

Original Paper

Recognizing the Misuse of Probabilistic Language and False Certainty in False Accusations of Child Abuse

Steven C. Gabaeff¹

¹Clinical forensic medicine, Emergency Medicine, Healdsburg, CA, USA

Received: November 26, 2020 Accepted: December 12, 2020 Online Published: December 21, 2020

doi:10.22158/jrph.v4n1p1

URL: <http://dx.doi.org/10.22158/jrph.v4n1p1>

Abstract

Probabilistic language is language used to convey mathematical probabilities in narrative form including terms like “highly likely”, “concerning for”, “suspicious of”, and many others. PL can be used in conformance with standards elucidated in Forensic Epidemiology or misused with intentional imprecision, when not justified, to promote a misdiagnosis of abuse, with dire consequences. The application of actual probability analysis using tested mathematical models, like Bayes Theorem, is essential to an analysis of the actual probability of abuse in a specific case to avoid false accusations of abuse. Consideration of the prior odds of abuse combined with calculations the reliability of nonspecific and/or unreliable criteria or “indicators”, is being disregarded by child abuse pediatricians to justify diagnosing abuse with statements of false certainty that depend on the misuse of probabilistic language. These suppositious statements of false certainty are the sine qua non of accusatory expert opinion. Currently, and unfortunately, only scientists and physicians with the requisite advanced knowledge of these issues only detect false certainty. When probabilities and evidence based science are studied and applied, deep flaws in the fund of knowledge of child abuse pediatrics have been exposed. On balance, there is an emerging reality that the collective suffering of falsely accused families may dwarf the horrific impacts associated with real abuse. It also exposes iatrogenic abuse as possibly the most common form of prosecuted child abuse in the legal system. A false accusation of child abuse is child abuse. The misuse of probabilistic language to convey false certainty and its ramifications for innocent caregivers is discussed herein and must be prevented.

Keywords

False accusations, child abuse, probabilistic language, Bayes, probability, AHT, SBS, misuse

1. Introduction

Probabilistic Language (PL) is language used to convey mathematical probabilities in narrative form. PL can be used appropriately in scientific and forensic writings, reports to Law Enforcement (LE), and testimony, to convey a range of probabilities defined in Forensic Epidemiology (very unlikely to very likely) as noted in Table 1 from Goodman and Royal (Goodman, Royal, & Royal, 1988).

Adherence to these agreed upon probabilities, and their mathematical weight, should be regarded as proper scientific and forensic/legal use. Disregard of the mathematics of these narrative terms is misuse. Beyond these terms, other vague probabilistic constructs, with no agreed upon mathematical reference especially in expert opinion, create other opportunities to misuse PL. Terms like “concerning for”, “suspicious of”, “consistent with” and many others discussed below, are very commonly used. In fact, they are, under the direction of the American Academy of Pediatrics (AAP), the preferred linguistic constructs in generating accusation of child abuse. These probabilistic constructs, advanced by authority figures, can have an outsized influence on colleagues, LE, and lay people that are unsupported by actual probability or science.

Probabilistic language is misused in pediatric child abuse by highly placed and titled, child abuse services providers, often working in concert for greater impact. This synchrony is often represented as “consensus”. With consensus, correct or not, evidentiary standards used in some jurisdictions, create an extraordinary potential to misrepresent the probability of abuse, with very negative impacts. The misuse, when not supported by actual probability, conveys “false certainty”.

False certainty, expressed with PL, has been shown to be effective in child abuse investigation in supporting pre-conceived suspicions (beliefs) of abuse when actual probability analysis of each finding and the final diagnosis, would not be supportive of abuse over nonabuse; often by huge margins (as discussed below).

Like-minded providers, within the child abuse establishment, in specialties linked to child abuse medicine, will misuse PL as well. The specialties of ophthalmology, radiology, orthopedics, neurosurgery, pediatric critical care, with formal, or informal training, have adopted the phrases, rhetorical strategies, and protocols advocated by the American Academy of Pediatrics (AAP), with cumulative impacts. The false certainty expressed by the imprecise misuse of PL, initiates legal accusations of abuse, followed by dependency and criminal proceedings. Innocent caregivers, falsely accused of abuse, regardless of legal end results, are severely damaged by false accusations. Innocent caregivers become victims themselves, with life altering legal and sociopathic impacts from which almost none recover. Whether intentional, or not, the misuse of PL in false certainty statements, is integral to the creation of an abuse narrative when no abuse has occurred.

Overall, the recognition of the use and misuse of PL, in combination with the application of mathematical probability analysis, using tools like Bayes Theorem (Bayes, 1764) applied in detail below, is the key to decreasing the misuse of PL and identifying actual low probabilities of abuse when abuse seems intuitively unlikely.

The misuse of probabilistic language to convey false certainty is the subject of this article.

2. Understanding Probabilistic Language and False Certainty

2.1 False Certainty

False certainty, by definition, is based on beliefs that are not supported by actual probability analysis. In pediatric child abuse medicine, calculated mathematical probabilities based on circumstances, findings, and evidenced based analysis, based on my review of over 3000+ contested cases of alleged abuse, are essentially never used. When such analyses are offered for consideration, and very unfavorable probability analyses are generated, child abuse pediatricians and their colleagues are dismissive. In the process, they eschew the scientific methodology and error analysis that they are legally required to subscribe to under the Daubert (Daubert, 1993) criteria for expert testimony, in favor of suspect statements of consensus (Choudhary, et. al. Consensus Statement, 2019). Establishment child abuse pediatricians and their juniors, and close colleagues and organizations with little or no direct involvement in child abuse research, populate this confederacy of the like-minded. Consensus or “general agreement/acceptance” is a lesser evidentiary standard, inappropriate for scientific testimony and is referred to as Kelly-Frye (Frye, 1923). The actual probability of abuse, even when under 5%, and even when nonabuse is supported by other evidence, is declared irrelevant. Beliefs, collated into consensus, from authority figures, can, and do, trump actual calculated probability, and this is effectuated with the misuse of PL.

It is axiomatic that beliefs are not science. Furthermore, false certainty statements, when subjected to actual calculated probability analysis, as demonstrated below, would not reach any accepted legal or medical thresholds of conventional evidentiary reliability. In child abuse pediatrics, false certainty leads to false positive medical determinations of abuse, and increased probability of a false conviction, based on false evidence. False certainty, once understood, is obvious, unsettlingly common, and should not, and cannot, be abided.

2.2 Probabilistic Language in Expert Opinion

When used in expert opinion, PL adds linguistic weight to opinions and conveys varying degrees of certainty. PL can create a sense of certainty at one end of the spectrum (“abuse is highly likely”) and implausibility on the other (“household falls are very unlikely, or never, cause serious injury”).

Current trends in child abuse documentation advocating the misuse, and inherent deception, of PL, have been precisely crafted, assiduously promoted, and widely adopted by the American Academy of Pediatrics (AAP) and its subcommittee: the Committee on Child Abuse and Neglect (COCAN). COCAN is the controlling entity in all matters in hospital based child abuse pediatrics (Associated Press, 2009).

PL has been fashioned to offer some indemnification from allegations of malpractice after diagnosis or more accurately, misdiagnosis of abuse has occurred. By using probabilistic language, each term, inherently acknowledges some exceptions and provides a modest defense to including abuse as a

diagnostic possibility, even when abuse is not probable, plausible, or has not occurred. However, and more importantly, in child abuse investigation the terms are misused very effectively in this forensic framework, as insider code for “diagnostic of”. The code word scheme is seen, case after case, and validated by the examination of the timing of documentation, the initiation of legal action, and the actions of prosecutors and police. This is combined with an expected willingness by the child abuse pediatrician to investigate, advocate for an abuse diagnosis, consult, formally accuse, testify, and defend their abuse diagnosis in court and other forums. From a forensic perspective, and to use PL as child abuse pediatricians do, there is a “concerning” array of conflicts of interests among different roles assumed by a single mesmerizing power: the child abuse pediatrician.

2.3 The Scope of Probabilistic Language (PL)

AAP actively promotes, formally, or via more informal communications, the use of terms like “suggestive of”, “evidence of”, “concerning for”, “suspicious of”, “suspicious for”, “consistent with”, “associated with”, “highly associated with”, “cannot exclude”, “sufficiently explained by”, “insufficiently explained by”, “significantly more”, “quite concerning”, “worrisome”, “strong probability”, “often”, and “rarely,” as well as other similar terms. Near simultaneous activation of the use of new rhetorical strategies is evidence of specialty-wide dissemination capabilities for directing matters like this, via AAP channels.

PL can be used with qualifiers like “almost, rarely, and very”; terms, which when used inaccurately, obfuscate actual probability even further.

Adverbs and adjectives like “most”, “extremely”, “more”, “quite”, “somewhat”, “very”, “very much”, “highly”, can appropriately, or inappropriately, add additional emphasis and weight to expert opinion.

Expressions of “doubt” and the use of terms like “reassuring” also convey imprecise expressions of actual probability, often used to eliminate diagnostic possibilities that may be quite possible, plausible, and, many times, more probable.

Words like “never”, “always”, can be used and misused as well. The absolute certainty (or false certainty) these words, never and always, express, invoke the circumstance of no exceptions allowed. “Short falls never cause serious injury”, is a commonly invoked myth in child abuse pediatrics. The child abuse pediatrician who states we should believe that accidental falls from 2 story windows are the only accidental falls that should be considered nonabuse, is supporting a deception. The statement is false. Caregivers and witnesses to household falls, reporting such events, are judged to be liars by child abuse pediatricians (Chadwick et al., 1991) after describing a falls that caused injury (serious or just minor bruises).

There are few, if any, “never” or “always” statements applicable in medicine and this is particularly true in child abuse pediatrics. Always and never, as PL in child abuse pediatrics, connote exclusion, or more accurately, disregard for exceptions, even when exceptions are known to occur. While “rare” exceptions may be off handedly acknowledged by accusers, they are then commonly dismissed as irrelevant. With falls, the exceptions (serious injuries), are rare as a percent of all falls, but the absolute

number can be tracked year after year and is substantial and relevant. The assumption by child abuse pediatricians is that a rare event should not be considered as legitimate or plausible, and can be dismissed. This is a grave mistake. This is what might be called the “rarity argument”. With a corroborating clinical and social history, and a reasonable nonabuse diagnostic possibility present, the rarity argument is yet another rhetorical trick based on projecting false certainty in code. To complete the sequence with falls, the history is dismissed and abuse is elevated, by default, as the only “possible” diagnosis. In real life, with 4 million newborns a year resulting in 8 million toddlers, if two fall in a day (out of 8 million) and the result is a serious injury, that is 730 serious falls a year, some of which will result in death. Essentially all such cases end up at a children’s hospital. There, the 730 serious accidents with injury are highly likely to be misdiagnosed as abuse under the fallacious rubric of the rarity argument (Chadwick, 2008). To be clear, short falls causing serious injury are rare, however, household falls produce an absolute number of cases that comports with some number of cases where accidents are misdiagnosed as abuse (Atkinson, van Rijn, & Starling 2018).

2.4 The Proper Use and Misuse of Probabilistic Language

Parenthetically, and somewhat ironically, in general medical use, (as compared with child abuse pediatrics), many of the vague probabilistic terms mentioned above, including the standardized forensic terms, are used legitimately in medicine to express a fundamental uncertainty.

Examples like “consistent with”, “concern for”, and “associated with”, connote possibility, plausibility, and probability, as in a differential diagnosis. They should not be used to connote legal or medical certainty, or a final diagnosis, simply because they do not. The danger of saying, relying on, and acting as if these terms are equivalent to “diagnostic of”, without objective evidence, is clear in this real life example. “Concern for” appendicitis in an emergency department, without testing to confirm the diagnosis, should never lead to a diagnosis of appendicitis and the requisite emergency appendectomy. With only about 1 out of 100 abdominal pain patients in the emergency department actually having appendicitis, acting on a “concern”, by operating, without objective testing, and disregard of known probabilities, will be wrong 99% of the time. This would lead to 99 cases out of 100 in which an unnecessary operation will be done and the doctor will be sued for malpractice for each of them. With this liability looming, testing is the standard of care and accuracy is near 100%.

The evaluation of the actual odds of abuse is central to eliminating the misuse of PL to generate legal action in cases where abuse is very unlikely (under 5% likely or has not occurred). Conversely, such analysis can strongly support, to near certainty (beyond reasonable doubt or to a reasonable medical probability; greater than 95% likely), that abuse has not occurred. After probability analysis it should be clear that there are many cases when abuse is barely worthy of being in a differential diagnosis.

3. Bayes Theorem...The Quantification of Probability and Reliability

3.1 Evidence

Evidence, has been defined by Goodman (Goodman, Royal, & Royal, 1988) as a “the property of data that makes us alter our beliefs about how the world around us is working...evidence is the basis on which we derive inference...Inference is the calculus of decision making”. The “weight of evidence,” a legal concept used to frame the analysis of how much dependence on a particular piece of evidence is warranted, is another more subjective measure of the evidence’s reliability and how it will be received. Weak, moderate, and strong are the adjectives used in the most general terms, but more precision is superior and attainable. The reliability of a piece of evidence can, and should be calculated with the 200+ year old Bayes Theorem. Bayes is the key concept in understanding false certainty.

3.2 Overview of Bayes Theorem

Bayes theorem in its most simplistic form, subject to more detailed analysis below, states the “prior odds” of subject event (abuse) occurring in a defined population, integrated with the reliability of a selected test or criteria (an “indicator”) to support the event (abuse), produces a probability (“posterior odds”) that the subject event actually occurred. An indicator can be a finding like retinal hemorrhage [RH] or subdural hemorrhage [SDH], a lab result, a behavior, a circumstance, etc. The indicator’s reliability to deliver the correct result (expressed as a “likelihood ratio-LR”) can be calculated using conventional sensitivity and specificity. When the measure of reliability (LR) is applied to the prior odds of the event in a defined population, the result is the actual probability of the event being studied (abuse) occurring in the defined population (the posterior odds). The higher the LR, the more likely the indicator will deliver the correct end result. What is key however, is the awareness that result of this calculation is not a binary answer (abuse, did or did not occur), but another ratio: the “posterior odds”. The posterior odds is not the definitive answer to the question of abuse or not: the binary choice. It is a probability that the indicator, test, or finding, for the defined population, will deliver the correct answer (a true positive determination of abuse; i.e., abuse is alleged and abuse occurred).

Prior odds x Likelihood ratio = Posterior odds

When posterior odds are high, the probability of the subject event (abuse) is high. Conversely, when it is very low, the probability of event can be near zero, and everything in between.

3.3 Understanding Odds and Probability

Odds and probability are two mathematical ways to express the same thing. Probability can be expressed as a percentage of certainty. Odds are expressed as ratios. Both odds and probability percentages are used to quantify reliability, and have similar narrative equivalents in forensic epidemiology (see Table 1). Probability of 5%, equal to odds of 1 in 20, is officially regarded as “very unlikely”. Such low odds are not evidence. Odds of 19 in 20 is 95% probability and are considered “very likely”, and this comports with the reasonable degrees of forensic certainty standard used in criminal proceedings and is “strong evidence”. The 66%, clear and convincing legal standard is regarded in forensic epidemiology as “slightly more likely than not”. The 50% preponderance standard

is “indeterminate” according to forensic epidemiology, yet oddly remains in wide use in dependency proceedings.

Anything under 95% technically defines reasonable doubt and as the probability of guilt decreases from that 95% threshold, not guilty is the only justified legal outcome in criminal proceedings. The lower legal standards, on this basis, are even more suspect as to reliability.

3.4 Prior Odds

Mathematics discriminates between reliable science, actual evidence, and beliefs projected as false certainty. False certainty, the key component of false accusations, is communicated by the misuse of probabilistic language. Strict adherence to the quantification of probabilities prevents beliefs from morphing into false testimony.

The first step in quantifying probability with Bayes is the determination of the calculated numeric probability of an event like abuse occurring in a defined population. The probability is called the “prior odds” of an event and each calculation of prior odds applies only to the defined population in that calculation. The prior odds change with different populations; the prior odds of abuse among all parents, is markedly different from say, parents who are previously convicted child abusers.

If the studied event is abuse, we must identify a rate of abuse for a defined population that would include a suspect. From the broadest perspective, we can include all parents, estimated here at approximately (~) 150,000,000, and since we know the number of prosecutable cases of abuse in the US (~4000 per year), we can calculate the maximum prior odds of prosecutable abuse for any parent in the US. It is $4000/150,000,000$, or .00003 or .003 % probability, or about 1 in 3300. Thus the prior odds of abuse by any parent, is practically zero. This does not mean abuse cannot occur in a future instance of a particular parent, but it does mean that defaulting to abuse and presumptions of abuse are in defiance of near zero probability and are not justified. Low prior odds however, combined with compelling inculpatory indicators (findings) point to abuse. Defining populations and subpopulations, and the impact to base rate or prior odds is exemplified with this example question: What is the base rate (prior odds) of illegal drug use among FBI agents, college students, and gang members? Each would yield very different numbers. In clinical medical use, the flexibility in defining subpopulations can be used constructively to yield reasonable prior odds for general and specific subpopulations. Bayes theorem accommodates more narrowly defined subpopulations, each with different prior odds. As applied to caregivers and abuse, subpopulations can be defined to include groups like apparently loving parents with no prior concerning behaviors, single or divorced parents, previously nonviolent caregivers, previously violent parents, experienced babysitters with no past history of child abuse, teenagers, criminals, drug users, and criteria like poverty, education, age, race, etc.

Any population or subpopulation’s prior odds can be defined if we know the frequency of the event being studied in that population. While we do not know exactly how many lifelong childcare workers who have never abused (a defined subpopulation) will abuse in a year, we can assume it is a low number. If 20 per year abuse unexpectedly out of ~350,000 non-abusing child care workers (Census

data 2019) the prior odds in that group is 20/350,000 and the probability .00006. If it were 1000 per year (which it is not), then 1000/350,000, returns the probability of .003; also a number very close to zero.

On this basis a childcare worker with experience and no prior abuse, would be so rare as to require only the most compelling inculpatory evidence or a pro forma dismissal of any reasonable alternative explanation, to then assume and promote guilt. In spite of these odds, and the absence of compelling evidence, and often a plausible nonabuse explanation, abuse will be alleged and with that, lives ruined, regardless of the downstream legal outcome.

3.5 An “Over Estimation Approach” to Assessing Frequency of Events (i.e., Abuse)

In contrast to well-regarded childcare workers, if poverty is used to define a population of potential abusers, we need to know the number of poor adults, and the frequency of prosecutable abuse in that population of poor people. This is really quite difficult to assess with 100% accuracy. However, with an “over estimation” approach, when we don’t know exactly what the exact frequency of abuse in that population is, we can make an assumption that would overestimate the prior odds of abuse for poor people. The assumption is that all prosecuted abuse cases (~4000 per year in the US) are caused by poverty. If the real number of poverty driven abuse cases is 1000, using 4000 will overestimate the frequency of abuse in poor people. Overestimating then increases the odds of abuse, but if the overestimated odds are low, the actual odds will be lower since we know that all abuse cases are not the result of poverty alone. To do this, we could use these estimates: there are about approximately 250 million people over 18 years old in the US and about 10% are living in poverty (Poverty data 2019). That yields about 25 million impoverished adults/parents/caregivers. Then, the key assumption is to use 4000 (all prosecuted cases) as the number of cases of abuse due to poverty. This, of course, is an overestimate, but there is value in using the maximum possible number to make the prior odds the highest possible when judging if poverty is a reliable indicator, proof, or even an applicable contributor to diagnosing abuse. It also informs, whether using poverty to infer abuse, is valid legal “evidence”, or a false statement being advanced with false certainty.

What we see with the overestimation approach is that the prior odds of poverty contributing to abuse, is very low. The calculation of prior odds of abuse in this “poor” subpopulation using overestimation, is 4000/25,000,000 or .00016 prior odds and probability of abuse well under 1%, .016%, or a frequency of about 1 in 6250 (1/.00016). This overestimate of prior odds for poor people abusing, despite being an intentional overestimation, make that same point as seen with prior odds of abuse by any parent: poverty is associated with infinitesimal prior odds of abusing. The levels are so low that the statement that “poverty is a relevant risk factor for abuse”, has no legitimate evidentiary value in making a determination of abuse. Suggesting that it would be, therefore, is prejudicial and a deception, and when presented as evidence, will mislead almost everyone who hears it.

The key here is that reasonable conservative overestimates of prior odds is adequate to make the point and 100% accuracy is not necessary to demonstrate low probability (low prior odds) over a broad range

of demographic information and subpopulations. Similar miniscule prior odds exist for domestic violence (DV) perpetrators, military personal, pot smokers, teenage parents, prior criminal history, etc. Evaluating DV shows with 10,000,000 incidents a year in the US (National Statistics 2018), the overestimated prior odds for DV being the reason prosecutable abuse has occurred is 4000/10,000,000 or .0004 or .04%. This means that using DV to support abuse will be correct in only 1 out of 2500 times it is invoked as a cause of prosecutable child abuse.

As we shall see with infinitesimal prior odds of abuse, common to many subpopulations (even with the overestimation approach), the reliability of the indicator becomes a key determinant of the probability of abuse. It should be legally accepted that low LRs combined with low prior odds/probability, yield low very posterior odds that clearly define reasonable doubt, if not legal innocence.

Next we can turn to the indicator and its LR, to see how these two variables work together to yield realistic posterior odds of guilt or legal innocence in child abuse cases.

3.6 The Indicator Component of Bayes Theorem

The next step is to factor in the reliability of the any indicator proposed to identify abuse (like SDH, RH, bruises, healing fractures, poverty, a lab test, race, age, prior criminal behavior, etc.) When we have statistics from studies about sensitivity and specificity of an indicator, we can calculate the reliability of an indicator expressed as a ratio of two probabilities; generally, the probability of being right compared to the probability of being wrong.

The indicator's inherent reliability, the Likelihood Ratio (LR) is a ratio, calculated separately and independent of the subpopulation's prior odds of event occurring. In general, the probability of the evidence to lead to correct diagnosis, a true positive (TP), is the sensitivity. The unreliability of a piece of evidence is a variant of specificity and can be expressed as the false positives (FP); how often is the test wrong (suggesting abuse, when it did not occur). The ratio of these numbers is called the "likelihood ratio" (LR).

$$\text{LR} = \text{percent true positive (right)/percent of false positive (wrong)}$$

Indicators (and their LR) can increase or decrease the posterior odds of an event (abuse) occurring, depending on the calculated reliability of the indicator. Unreliable indicators decrease the prior odds (the probability in the defined population) to levels, by any legal or medical standard, which have decreased, or no evidentiary value. Reliable indicators can increase the probability of a low probability event, or end result, having occurred, to thresholds of probability that can be probative by medical or legal standards. For example, a bloody baseball bat with the victim's DNA on it and a depressed skull fracture that matches the shape of the bat, with the perpetrator's fingerprints on the bat in blood and a DNA match, will trump the low prior low odds of abuse in the general population of parents.

In summary, with Bayes Theorem, two probabilities are combined to generate a probability that integrates a prior probability of the event (abuse) occurring, the reliability of the evidence used to diagnose the event (the indicator's LR), to determine the calculated probability (the posterior odds) of that event having actually occurred.

3.7 Examples of LR Calculations and Comparison of Evidentiary Value

With Bayes Theorem, indicators like SDH, RH and bruising can be used separately or in combination to determine the actual probability of the event (i.e., abuse). When present, multiple unreliable indicators, like those mentioned above, are amalgamated, by child abuse pediatricians, to deceptively suggest that abuse is more likely. However, according to the principles of logic, compounding of more than one rare (low probability) event leads to lower prior and posterior odds.

This example from Wood (Wood, 1996) shows LR development for two indicators (in this case two behaviors) and how they can be used in combination with prior odds to determine posterior odds. Furthermore, we can see how the two indicators can be compared to each other to determine which is more reliable.

Here is an actual example of creating a LR related to imitating sexual intercourse as an indicator that abuse occurred from statistics collected in a study by Friedrich, 1993 (Friedrich, 1993). From Wood the following is quoted. "According to reports from parents (Friedrich et al., 1992), this behavior [imitating intercourse] occurs in about 14% of abused children [true positive-TP] and 1% of nonabused children [false positive-FP]." TP and FP define the number of children who do the behavior; some abused (TP) and some not (FP). The ratio of these numbers is 14:1. This is the LR of imitating sexual relations as an indicator of abuse. What it delivers is not the diagnosis of abuse, but the knowledge that the imitation of intercourse is about 14 times more common among abused than nonabused children.

Wood goes on, "The evidentiary 'strength' represented by an LR can be expressed in common English terms... adapted from Goodman and Royall (1988)", where the accepted probabilistic language for the spectrum of LR's, along with their English "translations", are given. See Table 2. This LR, 14:1, represents "moderate to strong evidence".

As a second (contrasting) example, consider the potential indicator of overly aggressive or overly passive behaviors in a child as a sign (indicator) that abuse occurred.

According to parent's reports (Friedrich et al., 1992), about 35% of abused children, but only 10% of nonabused children, are overly aggressive or overly passive. The LR for the indicator (overly aggressive or passive), as a form of evidence of abuse, when calculated, is therefore 35:10, which reduces to 3.5:1, (the correct format for LRs). 3.5 to 1, as shown in Table 2, has been determined forensically to be between "weak" to "weak to moderate evidence".

Once in the LR format, the 2 LRs can be compared to get a sense of which is more reliable. The LR for imitation of intercourse (14:1) is about four times larger ($14/3.5$) than the LR for overly aggressive, overly passive behavior (3.5:1). The comparison indicates that imitation of intercourse is substantially stronger evidence (moderate to strong) of sexual abuse, than is overly aggressive/passive behaviors as an indicator of abuse (weak evidence).

3.8 Posterior Odds and Reliability of an Indicator

Using the percentages above, and a denominator of 100, when prior odds and the relevant LR are multiplied (14+1 of 100 children imitating sex: $15/100 \times$ the LR $14/1$) the posterior odds are $210/100$ or 2.1 to 1, rendering abuse “somewhat unlikely” (see Table 3). When “imitating sex” is used, among all children, as an indicator of abuse it would not qualify as valid evidence. So when the statement “imitating sex is a sign of sexual abuse” is invoked, it must be recognized. Here, “a sign of” is the PL being misused and generally indicates that the subject child imitating sex, will be part of the testimony to support abuse even though it is “somewhat unlikely”. The evidentiary value of the statement is far outweighed by the bias invoked by such poorly defined and often misconstrued probabilistic arguments. Only if the posterior odds/probability reaches accepted medical or legal standards, does it qualify as evidence. If it does not, but is portrayed as such, by misuse of probabilistic language with a false statement like “imitating sex is evidence of abuse”, or other overtly false statements (i.e., short falls never cause serious injury), it is not valid evidence; it is false certainty.

The medically acceptable relationships of posterior odds, to associated probability and applicable probabilistic language are shown in Table 3 (Wood, 1996). Any deviation from these standards in a forensic framework is a form of medical malpractice.

Table 1. Mathematical Probabilities and Forensically Equivalent Narrative (Probabilistic) Terms

Mathematical probability	Mathematical probability
5% or less	Very unlikely
5% to 20%	Unlikely
20% to 40%	Somewhat unlikely
40% to 60%	Undetermined
60% to 80%	Somewhat more likely than not
80% to 95%	Likely
95% or greater	Very likely

Table 2. Likelihood Ratios and Suggested Interpretations

Likelihood Ratio	English Translation
1:1	No evidence
3:1	Weak evidence
5:1	Weak-to-moderate evidence
7:1	Moderate evidence
14:1	Moderate-to-strong evidence
20:1	Strong evidence
55:1	Very strong evidence

Note. Based on Goodman & Royall, 1988.

Table 3. Posterior Odds as Related to Probabilities and Common English Terms

Posterior Odds of Abuse vs. No Abuse	Probabilities That Abuse Has Occurred	Conclusion (in common English)
1:19 or less	5% or less	Very unlikely
1:19 to 1:4	5% to 20%	Unlikely
1:4 to 2:3	20% to 40%	Somewhat unlikely
2:3 to 3:2	40% to 60%	Undetermined
3:2 to 4:1	60% to 80%	Somewhat more likely than not
4:1 to 19:1	80% to 95%	Likely
19:1 or more	95% or greater	Very likely

4. Discussion

4.1 The Unreliability of SDH and RH as Indicators of Alleged Shaking Abuse

Consider this question:

What is the probability of shaking abuse having occurred by a previously nonviolent caregiver as a member of a defined population of nonviolent parents, where the finding of SDH on an imaging study or RH on exam, are being used to diagnose that abuse?

Can the actual reliability of SDH or RH as indicators of shaking abuse be calculated? Perhaps not exactly, but in broader terms we can examine RH and SDH in light of known evidenced based science and gather information sufficient to generate and estimate of reliability and a link to the LR for each finding. In general, any contradictory evidenced based science, as we can see, quickly decreases the LR of these findings and almost immediately they fall solidly in the unreliable range as discussed below. As the LR decreases it generates substantial reasonable doubt.

If RH were shown to occur ONLY with shaking abuse, then the reliability of this indicator is 100%. There was a time in the late 70's and early 80's, when that was thought to be true (Cross, Walsh, Simone, & Jones, 2003). At that time, RH was considered a perfect indicator of abuse (LR = 55:1 or greater...essentially portrayed as always right, never wrong). Every patient with RH was assumed to have been abused by shaking, as in shaken baby syndrome, and almost all (97%) were convicted in that decade when it was present (Cross, Walsh, Simone, & Jones, 2003). This extreme reliance of RH, despite evidence brought forward after 1985 from prior and subsequent studies, that RH and SDH were not 100% reliable, has continued, more or less, unabated. In 2010, Alex Levin MD, for decades, the key historical promoter of the relationship of RH and SBS, published "Retinal Hemorrhage in Abusive Head Trauma" in AAP's official journal Pediatrics (Levin, 2010). The first line is:

“Retinal hemorrhage is a cardinal manifestation of abusive head trauma.”

Here the use of the word “cardinal” is yet another misuse of PL. In reality, a variety of studies dating back 120 years show that RH occurs in many circumstances that are substantially more common than all abuse cases (4000), and render the LR for RH from abusive shaking, based on a high number of false positives, well below 1:1. Some would even argue that RH is not valid evidence of abuse and it has no connection to abusive shaking at all. Challengers (Gabaeff, 2011) would say, based on existing well-constructed studies (discussed below), that RH is a secondary consequence of increased intracranial pressure (ICP) of any etiology (Frasier, Rauth-Farle, Alexander, & Parrish, 2006).

Increased ICP, and in many cases, associated RH, is something that occurs with many medical problems, including SDH itself. Examples include viral brain infection, at around 75,000 cases per year, birth trauma and complications causing SDH affecting about 1,800,000 cases per year, accidents with head injuries that occur in the millions per year, and a wide variety of other conditions that cause increased ICP of any etiology.

Terson (Terson, 1900) was clear that RH occurred in the context of increased ICP from a nonabuse etiology. Walsh (Walsh, 1951) showed ocular hemorrhage with increased ICP with multiple nonabuse etiologies. Muller (Muller, 1974) showed that intraocular and optic nerve sheath hemorrhage occurs in cases of sudden increased ICP with no abuse or shaking involved. Aoki (Aoki, 1984) showed SDH and RH in 100% of adolescents beaten about the head with tatami mats by their hyper-disciplinary parents and no shaking at all. Reddy (Reddy, 2010) showed that unilateral RH, in the context of stroke and increased ICP, can occur, revalidating Terson. Reece (Reece, 2003) showed that 45% of babies at birth have RH from increased ICP occurring during labor. Rooks (Rooks, 2008), separately showed that SDH at birth occurs in 46% of newborns after normal deliveries. It is highly likely, if not certain, that birth trauma causes both (RH and SDH) in the same patients among the ~1.8 million babies that each study separately represents. Both findings are the result of the same spikes in ICP that occur with contractions during labor; obviously with no shaking. Levin himself (Brown, Levin, & Serbaneasu, 2007) showed that small animals, shaken to death, with broken necks, did not get RH even when shaken to death. Binenbaum (Binenbaum, 2007) also published that baby piglets subjected to forces 100 times human capability by a machine, could not produce RH. Lantz, in the last 2 decades, has shown scores of pathology cases with RH from nonabuse etiologies, with the full spectrum of patterns, including 17% of terminal patients at any age have RH at autopsy (Lantz, 2006). Videotaped incidents of severe shaking, multiple times, have not been able to produce RH (Cheryl White, 2008) when followed by immediate exam. With some recent acknowledgement of Terson’s (increased ICP as a cause of RH), it has then been proposed that specific patterns of RH suggest abuse, when that too, as stated by Binenbaum in 2009 (Binenbaum, 2009), is invalid when he wrote:

“Currently there is no uniform grading system for the classification of retinal hemorrhages in abusive head trauma”.

The evidence is now clear, that RH is not a perfect or even reasonable indicator of shaking abuse. RH should be regarded as a secondary response to other biologic processes that cause impulse increases in ICP. Thus, from 1975 on, it was only statements of false certainty about the link of RH and abuse, propped up with misused PL that portrayed RH as diagnostic of abuse (implying very high LR). Evidence that RH is diagnostic of abuse using the false testimony, founded on false certainty, expressed by the misuse of probabilistic language, was often the “nail in the coffin” of a false and unjust conviction of child abuse when none had occurred.

To be reliable testimony (beyond a reasonable doubt), the LR for RH must be greater than 19:1. That would mean that for every 20 cases with RH (of any etiology) 19 would have to be true positives for abuse (abuse is what happened). With the high frequency of RH from many other common sources, and the commonality of RH across the medical spectrum, this is preposterous. Even a small number of other etiologies of RH render it nonspecific, less reliable, or if the LR is sufficiently low, not reliable at all as an indicator of abuse. If 50% of children with RH got RH from other causes than abusive shaking, then RH would have no probative value at all since half of the children with RH would be abused and half would not.

When probabilistic statements like, “RH is highly associated with abuse”, or “RH is diagnostic of abuse”, are dissected, 2 probability questions, emerge: first, what are the prior odds in the subpopulation of which the caregiver is member and second, what is the reliability of RH as an indicator? Beyond increased ICP there are other rare, but commonly identified causes of RH that further weaken RH’s probative value, (if any ever existed). Pre-terminal coagulopathies (DIC) are quite common and a few other rare diseases are often held up as the only alternatives to abuse, another mythology, and once ruled out, abuse is declared the default diagnosis. Default diagnoses should have no role in forensic medicine and legal decisions.

Once the predictability of other causes of RH are identified in the misdiagnosed categories of accidents, birth trauma and complications, brain infections, and other medical problems with increased intracranial pressure, RH as diagnostic abuse, is reduced to an incorrect myth (Matshes, 2010). Consider that even if RH identified abuse 50% of the time correctly (which it does not), and 50% of the time the presence of RH is used to misidentify abuse, the LR is 50%/50% or 1:1. This is defined in forensic epidemiology as “no evidence” (see Table 2), thus, having no evidentiary value. If RH correctly identifies abuse in only 10% of cases and misidentifies abuse 90% of the time, the LR is $10/(90+10)$ or 0.1:1, a number that is well below 1:1, reflecting less than the standard of “no evidence”. For contrast, at the other end of the spectrum think of a situation where an indicator has an LR of 55:1 and when you use it, you make the correct decision 55 times out of 56 cases; you are going to be correct 98% + of the time (55/56). 55:1 translates into “very strong evidence”. White blood cells invading the dura at autopsy, is a very good indicator of meningitis being present in essentially all cases of meningitis and hence is “very strong evidence” of meningitis with an LR in the range of 55:1 or higher.

4.2 PL in the Hospital Environment: Authority Figures, Abuse of Authority, and Legal Endpoints

In emotionally charged legal circumstances (i.e., child abuse suspicions), misused probabilistic language, spoken by an authoritative person, with stature related to title and experience, is, to law enforcement, other physicians, other health professionals, and certainly, lay persons, persuasive, compelling, and convincing. This can occur regardless of the validity of what is being said. In common situations and in court, the judgment of accuracy, or reliability, of various probabilistic statements, ultimately resides with the listener's assessment of the speaker.

In these situations and in general, authority figures are more likely to be believed (Authority & Credibility, 2018). Other specialists in the children's hospital environment from emergency medicine, critical care, radiology, ophthalmology, neurosurgery, forensic pathology, and outside child abuse experts, work in concert with this rhetorical accusatory scheme, and for each, the misuse of similar probabilistic language occurs. Even physicians familiar with the findings being misused to falsely diagnose abuse, normally self-reliant and confident in their own practices, ignore their own experiences, to infer abuse, when they know, or should know, there is reasonable doubt. Yet, such knowledgeable specialists (think of orthopedists testifying that acute postnatal traumatic bone fractures in babies can be considered not painful to allow for an endorsement of abuse), are loathe to challenge the statements the child abuse pediatricians offered to police and CPS. They hold back to avoid conflict, professional criticism and attacks (Greeley, 2011), possible censure (Guardian, 2016), entanglement in the legal system, and/or have an aversion to study the subject independently. In the most favorable light, these other specialists may have been mistrained and/or mesmerized in the parallel pseudoscientific world that is fabricated by senior child abuse pediatricians and endorsed by AAP. Among the darker tones, fear of damage to their own professional persona or any tacit admission of prior misdiagnoses, is anathema and provides a rationale to "stay out of it".

The reality is that any critics of modern child abuse pediatrics, whether in or out of the hospital, are the targets of ad hominin attacks to negate any incoming criticism of child abuse pediatrics (Choudhary et al., Consensus Response 2019). These attacks occur without the child abuse pediatricians addressing the science, the relevant probability analysis, or any acknowledgement of relevant negative investigations of their literature that should alter their behavior and undermines their allegiance to their flawed fund of knowledge (Elinder et al., 2018).

In child abuse pediatrics, "concern" and similar probabilistic terms, poorly defined, or undefined, lead to irreversible actions. The lack of definition however, does not prevent PL from being misused to create, in the first instances of documentation, an illusion of thoughtful diagnostic thinking. In child abuse pediatrics, "concern for abuse" might sound considered and appropriate, but if "concern" or "suspicion" is merely a semantic prelude and/or a stepping stone to declarations of certainty, or false certainty of abuse, that is a different matter. "Concern" by an in-hospital child abuse pediatrician is clearly sufficient for the activation of child protective services and law enforcement. Children are taken and dependency or criminal proceedings are initiated. The statements of false certainty have already

become “evidence” and will be testimony. There are few, if any, examples of child abuse pediatrics moving from concern to no concern, and the closing of a suspected abuse work-up.

On the other hand, a complete and objective analysis of the case to find out the underlying problem and treat it, is often terminated and may not be done at all. When possible, plausible and/or probable alternatives to abuse, supported by existing science, are offered by the defense experts, they will be disregarded and dismissed with words. Soon, in the hospital, and in the offices of law enforcement, it is clear that the abuse narrative is going to trump any, and all, exculpatory information.

4.3 Immunity and Indemnification of False Accusers

What should be of more “concern” however, for all caregivers and people of good will, is that being wrong in child abuse misdiagnosis, carries no legal liability (Gokor & Schlievert, 2020). Being wrong about “concerns” of child abuse, thanks to modern child abuse pediatricians lobbying legislatures, has been deemed legally indemnified (Child Welfare Information Gateway, 2019). With that unrestrained freedom to speculate, believe in, and to accuse caregivers of abuse, concern for abuse, right or wrong, quickly morphs into abuse. The unintended consequence of that inappropriate exemption from liability, which no other specialty in medicine enjoys, is clear. Without liability, without consequences, the cavalier misdiagnosis of abuse, when it occurs, destructive as it is, goes unrestrained. Rush to judgement, intentional disregard of the relevant positive social history (Choudhary et al., 2019), and various forms of confirmation bias, manifest. When belief trumps actual probability, the misuse of PL creates a toxic narrative of accusatory language that streams from the accusatory mind(s), often with psychopathologic overtones. Whether this is done knowingly or unknowingly, by definition, it has disregarded scientific methodology with no consequence to the negligent provider misdiagnosing abuse.

In almost all contested cases of child abuse, when all details are analyzed, the calculated probability of the case being an exception to false certainty, is very high and the probability of abuse very low.

In child abuse practice, however, using known statistics and calculating probabilities of abuse is avoided in favor of belief based thinking. Thus statements like “RH is diagnostic of abuse”, becomes a form of dogma (Matshes, 2010), unsupported by actual probability analysis. With inadequate probability, false certainty then becomes the only effective rhetorical strategy and the primary tool to initiate and win the legal cases after “concern”, as a belief, is expressed. These beliefs are ultimately pitted against valid alternative, evidenced based, nonabuse diagnoses that have a predictable rate of occurrence and have not been meaningfully ruled out. This occurs in spite of the fact that nonabuse possibilities are generally more common. When plausible alternatives exist, exculpatory past social histories are present, and academically qualified defense experts are expressing their own legitimate concerns about abuse, reasonable doubt is substantial. As a consequence, suppression of reasonable doubt with words and invalid criticism of critics remains the principle tactical move to dominate and attempt to control the legal process after accusations of abuse are leveled by child abuse pediatricians.

4.4 The End Result of False Certainty Revealed

The apparent invincibility of the child abuse pediatrician to allege and secure a conviction of child abuse through their own misdiagnosis of abuse, is an immense conflict of interest and obviously improper. Being an uninvited child abuse “specialist,” seeking to scrutinize presenting histories, examine patients, and diagnose abuse, seems more a police function and HIPPA violation (vs health care provider activity). Should parents be given a Miranda warning before talking to a child abuse pediatrician they have never met, who they assume is there to help them, but who is surreptitiously gathering information to use for an accusation of abuse? The answer is yes. Then, these “specialists” take on multiple other self-defined roles: as accuser of targeted caregivers, member of the prosecution team offering ongoing advice and consultation to police and prosecutors to support their own accusations. They function as critics of alternative nonabuse narratives and other experts who raise valid counterpoints. As trainer and also consultant to local social workers, CPS leaders, police and prosecutors, their contorted approach to these cases is baked into the system of investigating, charging, and prosecuting the very caregivers they have duped, from moment one, into talking to them. Their role as key expert witness and the coordinator of the collaboration with other like-minded in-house specialists to promote that their own abuse determinations are correct, is an unprecedented forensic situation. The importance of the child abuse pediatricians in this long sequence would be easy to see if they were removed from these conflicting roles; all the dynamics would shift and there would be no legal case.

In spite of the child abuse pediatrician’s power, at times, prosecutors confronted with strong evidence that substantial doubt is present, will jump off the prosecutorial tract. They rarely dismiss a case, but are willing to get a guilty end result by offering a plea deal. Plea deals can be legitimate and necessary tools. By offering reduced sentences that satisfy the necessities of justice, they can decrease the number of trials that would otherwise overwhelm capacity.

Then there is the “deal that can’t be refused”. After various compelling defense narratives have been aired, authorities can be occultly convinced that proving abuse may be more difficult than normal. In tacit acknowledgement of this and apparently in recognition of outright innocence, combined with a somewhat unethical aversion to dismiss a false accusation case, very serious charges (like homicide) can be reduced to lesser charges. Plea offers with none or very limited jail time, and the promise of the return of children to the parents, as if abuse had not occurred, are made. What once was called murder, or heinous abuse, is now getting an offer at the other extreme. “Pled out and go home’ is the offer. If not, face the threat inherent in the alternative: ‘try to prove your innocent and risk life in prison’. To the beleaguered and innocent caregiver, that deal will only rarely be refused. The defendant, after being falsely accused, now, with no faith in the system, takes the deal. Prosecutor gets the conviction, the child abuse pediatricians are rewarded for their misdiagnosis with yet another guilty verdict that will augment their abuse case count and provide the basis with each conviction for expanding their

programs. Only the caregiver's life is left in ruins. The ruination tracks back to only one person in this tragedy, the child abuse pediatrician, and their statements of false certainty.

5. Conclusion

The application of Bayes theorem in some form is essential to an analysis of the probability of abuse. Consideration of the prior odds of abuse and the unreliability of literally all of the nonspecific and unreliable indicators being used to justify diagnosing abuse, are the necessary components of the diagnostic construct/scheme in use by child abuse pediatrics. In contested abuse cases, without witness evidence, the child abuse pediatrician's expert opinion falls into the zone where the legal and medical criteria to determine reliable evidentiary value are not reached. Only suppositious statements of false certainty, currently only detected by scientists and physicians with the requisite advanced knowledge of these issues can expose the deep flaws in child abuse pediatrics (Lynoe, 2018) and the destructive path on which it walks. In 2006, Leestma (Leestma, 2006) commented correctly regarding analysis of the literature and the issues regarding child abuse.

“Unless the reader is very well informed on the issues and is intimately familiar with the literature, this component of an informed appraisal of an article almost never gets done, and the reader may accept the conclusions uncritically.”

With plausible medical explanations in the differential as well, abuse, without the most solid evidence, should not be regarded as the “most likely” etiology of the findings. It is just the opposite. At times, abuse is not even remotely likely to have occurred. When the mathematical perspective, demonstrated with Bayes Theorem, reveals the prior odds of abuse near zero, likelihood ratios for the most common indicators unreliable, and the posterior odds similarly low, the probabilistic statements used to prosecute cases, have no evidentiary value and are likely to lead to the wrong end result.

Continued use of statements that can be shown to be of marginal validity or false, defines either one of two states of mind which could precede it. One would be a preconceived bias to diagnose abuse in which a scientific methodology intentionally gives way to unsupported beliefs. The second would be an incompetence to recognize the inherent invalidity, lack of a valid scientific foundation, and corrosive impact of the probabilistic statements that child abuse pediatricians have been mistrained to use to advance a false abuse narrative. Neither relies on scientific methodology and both will generate a form of iatrogenic child abuse. The falsehood conveyed via the misuse of PL and the collective negative impact of being wrong, in case after case, is enormous (Hogberg, Eriksson, Hogberg, & Wahlberg, 2020). It is highly likely that among contested cases, based with experience in more than 3500 cases, false accusations cases may outnumber real abuse cases by a wide margin. On balance, we must reckon with the reality that the collective suffering of falsely accused families may dwarf the horrific impacts associated with real abuse.

A false accusation of child abuse is child abuse. It may be the most common form of child abuse in the legal system. Educating the judiciary, police, prosecutors, and social workers of the pitfalls of blind

faith in the flawed subspecialty of child abuse pediatrics, must continue (Gabaeff, 2016). False accusations are a destructive waste of money (Fang, Brown, Florence, & Mercy, 2012), a cause of unnecessary human misery (Hogberg, Eriksson, Hogberg, & Wahlberg, 2020), and societal disruption. The misuse of probabilistic language to advance allegations of abuse, with false certainty statements, must be stopped.

References

- Aoki, N., & Masuzawa, H. (1984). Infantile acute subdural hematoma: Clinical analysis of 26 cases. *J Neurosurg. Aug*, 61(2), 273-280. <https://doi.org/10.3171/jns.1984.61.2.0273>
- Associated Press, Doctors recommend new term for 'shaken baby syndrome', the Boston Globe*. (April 27, 2009). Retrieved from http://www.boston.com/news/nation/articles/2009/04/27/doctors_recommend_new_term_for_shaken_baby_syndrome/
- Atkinson, N., van Rijn, R. R., & Starling, S. P. (2018). Childhood Falls With Occipital Impacts. *Pediatr Emerg Care.*, 34(12), 837-841. <https://doi.org/10.1097/PEC.0000000000001186>
- Authority and credibility*. (2018). Retrieved from <https://www.nngroup.com/articles/authority-principle/>
- Bayes, T. (1764). An Essay Toward Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London*, 53, 370-418. <https://doi.org/10.1098/rstl.1763.0053>
- Binenbaum, G., Forbes, B. J., Raghupathi, R., Judkins, A., Rorke, L., & Margulies, S. S. (2007). An animal model to study retinal hemorrhages in nonimpact brain injury. *J AAPOS*, 11, 84-85. <https://doi.org/10.1016/j.jaapos.2006.11.074>
- Binenbaum, G., Mirza-George, N., Christian, C. W., & Forbes, B. J. (2009). Odds of abuse associated with retinal hemorrhages in children suspected of child abuse. *J AAPOS.*, 13(3), 268-272. <https://doi.org/10.1016/j.jaapos.2009.03.005>
- Brown, S., Levin, A. V., Ramsey, D., & Serbanescu, I. (2007). Natural animal shaking: A model for inflicted neurotrauma in children? *J AAPOS.*, 11-85. <https://doi.org/10.1016/j.jaapos.2006.11.077>
- Census data*. (2018). Retrieved from [https://www.census.gov/topics/families/child-care.html#:~:text=The%202018%20American%20Community%20Survey,and%20service%20workers%2C%20all%20other.&text=The%20majority%20of%20America's%2073.8,with%20mother%20only%20\(23%25\)](https://www.census.gov/topics/families/child-care.html#:~:text=The%202018%20American%20Community%20Survey,and%20service%20workers%2C%20all%20other.&text=The%20majority%20of%20America's%2073.8,with%20mother%20only%20(23%25))
- Chadwick, D. L. et al. (2008). Annual risk of death resulting from short falls among young children: Less than 1 in 1 million. *Pediatrics*, 121, 1213-1224. <https://doi.org/10.1542/peds.2007-2281>
- Chadwick, D. L. et al. (1991). Deaths from Falls in Children: How Far is Fatal? *Journal of Trauma-Injury Infection & Critical Care*, 31(10), 1353-1355. <https://doi.org/10.1097/00005373-199110000-00006>

- Cheryl White video forcefully shaking infant baby 5 times with no injury confirmed. (2008). *Click link and click the image to start*. Retrieved from <https://spaces.hightail.com/receive/83RoW>
- Child Welfare Information Gateway. (2019). *Immunity for reporters of child abuse and neglect*. Washington, DC: U.S. Department of Health and Human Services, Children's Bureau.
- Choudhary A. K. et al. (2019). Consensus statement on abusive head trauma in infants and young children. *Pediatric Radiology*, 49, 424-428. <https://doi.org/10.1007/s00247-019-04341-4>
- Cross, T., Walsh, W., Simone, M., & Jones, L. (2003). Prosecution of child abuse: A meta-analysis of rates of criminal justice decisions. *Trauma Violence Abuse*, 4, 323. <https://doi.org/10.1177/1524838003256561>
- Daubert v. Merrell Dow Pharm., Inc.* (1993). 509 U.S. 579, 589.
- Elinder, G. (Note 1), Eriksson, A. (Note 2, Note 3), Hallberg, B. (Note 4, Note 5), Lynøe, N. (Note 6), Sundgren, P. M. (Note 7, Note 8), Rosén, M. (Note 9), Engström, I. (Note 10), & Erlandsson, B. E. (Note 11) (2018). Traumatic shaking: The role of the triad in medical investigations of suspected traumatic shaking. *Acta Paediatr*, 107(Suppl 472), 3-23. <https://doi.org/10.1111/apa.14473>
- Note 1. Department of Pediatrics, Karolinska Institutet, Stockholm, Sweden.
- Note 2. Department of Community Medicine and Rehabilitation, Forensic Medicine, Umeå University, Umeå, Sweden.
- Note 3. The National Board of Forensic Medicine, Sweden.
- Note 4. Karolinska Institutet, Stockholm, Sweden.
- Note 5. Department of Pediatrics and Neonatology, Karolinska University Hospital, Stockholm, Sweden.
- Note 6. Medical Ethics, Karolinska Institutet, Stockholm, Sweden.
- Note 7. Diagnostic Radiology, Lund University, Lund, Sweden.
- Note 8. Department of Neuroradiology, Skåne University Hospital, Lund, Sweden.
- Note 9. Health Technology Assessment, Karolinska Institutet, Stockholm, Sweden.
- Note 10. Child and Adolescent Psychiatry, University Health Care Research Center, Faculty of Medicine and Health, Örebro University, Örebro, Sweden.
- Note 11. Medical Technology, Royal Institute of Technology, Stockholm, Sweden.
- Fang, X., Brown, D., Florence, L., & Mercy, J. (2012). The Economic Burden of Child Maltreatment in the United States and Implications for Prevention. *Child Abuse Negl.*, 36(2), 156-165. <https://doi.org/10.1016/j.chiabu.2011.10.006>
- Frasier, L., Rauth-Farle, K., Alexander, R., & Parrish, R. (2006). *Abusive head Trauma, a Medical, Legal and Forensic Reference*. Chapter 14 and 15 authored by Sirotnik called "Medical Disorders that Mimic Abusive Head Trauma" and G. W. Medical Publishing.
- Friedrich, W. N. (1993). Sexual victimization and sexual behavior in children: A review of recent literature. *Child Abuse and Neglect*, 17, 59-66. [https://doi.org/10.1016/0145-2134\(93\)90008-S](https://doi.org/10.1016/0145-2134(93)90008-S)
- Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir). (1923).

- Gabaeff, S. C. (2011). Challenging the Pathophysiologic Connection between Subdural Hematoma, Retinal Hemorrhage and Shaken Baby Syndrome. *West J Emerg Med.*, 12(2), 144-158.
- Gabaeff, S. C. (2016). Exploring the controversy in child abuse pediatrics and false accusations of abuse. *Legal Medicine*, 18, 90-97. <https://doi.org/10.1016/j.legalmed.2015.12.004>
- Gokor v Schlievert. (2020). *In the United States District Court for The Northern District Of Ohio Western Division*. Case: 3:16-cv-03038-JGC Doc #: 79 Filed: 07/23/20 1 of 14. PageID #: 1630.
- Goodman, S. N., & Royal, R. (1988). Mrs. Evidence and Scientific Research. *Am J Public Health*, 78, 1568-1574. <https://doi.org/10.2105/AJPH.78.12.1568>
- Greeley, C. S. (2011). A Wolf in Evidence Clothing: Denialism in Child Abuse Pediatrics. *AAP Grand Rounds*, 26, 24. <https://doi.org/10.1542/gr.26-2-24>
- Guardian. (2006). *Doctor Wins Appeal Over Shaken Baby Syndrome Trials Evidence. Waney Squier, Who Was Barred In March From Practising Over Evidence She Gave In Criminal Trials, Cleared of Dishonesty*.11-4-16.
- Hogberg, U., Eriksson, G., Hogberg, G., & Wahlberg, Å. (2020). Parents' experiences of seeking health care and encountering allegations of shaken baby syndrome: A qualitative study. *PLoS ONE*, 15(2), e0228911. <https://doi.org/10.1371/journal.pone.0228911>
- Lantz, P. E., & Stanton, C. A. (2006). *Postmortem detection and evaluation of retinal hemorrhages*. Am Acad Forens Sci (Abstract, presented at the AAFS Annual meeting, Seattle, WA, February).
- Leestma, J. E. (2006). SBS: Do confessions by alleged perpetrators validate the concept? *J Am Phys Surg*.
- Levin, A. (2010). Retinal Hemorrhage in Abusive Head Trauma. *Pediatrics*, 126, 961-970. <https://doi.org/10.1542/peds.2010-1220>
- Lynoe, N. et al. (2016). *The role of the triad in medical investigations of suspected traumatic shaking, a systematic review*. Swedish Agency for Health Technology Assessment and Assessment of Social Services. Traumatic shaking. SBU assessment. Report 255e.
- Matshes, E. (2010). *Retinal and Optic Nerve Sheath Hemorrhages Are Not Pathognomonic of Abusive Head Injury*. Proceeding of the American Academy of Forensic Sciences, Feb 24, Seattle, WA.
- Muller, P. J., & Deck, J. H. (1974). Intraocular and optic nerve sheath hemorrhage in cases of sudden intracranial hypertension. *J Neurosurg.*, 41(2), 160-166. <https://doi.org/10.3171/jns.1974.41.2.0160>
- National Statistics. (2018). *NCADV. National Coalition Against Domestic Violence*. Retrieved (5 October 2018).
- Poverty data. (2019). Retrieved from <https://www.usatoday.com/story/money/2019/11/06/united-states-poverty-rate-for-every-group/40546247/>

- Reddy, A. R., Clarke, M., & Long, V. W. (2010). Unilateral retinal hemorrhages with subarachnoid hemorrhage in a 5-week-old infant: Is this nonaccidental injury? *Eur J Ophthalmol.*, 20(4), 799-801. <https://doi.org/10.1177/112067211002000427>
- Reece, R. M., & Nicholson, C. E. (2003). Inflicted childhood neurotrauma. In R. M. Reece, & C. E. Nicholson (Eds.), *Proceedings of a conference sponsored by HHS, NIH, NICHD, ORD, NCMRR* (02 Oct 10). American Academy of Pediatrics.
- Rooks, V. J., Eaton, J. P., Ruess, L., Petermann, G. W., Keck-Wherley, J., & Pedersen, R. C. (2008). Prevalence and evolution of intracranial hemorrhage in asymptomatic term infants. *Am J Neuroradiol.*, 29(6), 1082-1089. <https://doi.org/10.3174/ajnr.A1004>
- Terson, P. D. A. (1900). Hemorrhage in the vitreous body during cerebral hemorrhage. *La Clinique Ophthalmologique*, 22, 309-312.
- Walsh, F. B., & Hedges, T. R. Jr. (1951). Optic nerve sheath hemorrhage. *Am J Ophthalmol*, 34(4), 509-527. [https://doi.org/10.1016/0002-9394\(51\)90294-2](https://doi.org/10.1016/0002-9394(51)90294-2)
- Woods, J. M. (1996). Weighing Evidence in Sexual Abuse Evaluations: An Introduction to Bayes's Theorem. *Child Maltreat*, 1, 25. <https://doi.org/10.1177/1077559596001001004>