Elementary Statistics on Trial
(the case of Lucia de Berk)

Richard D. Gill[*], Piet Groeneboom [†] and Peter de Jong [‡]

8 August, 2010

Abstract

In the conviction of Lucia de Berk an important role was played by a simple hypergeometric model, used by the expert consulted by the court, which produced very small probabilities of occurrences of certain numbers of incidents. We want to draw attention to the fact that, if we take into account the variation among nurses in incidents they experience during their shifts, these probabilities can become considerably larger. This points to the danger of using an oversimplified discrete probability model in these circumstances.

1   Introduction

We propose a model to explain the number of relatively rare events occurring to different subjects, where the incidence rate per subject varies. The latter assumption models the heterogeneity of the subjects, one of the issues we want to emphasize.

The model was inspired by the case of Lucia de Berk. The trial of the Dutch nurse suspected of several murders and attempts of murder was a very high profile case in the Netherlands. The initial suspicion rested mainly on statistical considerations, which produced (based partly on incorrect calculations) extremely small probabilities "under the null hypothesis". When the outcomes proved controversial the court dropped the statistical calculations from from the verdict. But still the verdict rested on intuitive notions as "very improbable". Statistics remained at center stage.

---

[*]Mathematical Institute, Leiden University; `http://www.math.leidenuniv.nl/∼gill`
[†]DIAM, Delft University; `http://ssor.twi.tudelft.nl/∼pietg`
[‡]WSO, Amsterdam; `mail:p.dejong@wiston.info`

The outcomes of applying the model to this case are in stark contrast with those of the first calculations which led to the initial suspicions and were instrumental in determining the atmosphere surrounding the trial and subsequent hysteria. Since then the trial of Lucia de Berk has resulted in acquittal and the trial is seen as one of the major miscarriages of justice in the Dutch Judicial system. For more background on the Lucia case see de Noo [2010] and de Berk [2010]. Accordingly the emphasis of this paper has shifted towards the lessons to be learned from this involvement of statistics with judicial matters.

The main result is that under the assumption of heterogeneity, the probability of experiencing a number of incidents (7) that led to Lucia's conviction is about *one in seven*. Moreover the probability of experiencing a far greater number of incidents, equalling the number of suspicious incidents (solely because occurring in the shift of Lucia de Berk) (13) is, although smaller, certainly not negligibly small: almost *one in twenty-six*.

Section 2 introduces the model and contains the actual calculations. Section 3 discusses the assumption of heterogeneity. This section not only deals with the data and the data collection, but also touches on more general issues. An important aspect is the analysis of the assumptions underlying the application of the hypergeometric distribution. We find that the initial conclusions are very sensitive to small errors in the observed data.

In an appendix 5 we present the data in some detail, since there has been some confusion over which incidents had been counted as 'suspicious' incidents. We try to settle this aspect.

2   The model and some calculations

2.1   The model

We model the incidents that a nurse experiences as a homogeneous Poisson process on the positive halfline, with a nurse-dependent intensity λ. As is well-known, a Poisson process is used to model incoming phone calls during non-busy hours, fires in a big city, etc. Since we believe the incidents to be rare, a Poisson process is an obvious choice for modeling the incidents that a nurse experiences.

Our approach models two separate phenomena. Firstly, the intensity of nurses seeing or reporting incidents (see e.g. subsection 3.2) is modeled parametrically, and we take as the distribution of the intensity λ over nurses the Gamma$(\rho, \rho/\mu)$ distribution. Secondly, the number of incidents happening to a nurse on duty depends on the intensity λ and the time interval she is

working, and follows the Poisson distribution. Using this model, our sample consists of realizations of the random variable

$$(L, T, N),$$

where L has a Gamma$(\rho, \rho/\mu)$ distribution, and N, conditionally on $L = \lambda$ and $T = t$, has a Poisson distribution with parameter $\mu = \lambda t$. The random variable T represents the time interval in which incidents take place (for a particular nurse). So, if there are n nurses, we deal with a sample

$$(L_1, T_1, N_1), \ldots, (L_n, T_n, N_n),$$

of independent random variables, all having the same distribution (L, T, N). The random variable N_i represents the number of incidents nurse i experiences in the time interval T_i. As a somewhat arbitrary choice, we will take $\rho = 1$, thus the intensity λ is exponentially distributed. This implies, among other things, that it can easily happen that one nurse has twice the incident rate of another nurse.[1]

The statistical problem boils down to the estimation of the Poisson parameter μ.

We use Derksen and de Noo's[2] revised data set Derksen [2007][3], taking account of incidents among the other nurses (which formerly were taken to be, by definition, *not suspicious*), and removing incidents and deaths for which Lucia was deemed innocent (not charged with murder or attempted murder, presumably because these events were medically speaking "expected to happen, when they actually did"). The numbers can be confusing: The verdict in 2004 is based on 7 murders and 3 attempts of murder. We 'only' use 7 incidents, disregarding two incidents that fall outside any of the data sets used for statistical calculations, (because these cases fell outside the timeframe or the hospitals for which data were available) and one incident that occurred outside Lucia's shift. The latter may seem somewhat strange but we believe the court ordered the same to the statistician for the prosecution (see 5).

We will also give results for the data set obtained when we do not treat Lucia so generously: a number of incidents removed by order of the courts as well as those removed by Derksen are put back. We will see that though this has a big effect on the number of incidents in her shifts (raising from 7 to 13), its effect on our final conclusion is small.

---

[1]The probability of this event is 2/3; in fact the probability of a incidence rate of a factor k times that of another nurse is $2/(k+1)$.

[2]Ton Derksen and Metta de Noo are the only people who had access to the complete files of the process. Their involvement has led to the final acquittal of Lucia de Berk.

[3]The underlying data, until now, have not been released for scientific research.

2.2 The numbers and the calculations

Combining the Juliana Children's Hospital and the two wards of the Red Cross Hospital, Lucia had 203 shifts, 7 incidents. It is not clear whether this combination works out pro or contra Lucia. This depends on whether she did proportionately more or less shifts at the different wards, and whether the overall mean incident rate is larger or smaller at each ward.

The total number of incidents is not clear: It varies between 26 and 30, depending on the total number of incidents at RKZ-42 (varying between 10 and 14). This constitutes a major flaw in the investigation: the data collection is irreproducible and lacks rigorous methods and definitions. It crucially depended on the memory of people who knew what was sought after. We will return to this in section 3. We follow the court by letting 4 incidents disappear from the shifts of Lucia; the incidents have *not* been moved to other shifts, they have been declassified as incident. We'll take the overall probability of an incident per shift to be the ratio of total incidents to total shifts, $\mu = 26/1734$. If we take a shift to be our unit time interval, then this would be a moment estimate of the mean intensity of incidents EL. This means, that, conditionally on $T = 203$, the number of incidents for Lucia follows a mixture of Poisson random variables with parameter $203L$, where the intensity L has a $\mathrm{Gamma}(1, 1/\mu)$ distribution, which is in fact the exponential distribution on $[0, \infty)$ with first moment μ. Thus on average, an innocent Lucia would experience $203 \cdot \mu = 203 \cdot 26/1734 \approx 3.04383$ incidents.

The probability of having 7 or more incidents is given by:

$$\frac{1}{\mu} \int_0^\infty \mathbb{P}\left\{N \geq 7 | L = \ell, T = 203\right\} e^{-\ell/\mu}\, d\ell$$

It is well-known from elementary calculus that, for a random variable N, which is distributed according to a $\mathrm{Poisson}(\lambda)$ distribution, we have:

$$\mathbb{P}\{N \geq n\} = 1 - \sum_{k=0}^{n-1} \mathbb{P}\{N = k\} = \frac{1}{(n-1)!} \int_0^\lambda e^{-x} x^{n-1}\, dx.$$

So we find:

$$\frac{1}{\mu}\int_0^\infty \mathbb{P}\left\{N \ge n | L = \ell, T = t\right\} e^{-\ell/\mu}\, d\ell$$

$$= \frac{1}{(n-1)!\mu}\int_0^\infty \left\{\int_0^{t\ell} e^{-y} y^{n-1}\, dy\right\} e^{-\ell/\mu}\, d\ell$$

$$= \frac{1}{(n-1)!}\int_0^\infty e^{-\{1+1/(t\mu)\}y} y^{n-1}\, dy$$

$$= \frac{1}{(1+1/(t\mu))^n} = \left(\frac{t\mu}{1+t\mu}\right)^n.$$

This is the geometric distribution with parameter $(1 + t\mu)^{-1}$. With $n = 7$ and $t\mu = 3.04383$ this yields 0.13690, just a bit smaller than *one in seven*.

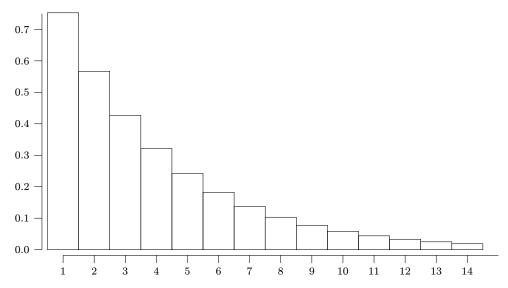A picture of the probalilities $\mathbb{P}\{N \ge k | T = 203\}$, $k = 1, 2, \ldots$ is shown in Figure 1.

Figure 1: Probabilities (in the model) that the number of incidents in 203 shifts for one nurse is at least 1,2,3,..., if $\mu = 26/1734$. The probabilities are given by the heights of the columns above $1, 2, 3, \ldots$, respectively.

A more unfavourable conclusion can be obtained for Lucia, if we do not follow first the court and then Derksen in removing incidents from the statistics. This results in two incidents occurring within a shift of Lucia, for which she was not found guilty. Further we reintroduce four incidents that had been omitted by Derksen and the court, adding these incidents to the total

number of incidents. This amounts to 13 incidents in Lucia's shifts, out of a grand total of 30 incidents in 1734 shifts. Taking again $\rho = 1$ we find the probability 0.03850 or *one in twenty-six*.

Heterogeneity of any kind increases the variation in the number of incidents experienced by a randomly chosen nurse over a given period of time (given number of shifts). From the well-known relations

$$\mathrm{E}(X) \;=\; \mathrm{E}(\mathrm{E}(X|Y)),$$

$$\mathrm{var}(X) \;=\; \mathrm{E}(\mathrm{var}(X|Y)) + \mathrm{var}(\mathrm{E}(X|Y)),$$

it follows that whereas for a Poisson distributed random variable variance and mean are equal, for a mixture of Poisson's (with different conditional means), the variance is larger than the mean. So if some nurses experience more or less incidents than other for "innocent" reasons as mentioned below, in all cases the end-result is *overdispersion* caused by *heterogeneity*. Applied to the current model which is geometric with parameter $(1 + t\mu)^{-1}$:

$$\mathrm{var}(N) \;=\; (1 + t\mu)^2 - (1 + t\mu) \;=\; t\mu + (t\mu)^2,$$

where the latter term neatly splits over the expected variance of the Poisson process plus the variance of the conditional parameter of the Poisson process which we assumed to be exponential[4].

2.3 Conclusion

A modest amount of variation makes the chance that an innnocent nurse experiences at least as many incidents as the number Lucia actually did experience, the somewhat unremarkable *one in seven*. Making some less favourable choices to her in the data cleaning process, only decreases this chance to *one in twenty-six*.

The fact that a modest amount of heterogeneity turns an almost impossible occurrence into something merely mildly unusual, is strong support for further empirical research whether and if so in what forms heterogeneity plays a role in healthcare. It can have major implications in different areas, such as medical research (representing an extra source of variation) and training of medical staff.

Another aspect is concerned with the contrast between the outcomes of the different model assumptions. When reporting outcomes of calculations, some kind of robustness analysis should be included. Robustness not only of data and data acquisition, but also robustness of the underlying model and calculations.

---

[4]If one wants a model with less overdispersion, the second parameter of the Gamma distribution could be used to accomplish this.

3 Extended discussion of heterogeneity

3.1 Preliminary remarks

In the previous sections we showed that a modest amount of heterogeneity leads to different orders of magnitude in the outcomes of crucial calculations. In this section we address some of the underlying mechanisms which may lead to the postulated heterogeneity. We describe two general mechanisms causing heterogeneity. The first one concerns properties of subjects directly related to the intensity of the rate of incidents. The other mechanism is more indirect and results from 'spurious correlations', in which properties not related to the underlying intensity, influence the measurement via unexpected dependencies and systematic variations in variables assumed to be independent and uniform.

Related to this is another aspect of the data: the degree into which a specific model or null-hypothesis is susceptible to small variations in the data. We will show this to be the case in the original calculations. Although our example is tuned to this very specific case, it refers to a much more general *caveat*. It should be established how stable certain models are under small perturbations of the data.

3.2 Are nurses interchangeable?

According to medical specialists we have spoken to, nurses are indeed completely interchangeable with respect to the occurrence of medical emergencies among their patients. However according to nursing staff we have consulted, this is not the case at all. Different nurses have different styles and different personalities, and this can and does have a medical impact on the state of their patients. Especially regarding care of the dying, it is folk knowledge that terminally ill persons tend to die preferentially on the shifts of those nurses with whom they feel more comfortable. As far as we know there has been no statistical research on this phenomenon.

There is another obvious way in which the intensity of incidents depends on characteristics that vary over the population. Any event that can turn out to be an 'incident' starts with the call of a doctor. And in all cases it is the nurse who decides to call a doctor. This decision is influenced by professional and personal attitude, past experience and personality traits like self-confidence. It seems obvious to us that these characteristics vary greatly in any population. Hence we assume that the intensity of experiencing incidents varies accordingly.

3.3 Inadequacy of the hypergeometric distribution as a model and spurious correlations

The model underlying the null-hypothesis (which led to the hypergeometric distribution) depends on two assumptions: Both the incidents and the nurses are assigned to shifts uniformly and independently of each other.

Above we have established two ways in which characteristics of individual subjects may lead to variation in the intensity of experiencing an incident. This variation is in contrast with one of the assumptions underlying the hypergeometric distribution: uniformity.

Next we discuss sources of correlation which correspond to indirect rather than direct causation: we speak then of spurious-correlation, correlation which can be explained by confounding factors, by common causes.

There are serious reasons to doubt the uniformity of incidents over shifts. There may occur periodical differences. The population of a hospital ward may vary over the seasons. The patients may differ in character and severity of illness due to seasonal influences. There are differences between day and night shifts and between weekend shifts and shifts on weekdays. Recently an extensive study of Dutch Intensive Care Units admissions shows a marked increase in deaths when the admission falls outside "office hours" Kuijsten et al. [2010]. Recall that there have to be nurses on duty throughout the night and throughout the weekends, while the medical specialists tend to have "normal working hours". Finally there is the periodical cycle of the circadian rhythm, influencing the condition of the patients and the attention of the medical staff Kuhn [2000]. There may be other, non-periodical variations that affect the uniformity of incidents. In the case of the Juliana Children's hospital there has been a rather sensitive matter of policy: whether very ill children, who are not going to live for very long, should die at home or in the hospital wards. We understand that this policy did change at least once at the JKZ in the period of interest. Presumably a change in policy concerning where the hospital wants children to die, will have impact on the rate of incidents. Further, incidents may be clustered, since one patient can give rise to several incidents.

On the other hand the way nurses are assigned to shifts is certainly not uniform and 'random'. Nurses take shifts in patterns, for example several night shift on a row, alternated by rows of evening or day shifts. Nurses are assigned to shifts according to skills, qualification and other characteristics. Maybe some nurses take relatively more weekend shifts than others, because of personal circumstances.

Although both the assignment of nurses to shifts and the assignment of incidents to shifts are not uniform processes, one could hope that there might

be some 'mixing' condition that makes the ultimate result indistinguishable from the postulated independence and uniformity. Certainly one may hope, but this magical mechanism should at least be made plausible.

Taken together, even if we consider both the shifts of a given nurse as a random process, and the incidents on a ward as a random process, and even if we consider the two processes as stochastically independent of one another, the assumption of constant intensities of either is a guess, not based on any evidence or argument. There may be patterns in the risk of incidents and there are certainly patterns in the shifts of nurses. These patterns may be correlated, through the process by which shifts are shared over the different nurses according to their different personal situations, their different wishes for particular kinds of shifts, their different qualifications, and the changing situation on the ward.

3.4   How stable is the null hypothesis under small changes in the data?

Consider the data (see section 5) of the ward at JKZ before revision. These numbers and their interpretation are at the root of what turned out to be one of the gravest miscarriages of justice of the Dutch Juridical system. Under the assumption of the hypergeometric distribution the probability of this configuration is very small, less than one in nine million. The configuration is in some respects extreme: eight out of eight incidents occur in the shift of one nurse. However the data are in another respect also conspicuous: no incidents occur in the 887 shifts where this nurse was not present. The data collection had been far from flawless, with no formal definition of incident, no or incomplete documentation, and rested at least in part on recollection of witnesses who were aware of which facts were looked for. Assuming the possibility of tiny flaws in the process of data acquisition, it is legitimate to investigate the effect of $1, 2, \ldots, 8$ incidents that could have been forgotten, or overlooked. This amounts to allowing a maximal error of less than one percent. The results are quite remarkable; in table 1 we give the inverses of the probabilities rounded to integers.

The large numbers vanish easily. Six or more incidents not remembered, not reported, or just defined away make the difference between astronomically small on the one hand and very unusual on the other. This shows that the model underlying the null hypothesis is sensitive to small errors in the data.

A judgement on data quality is not only the concern of a statistician. Judges are used to inconsistent and incomplete data (statements), psychologists are very well aware of the possible fallacies of memory. Both groups

Table 1: The effect of perturbations on the probabilities

Shifts with incidents outside Lucia's (postulated)	0	1	2	3	4
Inverse probability	9043864	1137586	257538	79497	29989
Shifts (continued)	5	6	7	8	
Inverse probability	13051	6329	3341	1889	

have their own professional standards of how to deal with these phenomena. A statistician, however, should point out what the effects of these phenomena can be on the outcome of his models.

If this model is used to corroborate evidence this sensitivity should be made explicit, just as adverse workings of a medicine are mentioned explicitly for the users.

4   Concluding remarks on the effects of heterogeneity

In the body of this paper we have shown the dramatic effect a modest amount of heterogeneity can have on tail probabilities, the probability that one nurse would experience a strikingly large amount of incidents. The data heterogeneity may stem from many different origins, be it variation in the characteristics of nurses, variation in the circumstances in time, or correlation between these. Room for subjective bias in data acquisition is, among others, a form of variation over the nurses. To put it bluntly, being a suspect in a murder trial is a major source of heterogeneity. One of the outcomes of this section is, rather surprisingly, that a simple discrete probability model like that of the hypergeometric distribution is based on assumptions that have to be verified, restricting it applicability in general.

Given the effects of heterogeneity on the outcomes, an analysis of possible sources of heterogeneity and their possible outcomes should accompany statistical reports, especially when there is so much at stake.

We conjecture that a more open minded and careful analysis of the data on the early stages of what was to become the case 'Lucia de B.' would have dampened the initial panic. The initial analyses which produced the astronomical small probabilities have been decisive in creating a very narrow mindset and tipped off the disastrous course of events.

5 Appendix: The data

In this section we present the data that were used in the various stages of the trial of Lucia de Berk. We rely mainly on the work of Derksen and de Noo Derksen [2007] and use also data from the unpublished reports of Elffers Elffers [May 8 2002] and Elffers [May 29 2002].

One of the key features of the data was the flawed data collection. Here different 'research' disciplines came into conflict: criminal investigation and 'scientific' data gathering are very different. Their objectives, methods and results are not compatible. Criminal investigation is started when there is (suspicion of) a crime, hence it is looking for or hunting down a suspect; it is not looking for a crime. If there is need for meaningful statistics another approach is needed guaranteeing clear definitions and uniformity of the data collection. In the case of Lucia de Berk this clash of cultures proved disastrous. Incidents without Lucia were discarded, incidents were declassified without clear reasons. The defense lawyer discovered extra shifts without incidents and incidents outside shifts of Lucia. Moreover the data collection rested for a large part on memory. Clearly, the context of a criminal investigation is mind setting: on the one hand the witnesses know what is looked for (and some of them may already be convinced of the guilt of the suspect), on the other hand fear of implicating one's self and friends can considerably distort memory.

We will summarize the data for each ward and give an overview of which incidents were counted in the various calculations. We start with an overview of all incidents followed by the tables for each ward, one for the original data one for the corrected data.

Table 2: All incidents

Name[a]:	V. 2004[b]	E9[c]	E8[c]	E7[c]	GGJ7[d]	GGJ13[d]	Remarks[e]
Eda(18/09/00)	A	-	-	-	-	-	Dropped
Ka I(10/10/00)							New, out
Jouad(10/10/00)	M	X	X	X	X	X	
Ka II(25/10/00)	M	X	X	X	X	X	
Kemal I(27/10/00)		X	X	X		X	
Kemal II(20/12/00)		X	X	X		X	
Sadia(17/01/01)[f]							Moved, out
Achmad I(25/01/01)	A	X	X	X	X	X	
Achmad II(23/02/01)	M	X	X				Moved, out
Kemal III(02/03/01)							New, out
Sarah(18/04/01)	-	X	-	-	-	-	Dropped
Achraf(01/09/01)	A	X	X	X	X	X	
Amber(04/09/01)	M	X	X	X	X	X	
Zonneveld(27/11/97) R	M				X	X	
Wang(12/11/97) R	M				X	X	
De Koning(09/05/97) L	M	-	-	-	-	-	Dropped
Unknown RKZ-42						4	Extra
Total:					26	30	

One incident at JKZ, although judged to be an attempted murder, has been excluded from all statistical calculations (18/09/00) since it fell outside the timeframe for reference. Another incident has been excluded in the early stages of the trial (18/04/01). A third incident has been moved on formal grounds from the statistical calculations, but was still included in the verdict of the court. The incident of (09/05/97) in table 2 is included in the verdict; by lack of data on the Leyenburg Hospital the case is left out of any computations.

---

[a]Name and date of incidents, R: RKZ, L: Leyenburg Hospital

[b]V. 2004: Verdict 2004, A: Attempt; M: Murder

[c]E9 Elffers with 9 cases etc. (All in JKZ)

[d]GGJ7 Gill, Groeneboom and De Jong with 7 cases etc. (in all wards)

[e]Remarks: Dropped: dropped from statistical calculations; out: outside Lucia's shift.

[f]We are not aware of any set, including this case.

Table 3: JKZ, original and corrected data

Shifts in JKZ	with incident	without incident	Total	Shifts in JKZ	with incident	without incident	Total
Lucia	8	134	142	**Lucia**	7	135	142
Others	0	887	887	**Others**	4	883	887
Total	8	1021	1029	Total	11	1017	1029

The original data in table 3 are from the revised report to the court Elffers [May 29 2002]. In the first report to the court Elffers [May 8 2002] the case of (18/04/01) was included. To test the stability of the calculations (see table 1) we have used the original data presented here.

Table 4: RKZ-42, original and corrected data

Shifts in RKZ-42	with incident	without incident	Total	Shifts in RKZ-42	with incident	without incident	Total
Lucia	5	53	58	**Lucia**	1	57	58
Others	9	272	281	**Others**	9	272	281
Total	14	325	339	Total	10	329	339

The original data in table 4 are from revised report to the court Elffers [May 29 2002]. The revised data follow the court, the remaining incident being that of (12/11/97). The other four incidents have been declassified, not moved.

Table 5: RKZ-41, corrected data

Shifts in RKZ-41	with incident	without incident	Total	Shifts in RKZ-41	with incident	without incident	Total
Lucia	1	0	1	**Lucia**	1	2	3
Others	4	361	365	**Others**	4	359	363
Total	5	361	366	Total	5	361	366

Since the original data contained only one shift with Lucia, the statistician of the court estimated the probability of having an incident on one shift directly. The two extra shifts without incidents have been discovered by the defense lawyer.

References

Lucia de Berk. *Lucia de B.* Uitgeverij De Arbeiderspers, 2010. ISBN 978-90-295-7262-0.

A.C. Brantley and R.H. Kosky. Serial murder in the Netherlands: A look at motivation, behavior, and characteristics. *FBI Law Enforcement Bulletin*, 74(1):26–32, 2005.

A.A. Derksen. *Lucia de B. Reconstructie van een gerechtelijke dwaling.* Veen Magazines BV, 2007. ISBN 978-90-857-1048-6.

A.A. Derksen and M. Meijsing. The fabrication of facts: The lure of the incredible coincidence. In *Legal Evidence and Proof,* H. Kaptein, H. Prakken and B. Verheij (eds), pages 39–70. Ashgate, 2009. ISBN 978-0-7546-7620-1.

H. Elffers. Nadere analyse: Verdeling reanimatie- en overlijdensgevallen in het juliana kinderziekenhuis en rode kruisziekenhuis. Unpublished report to the Court, May 29 2002.

H. Elffers. Verdeling reanimatie- en overlijdensgevallen in het juliana kinderziekenhuis en rode kruisziekenhuis. Unpublished report to the Court, May 8 2002.

S.E. Fienberg and D.H. Kaye. Legal and statistical aspects of some mysterious clusters. *J. R. Stat. Soc. (A)*, 154:61–74, 1991.

R.D. Gill. Lies, damned lies, and legal truths. In L. Mommers, H. Franken, J. van den Herik, F. van der Klaauw and G.J. Zwenne *Het Binnenste Buiten (Liber Amicorum ter Gelegenheid van het Emeritaat van prof.dr. Aernout H.J. Schmidt, hoogleraar Recht en Informatica de Leiden)*, pages 39–50. eLaw@Leiden, Law Faculty, University Leiden, 2009. ISBN 978 90-815196-1-8.

G. Kuhn. Circadian rhythm, shift work, and emergency medicine. *Ann Emerg Med*, 37:88–98, 2000. doi: 10.1067/mem.2001.111571.

Hans A. J. M. Kuijsten, Sylvia Brinkman, Iwan A. Meynaar, Peter E. Spronk, Johan I. van der Spoel, Rob J. Bosman, Nicolette F. de Keizer, Ameen Abu-Hanna, and Dylan W. de Lange. Hospital mortality is associated with icu admission time. *Intensive Care Med*, online first, 2010. doi: 10.1007/s00134-010-1918-1.

R. Meester, M. Collins, R.D. Gill, and M. van Lambalgen. On the (ab)use of statistics in the legal case against the nurse Lucia de B. *Probability and Risk*, 5:233–250, 2007. With discussion by David Lucy.

M. de Noo. *Er werd mij verteld, over Lucia de B.; de klok, de klepel en de klokkenluider.* Aspekt, 2010. ISBN 978-90-591-1974-1.

B.C. Yorker, K.W. Kizer, P. Lampe, A.R.W. Forrest, J.M. Lannan, and D.A. Russell. Serial murder by healthcare professionals. *J. Forensic Sci.*, 51: 1362–1371, 2006.