different hospitals cannot be an acceptable move within the legal system: the mathematics alone makes murderers out of people. Nothing untoward needs to happen, and yet the nurse who changes hospitals once too often, will face a life sentence.

This then is a serious problem with the method as used by Elffers: the multiplication of the probabilities of the different hospitals and wards is both statistically and legally wrong.⁵⁸

(11) Recalculation with Fisher's Exact test

We have seen that we have to adjust the data and the method. With respect to the Juliana Children's Hospital we have to add two recently discovered incidents outside Lucia's shifts⁵⁹, we have to remove two incidents from her shifts (incidents outside her shifts had been treated as being incidents within her shifts), and we should not multiply the probabilities of the different hospitals.

Let us redo the calculation for the Juliana Children's Hospital alone.

Juliana Children's Hospital, Medicum Care Unit-1 11 October 2000 – 9 September 2001	no incident	incident	total
shifts of Lucia	137	5	142
shifts without Lucia	883	4	887
total number of shifts	1020	9	1029

meaning of p-value. He multiplies three independent p-values (from three wards where Lucia has worked) and appears to present the product as a p-value rather than using one of the well-known ways to compensate for the *number* of statistical tests being combined. This error biases the result against Lucia'.

⁵⁶ The nurse has 8 hour-shifts.

⁵⁷ Gill subscribes to this conclusion in his Elffers' method and Elffers' mistake' (on his website): 'Elffers' method of combining wards by multiplying p-values is blatantly incorrect, since data from a large enough number of wards would make *any* nurse eventually guilty'.

⁵⁸ The second problem is that Elffers assumed that the balls were all equally likely to be drawn. Richard Gill objects to this assumption, the balls were sticky: we should not assume that all nurses have the same chance to get shifts-with-incidents. He may well be right, but in a recalculation we do not need this extra complication. As we will see in a moment, the simplified version of equally sticky balls does already yield a clear enough verdict: there is no statistical reason whatsoever to suspect Lucia because of the coincidence of incidents and her shifts.

⁵⁹ The third recently discovered incident outside Lucia's shifts was in the Red Cross Hospital, not in the Juliana Children's Hospital..

Fisher's Exact test now yields the result: in those circumstances the probability that by accident someone would be present at (at least) five out of nine incidents is 0,003704, or 1 in 270. This probability is way above Elffers' significance level of 1 in 10.000.

When we redo the calculation for all three wards, Fishers' Exact Test gives a chance 1 in 44, and with a more sophisticated calculation Richard Gill, professor of mathematical statistics, Leiden, gets an out come of 1 in 9.

That is, in all calculations, the Chance Hypothesis should not be rejected, it should be accepted.

(12) Problems with the method: the wrong question is being asked

Let us note that on 4 September 2001, the day when the hospital and the police started their investigation, there was no direct evidence of any murder. There was just a coincidence between incidents and Lucia's shifts,⁶⁰ and the powerful intuition that this was very fishy. So a statistician was called in, and he was asked: could this coincidence just be a matter of chance?

The statistician reformulated this as:

"What is the probability that such a coincidence of incidents and shifts would occur by chance (given the number of Lucia's shifts and the total number of shifts)?"

The question thus actually is:

"Assuming the innocence of the nurse, what is the probability that all these incidents occurred during the services of Lucia by chance?"

But *this is the wrong question*. We want to know

"Given the coincidence, what is the chance that the nurse Lucia is innocent or not innocent?"

The complication is that Elffers answered the Wrong Question in such a way that it seemed to yield an answer to this Right Question (about Lucia's guilt or innocence). Elffers found that the probability that the coincidence was a matter of chance, was below the significance level he set in advance. So he told the court: "This cannot be a matter of chance". He added that this does not entail that Lucia is the murderer. There were some alternative explanations of why the coincidence is not a matter of chance. However, the court had no trouble in getting rid of these alternative explanations. So Elffers' question led to the conclusion that Lucia must be guilty. It seems the right question after all, for it gave the right conclusion, - or so it seemed.

However, does Elffers' conclusion "This cannot be a matter of chance!" follow from his extremely low probability that the coincidence was a chance affair?

Prima facie the implication seems to be self-evident. There is an extremely small probability (in Elffers' calculation) that the coincidence occurred by chance. And it is an everyday rule of detachment that if it is very unlikely that p, it is reasonable (in normal circumstances) to believe that not-p. Elffers told the court that there was only a very small probability that the coincidence occurred by chance, so it is reasonable to believe that the

⁶⁰ Much later the prosecution allegedly found evidence for two intoxications.

coincidence was not a matter of chance. And after having rejected the alternative explanations, the court may conclude: Lucia is a serial killer.

Yet, in spite of appearances the conclusion does not follow. It neglects the base rates and the prior probability of the opposite hypothesis, namely that a Dutch nurse is a serial killer. Rather than using a theoretical argument to support this claim we give two examples illustrating what went wrong.

The first example is brother Piet. In his parental home, during card games, he once got terrific hands in a row. The chance that he would get such hands was incredibly small, say 1 in 44.000. (This is the chance Elffers calculates for the JKZ alone). Given Elffers' significance level of 1 in 10.000 we should reject the Chance Hypothesis (Brother Piet got these incredible hands by pure chance), and after having excluded alternative explanations (His mother gave him these cards intentionally), we conclude (if we follow Elffers' method) that brother Piet cheated.

But that is not what we concluded. We knew that he has not the ability to cheat. The probability that he cheated is therefore so small that we stick to the conclusion: those incredible hands happened by accident. The Chance Hypothesis has a very small probability (1 in 44.000), but the probability of brother Piet cheating us, is even smaller, - or so we think.

Similarly in Lucia's case, even if the (prior) probability that all those incidents during Lucia's shifts occurred by chance, is extremely small (quod non), that in itself does not imply that we should conclude that therefore the coincidence is not a matter of chance. That depends on the prior probability of the alternative, namely that a Dutch nurse is a serial killer. So, whatever this prior probability may be, our first example brings out the following:

Elffers' conclusion "The coincidence cannot be a matter of chance!" does not follow *automatically* from the extremely small probability that Lucia's coincidence happened by chance.

What is this prior probability of a Dutch nurse being a serial killer? We have no idea, but we may reasonably assume that there have been very few serial killers among Dutch nurses throughout the years. As far as I know Lucia is the only Dutch nurse who has been convicted as a serial killer, and we may assume further, I think, that serial killers would not go unnoticed too often. So let us take the ridiculously high estimate that every year there is a serial killer-nurse in the Dutch hospitals. There are 60.000 nurses in the Netherlands, so the prior probability that some nurse would be a serial killer is 1 in 60.000. (A more reasonable guess would be 1 in 600.000 or 1 in 1.200.000).

Given these prior probabilities it is not so clear anymore why we should reject the Chance Hypothesis that Lucia's coincidence happened by chance. The prior probability that the coincidence would happen by chance (as calculated by Fisher's Exact test) was extremely small, but the prior probability of a Dutch serial killer-nurse is even smaller.

This may seem very abstract but we can translate all this into a *decision strategy for courts*. And this is our second example. Assume that the courts follow Elffers' strategy: If the probability of the coincidence (of incidents and shifts) happening by chance, falls below the significance level of 1 in 10.000 and there is no other explanation for the coincidence, then convict the nurse to a life imprisonment. We have seen that Elffers calculated an extremely small probability. We have also seen that 1 in 44 or in 1 in 9 are closer approximations of that probability in Lucia's case. To get a simple

example, let us choose a probability of 1 in 44.000, the probability Elffers calculated for the JKZ alone.⁶¹

Let us now ask what this 1 in 44.000 implies about the number of nurses the court may expect to see. Remember there are 60.000 Dutch nurses. So each year — on average — there will be at least one nurse with seven incidents⁶² during her shifts. Assuming that the hospitals are as vigilant as the JKZ and catch any nurse with a high coincidence of incidents and shifts, at least one innocent nurse will be convicted each year. With one serial killer-nurse each year, the court would also — on average — see one serial killer before the bench. Also this serial killer will be caught with a coincidence of incidents and shifts, — and not by chance this time.⁶³ This nurse will be convicted as well.

Note now that in this situation in which we loaded the dice against Lucia (1 in 44.000 rather than 1 in 44; 1 in 60.000 rather than 1 in 600.000), the court will on average convict two nurses, one guilty and one innocent, each year. There is then no reason to rejoice when the court convicts a nurse. Its verdict will be mistaken in more than 50 % of the cases.

It can be seen easily that the situation becomes even more dramatic when the ratio is 1 in 44 or 1 in 9. With a ratio of 1 in 44 the court may expect to see $60.000 : 44 = 1366$ innocent nurses with Lucia's coincidence, against 1 serial killer-nurse.⁶⁴

So when the court's decision-strategy is "Follow the advice of the statistical expert", we find that it will very often convict innocent nurses, simply because it neglects the prior probability of a Dutch nurse being a serial killer.⁶⁵

Rather than following Elffers the court should have asked the *right question*: given that Lucia encountered x of the total of y incidents and had z shifts (of all r shifts), how likely is it that she has committed murders?

To answer this question the court should consider both the prior probability of such a coincidence occurring by chance (1 in 44.000 or 1 in 44), *and* the prior probability of a Dutch nurse being a serial killer, *and* of course *other*, additional evidence.

Asking the wrong question is called the *prosecutor's fallacy*. Apparently, worldwide prosecutors are being advised by Elffers-look-alikes. In the discussion it sometimes seems as if one's own statistical presuppositions determine whether Elffers' argumentation is a fallacy or not. Given Elffers'

⁶¹ The 1 in 44.000 is the probability related to the Juliana Children's Hospital after Elffer's post-hoc correction

⁶² The number of incidents in the JKZ Lucia is convicted of.

⁶³ I assume a stupid serial killer who kills only during his/her shifts.

⁶⁴ But how about the 1 in 342.000.000, the number Elffers came up with? Well in that case we should take a more realistic guess of the numbers of nurse-serial killers in Holland. Let us say: one in every twenty years, that is 1 in 1.200.000 nurses. Even with this (*totally wrong*) number, and *discounting further statistical data* we will mention in a moment, we may expect one innocent nurse among every 300 convicted nurses. That is, the probability of making a wrong conviction is 1 in 300 rather than 1 in 342.000.000.

⁶⁵ The court did not know more about Lucia than that she is a Dutch nurse with so many shifts and so many incidents. The court did make the claim that Lucia was a liar because she denied that she had killed people. Her alleged mendacious character was further evidence against Lucia, or so the court claimed. The court also concluded that there was independent evidence that Lucia fatally intoxicated two children. If the court would have had good reason for doing so (*quod non*), the statistical argumentation would have been superfluous.

statistical theory it is not a fallacy, given the statistical theory of his opponents, it is. However, we need not take a position in the general statistical debate to see that Elffers' approach does commit a fallacy indeed. Both when the case is perceived from everyday experience (my brother Piet) and when it is viewed from the perspective of a general decision-strategy, we see that that Elffers' approach leads to the wrong outcome: we do not think that our brother cheated, nor do we appreciate sending many innocents nurses to jail. So Elffers' approach is wrong, his question is wrong, his conclusion is based on a fallacy and it has led to a disastrous result in the case of Lucia de B.

Summarizing, the court should not ask: "Assuming Lucia's innocence, what is the probability that she meets with such a coincidence by chance?", and then, when that probability turns out to be smaller than the significance level, conclude: "She is the cause of the coincidence" (once the alternatives have been quickly excluded). The question should be: "Given the coincidence, is there reason to convict Lucia?" And the court has to take into account both the prior probability of the coincidence occurring by chance, and the prior probability of a Dutch nurse being a serial killer, and *other* evidence.

(13) Other Relevant Data: Total Collapse of the Case

But Lucia's position is much stronger. There was no incriminating evidence against her,⁶⁶ but actually there was exculpatory evidence, as we have already seen in Part I: during the three years (1996 – 1998) before Lucia worked at the JKZ seven children died in the ward MCU-1, during the three years that she was employed there (1999 - 2001) six (6!) children died at that ward. This should cast serious doubt on the idea that a serial killer was active during the years that Lucia worked at MCU-1.⁶⁷

	number of deaths in ward MCU-1 (which had the name Internal-1 before 1999)
1996-1998 (Lucia not in MCU-1)	7
1999-2001 (Lucia in MCU-1)	6

Nobody noticed these numbers. How is this possible? The only reason we can think of, is that the prosecution before the District Court gave a devious turn to these numbers. It used a semantic trick. The ward MCU-1 changed its name from IN-1 to MCU-1 on the 9th of November 1999. So we could misleadingly present the statistics of deaths as follows.

	number of deaths in ward Internal-1	number of deaths in ward MCU-1
1996-1998 (Lucia not in MCU-1)	7	0

⁶⁶ Ton Derksen argued for this claim in his *Lucia de B. : a reconstruction of a miscarriage of justice* (in Dutch, 2006, Diemen: Veen Magazines)

⁶⁷ Lucia did not work all 36 months of 1999 – 2001 in the hospital. In my book *Lucia de B.* I make a precise calculation in terms of months (rather than in years) that she was employed by the hospital. The outcome is similar: during the months that Lucia worked in the Juliana Children's Hospital, there were (proportionally) *fewer* deaths than in the months that she did not work there. Again, this should cast serious doubt on the idea that a serial killer was active during the years that Lucia worked at MCU-1.

1999-2001 (Lucia in MCU-1)	1 ⁶⁸	5
----------------------------------	-----------------	---

The prosecution could now state (almost) truthfully: during the last six years only five people died in ward MCU-1 and they all died during Lucia's shifts. But this is a chutzpah. This turns highly exculpatory evidence into incriminating evidence. For in the ward MCU-1, which changed its name in November 1999, there were 13 deaths, 6 of which occurred in the years that Lucia worked at the JKZ.⁶⁹

So not only is the popular figure of 1 in 342.000.000 based on a fundamental mistake and the wrong data, there is other statistical evidence, which was deviously misinterpreted, and which should have cleared Lucia.

(14) Conclusion

In Part I we argued that it was the Incredible Coincidence which controlled the minds of all who played a crucial role in the process of Lucia de B. It is the Small Chance Instinct, an evolutionary-evolved reasoning instinct, that is the real culprit (we assume), because we cannot imagine that so many people were either just wicked or plain stupid.

In part II we gave a quick overview of the reasoning and other mistakes involved. Apart from a biased data-set we noticed a mistaken method and mistakes even in the way the mistaken method was used. The infamous coincidence was neither incredible nor incriminating. We also found that important exculpatory statistical evidence was not taken into consideration.

Once we remove the bias and the mistakes and add the neglected statistical evidence there is no statistical reason to suspect Lucia of any crime.

⁶⁸ The patient died in March 1999 when the ward was still called Internal-1. At that time Lucia did not yet work in the Juliana Children's Hospital. The ward changed its name in November 1999.

⁶⁹ To be exact, there were five deaths during the months that she worked there.