University of Michigan Journal of Law Reform

Volume 50 | Issue 3

2017

Short Fall Arguments in Court: A Probabilistic Analysis

Maria Cuellar
Carnegie Mellon University

SHORT FALL ARGUMENTS IN COURT: A PROBABILISTIC ANALYSIS

Maria Cuellar*

I will be talking today about how statistical arguments are used in court, specifically in cases of Abusive Head Trauma in which the defendant has claimed that an accidental short fall, and not shaking or child abuse, has caused the child's injuries. So actually the *Johan* case[1] that Peter Aspelin was talking about leads perfectly into this. In particular, I will be talking about one specific paper by David Chadwick et al. from 2008.[2] In this paper, he and his colleagues calculate the risk that a child, a young child, will die from a short fall.[3] They find that the risk is less than one in a million.[4] In fact, 0.48 in a million.[5] I will be providing some criticism of how this quantity gets used in the court; then I will provide an alternative method for figuring out whether a short fall could have caused this specific child's injuries, and I will close by talking about the challenges that we have today with respect to data.

What I will call the Chadwick paper is a study that was written by David Chadwick, from the University of Utah, with seven co-authors in 2008.[6] It has the unambiguous title of "Annual Risk of Death Resulting from Short Falls Among Young Children: Less than One in a Million."[7] And as I said, 0.48 in a million is the quantity they calculate.[8] So to calculate this value, they used a database called the EPIC database, which stands for the Epidemiology and Prevention for Injury Control database. I will talk in detail about how they found the value shortly.

---

* PhD student at Carnegie Mellon University; joint program in Statistics and the Heinz School of Public Policy and Management. BA, Reed College; MS, Carnegie Mellon University.

1. Joseph Shapiro, *Dismissed Case Raises Questions on Shaken Baby Syndrome.* NAT'L PUB. RADIO (Dec. 21, 2012), http://www.npr.org/2012/12/21/167719033/dismissed-case-raises-questions-on-shaken-baby-diagnosis.

2. *See generally* David L. Chadwick et al., *Annual Risk of Death from Short Falls among Young Children: Less than 1 in 1 Million*, 121 PEDIATRICS 1213 (2008) (describing results from a systematic review that found only six possible fall-related fatalities of young children in a population of 2.5 million young children over a five year period).

3. *Id.* at 1214.

4. *Id.*

5. *Id.*

6. *See generally id.*

7. *Id.* at 1214.

8. *Id.*

This paper is very careful in defining what a short fall is, specifically. They define it as a fall from a height of less than five feet in which the horizontal velocity is no faster than that which a child could have achieved alone.[9] They define their population very specifically. They choose the population of infants, ages zero to five.[10] This is the only paper of which I am aware that provides a numeric estimate for how likely it is for a short fall to cause death. For this reason, it is used widely both in the literature and in court.

A few other papers have also looked at the issue of short falls. I will talk about two of them here. The first one is John Plunkett's paper from 2001.[11] In this paper, he reviewed eighteen cases of deaths in infants that had been caused by short falls,[12] and he reviewed each case in detail. What is important about this paper is that he showed that deaths can indeed be caused by short falls.[13] But he does not provide a specific numeric estimate that will tell us how common it is for a child to die from a short fall. The paper by David Moran, Keith Finley, Patrick Barnes, and Waney Squire from 2012 titled "Shaken Baby Syndrome, Abuse Health Trauma, and Actual Innocence: Getting it Right"[14] has a short section that criticizes the Chadwick paper.[15] The Moran et al. paper also does not provide a numeric estimate for how common it is for a child to die from a short fall.[16]

To arrive at their estimate of 0.48 in a million, Chadwick et al. used the EPIC database from California, which has death records from various medical examiners, in addition to other sources.[17] They used data from 1999 to 2003,[18] and they counted the number of infants who had died from a short fall in California in this time period.[19] Just six were not disproven as short fall deaths.[20] Then

---

9. *Id.*

10. *Id.*

11. *See generally* John Plunkett, *Fatal Pediatric Head Injuries Caused by Short-Distance Falls*, 11 AM. J. FORENSIC MED. & PATHOLOGY 1 (2001) (concluding that an infant or child may suffer a fatal head injury from a fall of less than three meters (ten feet)).

12. *Id.* at 2.

13. *See id.* at 10 (concluding the falls from less than three meters may be fatal).

14. *See generally* Keith A. Findley, Patrick D. Barnes, David A. Moran & Waney Squier, *Shaken Baby Syndrome, Abusive Head Trauma, and Actual Innocence: Getting It Right*, 12 HOUS. J. HEALTH L. & POL'Y 209 (2012) (providing a history of SBS, its evidence, and calling for more collaboration between medical and legal communities to "get it right").

15. *See id.* at 247–48.

16. *See generally id.*

17. *About the Fatal (Death) Data, Nonfatal Patient Discharge (Hospitalization) Data, and Nonfatal Emergency Department (ED) Data*, CAL. DEP'T OF PUB. HEALTH, http://www.cdph.ca.gov/HealthInfo/injviosaf/Pages/EpiCenterdata.aspx#fatal (last visited Jan. 13, 2017).

18. Chadwick et al., *supra* note 2, at 1214.

19. *Id.*

20. *Id.*

they divided six by the number of all infants in California during this time period.[21] So if you wanted to convert this to a probability statement, you might say this is the probability of an individual having a death and short fall, given that this individual is an infant in California, in the period of 1999 to 2003. They found that this value was 0.48 in a million.[22]

What I claim is that in court, implicitly, what is being argued is that this value of 0.48 in a million is equivalent to the probability that the child had a short fall given the evidence in that criminal case, the evidence being this is an infant with head trauma and death. Assuming that the 0.48 in a million is indeed this probability implies that the probability that the child was shaken, given the evidence, is the complement of that. That is, one minus 0.48 in a million. When one computes this number, one gets a probability of 99.999% that the child was shaken, given the evidence.

The problem is this argument is incorrect. The argument is, in other words, that short falls almost never cause deaths, and, therefore, this defendant must have shaken the baby or must have abused the baby. As opposed to other authors who have criticized Chadwick's quantity, I actually believe that Chadwick's quantity was calculated correctly. I, myself, went into the database, the EPIC database and counted the number of short falls in this time period and found that, indeed, it is 0.48 in a million. But, my argument is that it is the implicit argument used in court that is incorrect.

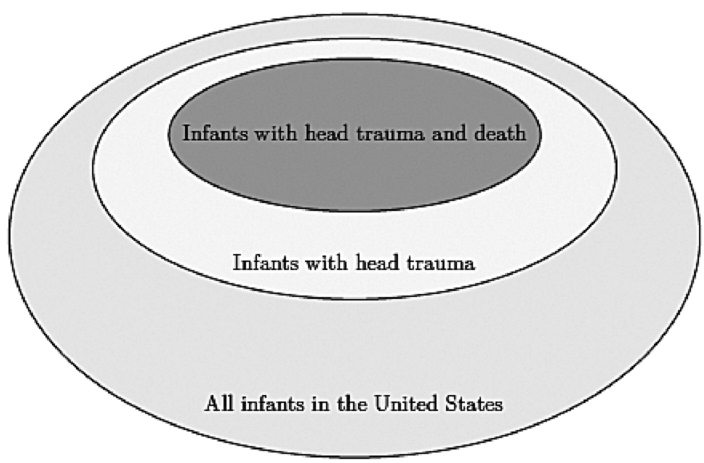
Now, I will talk about a few of the criticisms I have about using the Chadwick quantity in court in the way in which it has been used. The first criticism is that, as we have seen in the previous talks, rare events are not impossible. Chadwick and co-authors found the deaths due to short falls are 0.48 in a million, but today in the United States there are twenty-four million infants. So, we would expect to see at least twelve infants die from a short fall this year. A child in a specific legal trial could be one of these twelve. This argument was mentioned earlier this morning by Peter Aspelin, and it was also mentioned in the paper by Findley, Moran et al. from 2012[23] that I mentioned earlier. It has also been used by various expert witnesses in the courtroom. For illustration purposes, I calculated a few other values to show that they are also very small. The prevalence of having Abusive Head Trauma (including cases of Shaken Baby Syndrome), according to the specific database that I

21. *Id.*
22. *Id.*
23. *See generally* Findley et al., *supra* note 14 (providing a history of SBS, its evidence, and calling for more collaboration between medical and legal communities to "get it right").

will mention later, is eighty-five in a million. All of these are very small values. But again, they are not zero.

The second criticism I have for using the Chadwick quantity in the way in which it has been used is that one must restrict the population in light of the evidence that has been provided in the case. The way that Chadwick et al.[24] calculated their quantity was for the entire population of infants. If you recall, I said that they divided six by the number of all infants in California in that specific time period. But we know some more information about this infant, not just that he is a child in California. This is not a healthy child. This child has head trauma and has died. Or, in a different case, we might have a child who has not died, but also has head trauma. What we must do is restrict the population in light of this additional information. So, in a probability statement, we would write this as the probability that a child was shaken, given that this is an infant with head trauma and death. This "given" part is called "conditioning." By conditioning, we are restricting the population to a smaller subset, as the figure below shows.

$$P(\text{ Shaken } \mid \text{ Infant with head trauma and death })$$



The large oval represents all the infants in the United States. A small group will be those who have head trauma and an even smaller subset of that will be the ones who have head trauma and death.

This argument has been used by various experts in court. Some have called it the problem of having the wrong denominator. What

24. *See generally* Chadwick et al., *supra* note 2.

they mean is that one needs to focus on the population that is relevant in light of the evidence.

The third criticism is that this quantity calculated by Chadwick et al.[25] is calculated in isolation. What I mean is that they calculate the probability that one event happens—that is, that a child will die from a short fall. But they do not calculate any of the probabilities for any other possible events that might have caused the outcomes, in order to provide a comparison of the two occurrences. We need to find out how likely these other possible causes are before we make any conclusions about what is more likely. So for example, if we have an event that has a probability of happening of one in a million, and we compare it to another one that has a probability of one in four hundred million, all of a sudden, one in a million seems very common. So, what we would really like to estimate are fractions like

$$\frac{P(\text{Short fall} \mid \text{Evidence})}{P\ (\text{No short fall} \mid \text{Evidence})} \text{ and } \frac{P(\text{Shaken} \mid \text{Evidence})}{P\ (\text{Not shaken} \mid \text{Evidence})}$$

The one on the left is the quantity that answers the question, "Were these injuries caused by a short fall?" Here we are dividing the probability of that a child had a short fall, given the evidence, over the competing hypothesis, which is that the child did not have a short fall, given the evidence. Eventually what we would like to get is a quantity for whether these injuries were caused by shaking or by child abuse with intentional actions. That ratio is the fraction on the right: the probability of shaking, given the evidence, divided by the probability of not shaking, given the evidence. Again, the evidence being that the child has had trauma and death, or maybe just had trauma, depending on the child's specific circumstances. So, we can expand this ratio that I mentioned, the probability of shaking, given the evidence, over not shaking, given the evidence, by looking at the denominator specifically. We can do this by expanding the probability that the child was not shaken into the probability that the child had an accidental short fall, given the evidence, plus the probability that the child had other causes, given the evidence. Other causes might be, for instance, some of the conditions that have been described by Patrick Barnes and others this morning, conditions that might mimic the symptoms of Abusive Head Trauma or Shaken Baby Syndrome. These include things like rickets or Vitamin D deficiency. A question that might follow is, "How do we really obtain a value for this ratio that the probability

25. *See generally id.*

of the child was shaken, given the evidence, divided by the probability the child was not shaken, given the evidence?"

For the purposes of this talk, and just for this talk, I will define the evidence as the triad of symptoms mentioned previously, defined here as retinal hemorrhage, cerebral edema, and subdural hemorrhage.[26] I am aware that there has been some debate about what are the symptoms that really indicate Shaken Baby Syndrome and Abusive Head Trauma. Here, I just included the classical triad, but we could include any number of symptoms that a researcher would like to include in the analysis.

What we can do with this quantity is that we can expand it by using Bayes' rule. So the fraction on the left-hand side becomes a product of two fractions on the right-hand side.

$$\frac{P(\text{Shaken} \mid \text{Evidence})}{P(\text{Not shaken} \mid \text{Evidence})} = \frac{P(\text{Evidence} \mid \text{Shaken})}{P(\text{Evidence} \mid \text{Not shaken})} \frac{P(\text{Shaken})}{P(\text{Not shaken})}$$

On the right-hand side, the first fraction is the probability (with the evidence given) that the child was shaken over the probability (with the evidence given) that the child was not shaken. Then, that is multiplied by the second fraction—the odds that the child was shaken. So, the first factor on the right hand side is the quantity that a statistician might be able to address. This is the quantity that one could, in theory, obtain from the data. And that is all a statistician can calculate. The second factor, the odds, is what can be calculated for a specific case. This quantity is up to the jury and the judge to decide, not the statistician. The evidence that could be relevant in the odds factor is, for example, the amount of time that this person was spending with the child, or whether the child has injuries from prior abuse.

	Evidence	Does not have the evidence
Shaken		
Not Shaken		

26. Goran Hogberg et al., *Circularity Bias in Abusive Head Trauma Studies Could Be Diminished with a New Ranking Scale*, 6 Egyptian J. Forensic Sci. 6, 7 (2016).

The table above shows the quantities we would need in order to calculate the first fraction on the right-hand side. In order to calculate this properly, we would need to fill in this table with the number of children who were shaken and had traits corresponding to the evidence (e.g. triad and death), the ones who were shaken and did not have traits corresponding to the evidence, the number of children who were not shaken and have traits corresponding to the evidence, and the ones who were not shaken and did not have traits corresponding to the evidence. To do this properly, we would need to use a database that has a sample that is representative of the population at large. We could not just use a few examples of similar cases in a hospital, for instance, since this could lead to selection bias. Since they are rare cases, the sample should be large as well, because when you have very few events, there might a great deal of variability, which could lead to misleading results. When you have a larger sample size you might be able to arrive at more stable results.

Now I do not know if filling in this whole table seems suspicious to any of you, or at least like a difficult task. For example, if the evidence is that the child had the triad, one of the values necessary is the number of children who were shaken and did not have the triad. You would need to have a database that correctly reports the values necessary for that calculation, and it is possible that although a child could have been shaken and did not acquire the triad, this might not be recorded in the database. So, there are many challenges with filling in this table in order to calculate the value properly. If you were dealing with a case in which the child had died in addition to having the triad, for instance, you would need to not just include the cases that had the triad and no triad, but the cases that had the triad and death and no triad and death in order to calculate these values properly.

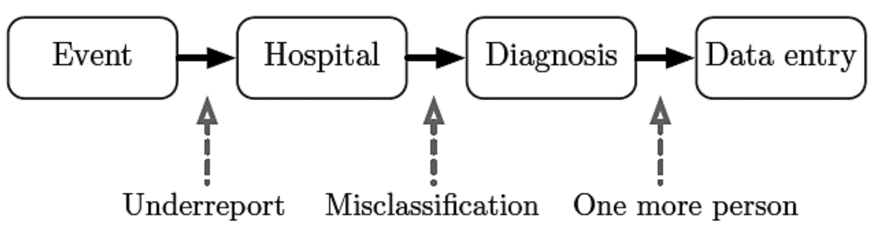
I looked at some major databases that one might use to fill in this table. The first one is the one that Chadwick et al.[27] used, the Epidemiology and Prevention for Injury Control database from California.[28] There are a few problems with this database from a statistical point of view. It has a very small sample size, which means it might not be very useful for observing rare events. It is not representative of the United States population because it is limited to California. I looked at another database called the Kids' Inpatient

27. *See generally* Chadwick et al., *supra* note 2.

28. EPICENTER: CAL. INJURY DATA ONLINE, http://epicenter.cdph.ca.gov (last visited Jan. 13, 2017).

Database.[29] To clarify, this is not the same one that Patrick Barnes mentioned this morning, the KidsData.org.[30] They are two different databases but have similar names. The Kids' Inpatient Database is a very large database that has a sample size of about three million children, and it has been updated every three years since 1997. It has information from hospital records all around the United States. The database is built from a random sample of hospital records in about 4500 hospitals, and it is very rich. It contains information about the primary and secondary diagnoses for each record, which include up to fifteen diagnoses. These might include things like the triad. It contains information about external causes of injury, which is very detailed and varied data. For instance, it has information about railway accidents, poisoning, and short falls. It contains information about whether the child died, and many other variables.

When I was starting this project of analyzing the Chadwick[31] quantity and how it is used it court, I thought I could provide a new estimate and fill in this table with the Kids' Inpatient Database because it seemingly has all the variables I would need, and it has a large enough sample size. I started doing that and then I realized that even this dataset, which is so rich and large, is highly biased. As I mentioned before, the cases where the child has fallen but does not have the triad would not be included there. To visualize why there is bias in the data, we draw a possible sequence of events below about the data gathering process.



Maybe the event happened at home—that is, maybe someone dropped the child at home. It makes sense that then the child would be taken to a hospital, but perhaps some children will not be because the parents might believe there has been no real harm. At that point, the children who were not taken to the hospital are missed in the database. That is, there is some underreporting at that stage.

29. *Kids' Inpatient Database*, Dep't of Health and Human Servs.: Agency for Healthcare Research and Quality, http://www.ahrq.gov/research/data/hcup/kid.html (last visited Jan. 13, 2017).

30. Kidsdata.org, www.kidsdata.org (last visited Jan. 13, 2017).

31. *See generally* Chadwick et al., *supra* note 2.

Then, after the child arrives at the hospital, the physician might make a diagnosis, and this is subject to cognitive or human bias, like the one discussed this morning. In some hospitals, some cases might be categorized as a short fall, and some cases might be categorized as shaken. That is a misclassification problem with the diagnosis. Then, in the final stage of the data gathering process, someone needs to enter this information about the diagnosis into a database using codes used in the Kids' Inpatient Database (i.e. the ICD9 or ICD10 codes, which are the International Classification of Diseases). If this gave rise to data entry errors, we might say this does not generate a bias because the errors could go in either direction—that is, the cases might be both underreported and overreported. But we could also imagine that a person in charge of data entry might have a certain opinion about Shaken Baby Syndrome or about this specific case, and so he or she might enter the data in a specific way that is systematically biased. This chain of events shows that the data from one of the best sources that I could find is flawed and has inherent biases.

To summarize, the Chadwick[32] quantity should not be used in court because of the criticisms I mentioned. The first is that rare events are not impossible. The second is that we need to restrict the population in light of the relevant evidence, and the Chadwick calculation does not do that. The third is that one must compare competing hypotheses, again, in light of this evidence. Otherwise, this quantity, the Chadwick quantity in isolation, is of no use to the court. I argued that a better alternative to the Chadwick quantity to determine whether a child had a short fall would be to calculate this ratio,

$$\frac{P(\text{Short fall} \mid \text{Evidence})}{P(\text{No short fall} \mid \text{Evidence})}$$

But what the court is really concerned with is whether the defendant shook the child, so we need to calculate this ratio,

$$\frac{P(\text{Shaken} \mid \text{Evidence})}{P(\text{Not shaken} \mid \text{Evidence})}$$

To calculate it, the information must be combined in a very specific way, with additional information from the case, to arrive at the proper final ratio.

32. *See generally id.*

But, as I found in my analysis, the data that are available today are not of high enough quality to allow for a proper calculation of this quantity. So, any calculation that I make with this table today, if I were to show it to you, would be wrong. It would be false. I would like to say that a correct statistical argument about the United States population cannot be made to answer the question, "Could this child's injuries have been caused by a short fall?" I would like to close by saying that if you do use a statistical argument in court, you need to make sure it is based on high-quality data—which is not available today—and an appropriate probabilistic analysis. You need to include comparisons of competing hypotheses and not just one rare value in isolation, because that comparison is what is relevant in court. Thank you very much.