The Eighth Circle of Fire

ELAINE WHITFIELD SHARP, TLC '98

Junk science lurks in child abuse cases of every variety, whether they are alleged sexual molestation, or other physical and mental abuse cases. Parts I and II of this series covered the scientific flaws of a specific area of child abuse: so-called "shaken baby syndrome" (SBS) cases, and the scientific flaws in short falls cases.[1] This article[2] will cover some of the ways to characterize and challenge those flaws in a pretrial Daubert or Frye motion using the rules of evidence governing the admissibility of scientific evidence. (If you do not have a SBS case, you may still find the section on scientific evidence rules helpful because it applies not only to cases of alleged SBS, but to all scientific evidentiary issues in all civil or criminal cases.)

To the extent that the judge agrees to exclude testimony that is not based on reliable science, testimony that is not based on the correct application of reliable science to a case, or testimony proposed by an 'expert' who is not qualified to give it, a pretrial motion serves to: (1) possibly minimize the use of some of, if not all, unfairly prejudicial testimony before a jury; (2) educate the trial judge that there are issues of junk science that are more unfairly prejudicial than probative; (3) flush out a refinement of the State's theories on causation and timing of injury; (4) create an opportunity to discuss a possible plea by educating the State about the weaknesses in its scientific case; and/or (5) build a record for appeal on the junk science issues.

As a matter of strategy, if you do not want to make a pretrial motion challenging the science, consider using the information in this article to help construct cross examination for the purposes of moving for a directed verdict and/or motion to set aside the verdict in cases where the State relies predominantly on medical testimony to establish causation and timing.

Dante Alighieri with his Poems (detail), by Domenico di Michelino, c 1460. Dante placed the junk scientists of his time—alchemists and other deceivers of men—in the eighth circle of fire in Hell.

Child abuse is as ugly as it gets. The bruises on the swollen faces and battered bodies of dead tots on the autopsy table, the broken bones, the blood, the burns from irons, cigarettes and scalding water, might make you want to kill the culprit. It might even be enough to make you believe the Devil actually exists. But the expert witness who brings junk science to the courtroom to assist in the wrongful convictions of innocent men or women is also enough to make you believe in the 'Father of Lies.'

Real child abusers and pseudo-science experts indulge in the same vice: they abuse the power and trust we vest in them. In "The Inferno," Italian poet and philosopher Dante Alighieri placed the pseudo-science deceivers of his day—diviners, magicians and alchemists—in the eighth circle of fire in Hell.

Like an accusation of witchcraft, a publicized child abuse accusation is tantamount in the minds of many prospective jurors to a conviction. Many view a trial as just a Constitutional nicety before the inevitable conviction. Here are a few choice comments from prospective jurors during voir dire in a so-called SBS case in which all but one of the 98 people on the venire had read negative pretrial press reports about the case:

"I am biased against child abusers."

"I feel she should be put away."

"I've followed this from the start. She probably is guilty."

"I've got three kids of my own and when anger gets out of hand, things happen."

"My wife quit work because of this case."

And, then there was this classic example of open-mindedness:

Juror: "She's guilty."

Judge: "Would evidence to the contrary make any difference to you?"

Juror: "*If* I listened, probably not."

Against this hostile backdrop, it is crucial to try to rewrite some of the established, predictable scenes by identifying unreliable science promoted by the State through expert testimony about causation and timing of injury in SBS cases.

Deep in the Mouth of the Wolf

This area of practice is still as dark as the inside of a wolf's mouth, but paradigm shifts in law and medicine now empower us to try to bring in some light. In law, under Daubert v Merrell Dow Pharmaceuticals[3] and its progeny, the responsibility and power to determine the reliability of scientific evidence has been placed squarely in the trial judge's hands.

In medicine, physicians are now expected to ensure that their diagnostic and treatment decisions are based on reliable science, that is, on evidence-based medicine (EBM) and not merely upon anecdotes and case reports. Evidence-based medicine involves the use of well-designed studies of large numbers of people to better guide treatment decisions.

Both paradigm shifts are shifts in emphasis, from potentially unreliable to reliable information, representing an evolution in

our attempt to increase trust in the system of justice and the practice of medicine. When we are responsible for the lives of others in either the legal or medical arenas, the practice must be as reliable as we can humanly make it. That is what is fair. That is what is just.

SCIENCE GETS ITS DAY IN COURT

The growth of science and technology leading up to and beyond the Industrial Revolution has increasingly landed scientific issues in court at an exponential rate. Especially in the past 150 years, the courtroom has become the proving ground for many new areas of science in order to determine criminal and civil liability.

Over the last two centuries trial judges have been asked to decide: Are there unique human indentifiers, such as fingerprints? If so, are fingerprinting techniques reliable? Is the identification of a person through facial measurement (anthropometry) reliable? What about bite marks? Lip prints? Voice recognition? Handwriting? Ear prints? Is DNA typing reliable? Is there a machine that can tell if someone is lying or telling the truth? Is there a machine that enables us to see a person's bones right through their skin, their brain, or injury to their brain? Is there an objective way to tell how much pain a person suffered before death? How do we know if this bullet was discharged from this gun?

Junk science can seem like a laughing matter—until it is used against another human being to take their liberty or life. Copyright, 2004 by Sidney Harris. Reprinted with permission.

Experts became a permanent fixture, and often as equally powerful, as the trial judge's gavel. Despite the widespread use of experts on every subject from Antibodies to Zoology, most state and federal trial court judges had few standards, if any, by which to independently judge the question: "Is this 'science' reliable enough for the trier of fact to consider in helping to resolve this case justly?" It is a maxim of our law that anything that is not reliable is not relevant and, as such, may be more unfairly prejudicial than probative on issues that impact a finding of guilt or innocence. A lack of standards left judges adrift in uncertainty about what to admit and what to exclude.

In 1923, a federal appeals court announced a working rule of thumb, for which we have James Alphonzo Frye[4] to thank.

Frye was on trial for murder and, to prove his innocence, or at least to raise reasonable doubt, he took a systolic blood pressure deception test, a crude precursor of the polygraph examination. His defense lawyer offered the results of the test at trial and, when the trial judge sided with the government and refused to let the jurors hear the results, Frye's attorney offered to bring the 'scientist' who ran the test to run it again on the defendant right in front of the jurors. Again, the trial judge ruled for the prosecutor and the jurors never heard that 'scientific' evidence.

Frye was convicted of second degree murder and, in a single-issue appeal, he raised the question of whether, in excluding the evidence of the test, the trial judge abused his discretion, creating reversible error that entitled him to a new trial.

Frye's lawyer claimed that the 1923 "deception test" measured systolic blood pressure, the body's strongest blood pressure and that, by measuring changes in blood pressure triggered by a witness' changing emotional state, the test was able to detect truth or falsehood. Describing the underlying 'scientific' theory of the test, the Frye Court wrote:

> Scientific experiments, it is claimed, have demonstrated that fear, rage and pain always produce a rise in systolic blood pressure and that conscious deception or falsehood, concealment of facts, or guilt of crime, accompanied by the fear of detection when the person is under examination, raises the systolic blood pressure curve, which corresponds exactly to the struggle going on in the subject's mind, between fear, as the examination touches the vital points in respect of which he is attempting to deceive the examiner. In other words, the theory seems to be that truth is spontaneous, and comes without conscious effort, while the utterance of a falsehood requires a conscious effort, which is reflected in the blood pressure.

Although there were no guiding standards about what scientific evidence should be admitted and excluded, a practice in some courts in 1923 was to allow an expert to testify about scientific or technical knowledge when, in the judge's discretion, the facts

needed the interpretation or opinion of an expert to assist the jurors in understanding and deciding important issues in the case.

Confusion occurred when a party, like Frye, wanted jurors to hear scientific testimony about novel or cutting-edge "science." How was the trial judge to determine whether the "thing" that was being proffered should be heard or considered by the trier of fact?

The Frye Court held that the standard was in the hands of the scientists, writing:

> Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained *general acceptance in the particular field in which it belongs.*" Frye, 1014. (Emphasis added.)

The systolic blood pressure test had not gained the kind of standing and recognition among physiologists and psychologists—the relevant scientific community—to warrant presenting the results of the tests to jurors, the appeals court explained. Holding that the trial judge had not abused his discretion, the appeals court affirmed Frye's murder conviction.[5]

Thus was born the test in the federal courts by which any evidence that a litigant claimed was novel "science" was to be judged: If the scientists said it was generally accepted in the relevant scientific community, the trier of fact could consider it to resolve factual issues in criminal and civil cases. While the Frye Court only addressed the question of novel scientific evidence, later courts extended the "general acceptance" test of admissibility to all scientific evidence.[6] For 70 years, the "Frye test," as it became known, was the standard for judging all science in the federal courts. Most state courts adopted the Frye test, although some adopted hybrids of the test that imposed a duty on the trial judge to independently assess the reliability of proffered scientific testimony.[7]

In an ideal world, if a scientific proposition has been generally accepted in the relevant community, one would expect it to be based on a well-reasoned hypothesis shown to be valid by reliably designed research consisting of sufficiently reliable data, and subsequent tests validating initial results. And, one would expect that any expert testimony about the proposition in a specific case would be the product of reliable methods and conclusions that would be properly applied to the facts of the case. Frye's admissibility test of "general acceptance in the relevant scientific community" would have implicitly contemplated all of these expectations because reliability in hypothesis formation, study or experiment design, testing and interpretation of results, have historically been the hallmarks of trustworthy science.

But that is an ideal world. In the real world, scientists are plagued by problems of ego, turf wars, the politics of scientific funding, faulty hypothesis formation, poor study design, poor data analysis, and otherwise intellectually corrupt orientations and conclusions. Like any other human endeavor, science is far from pure, neutral and unbiased.

Nevertheless, for more than 70 years, the Frye test held the world of law hostage to the idiosyncrasies of the world of science.

But three relatively recent cases are designed to change all that. The United States Supreme Court laid out the rules for the admissibility of scientific evidence in Daubert, Kumho and Joiner.

JASON DAUBERT AND THE SEARCH FOR RELIABILITY

Jason Daubert and Eric Schuller were born with serious birth defects. Because the boys' mothers had taken the anti-nausea drug "Bendectin" during pregnancy, they blamed their sons' birth defects on the drug. The boys and their parents sued Merrell Dow, the pharmaceutical company that made the drug. The company prevailed in a motion for summary judgment when the trial judge found that the boys had no generally-accepted evidence under Frye to take to a jury about what caused their birth defects. Juries resolve factual disputes and here there was no dispute about causation because the plaintiffs had no evidence to create one, the court held.

Lawyers for Merrell Dow argued that all the human statistical studies about Bendectin showed the drug did not cause birth defects. Lawyers for Jason and Eric argued that their evidence showed that Bendectin did cause the boys' birth defects. They offered (1) "in vitro" (test tube) and "in vivo" (live animal) studies that found a link between Bendectin and malformations; (2) pharmacological studies that showed chemical structures similar to the chemical structure of Bendectin that they claimed caused birth defects; and (3) a "re-analysis" of previously published human statistical—i.e., epidemiological—studies. (The third type of proof is sometimes called "meta-analysis" or "data pooling.")

The trial court judge dismissed the case because, under the Frye test, "scientific evidence is admissible only if the principle upon which it is based is 'sufficiently established to have general acceptance in the field to which it belongs.'"

Jason and Eric had no way to challenge the accuracy of Merrell Dow's human statistical studies, the trial court found. The plaintiffs' re-analysis or recalculation of those studies had not been available for review by the relevant scientific community in order to be generally accepted. Therefore, the trial court found, Jason's and Eric's animal-cell (test tube) studies, live-animal studies, and the chemical structure analyses, did not establish, and could not be admitted to show, that Bendectin caused birth defects.

Jason and Eric appealed to the United States Court of Appeals for the Ninth Circuit. That court agreed with the trial judge: Without *generally-accepted* evidence of causation, an element of the plaintiff's claim, the trial judge was correct to dismiss their lawsuit.

Jason and Eric appealed to the U.S. Supreme Court to review their case on the question of the correct standard to apply in

determining what is reliable "science." By this time, some of the 13 federal circuit courts of appeal were in conflict over how to answer this question. Should they use the "general acceptance" Frye test, or the rule of evidence enacted by Congress in 1975? The 1975 version of the rule stated:

> If ***scientific***, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise." Federal Rule of Evidence 702. ("Rule 702.") (Emphasis added.)

How was this rule, enacted by Congress 52 years after Frye, to be applied? Was Frye's general-acceptance test part of it? Or was the rule independent of that? If Rule 702 was independent of Frye, how were trial judges to interpret and apply the rule?

Daubert presented not only an unusual conflict of evidence laws and a bitter battle of scientific experts, but also a classic conflict of decisions among the 13 federal circuit courts of appeal.

The U.S. Supreme Court took the case.

In a nutshell, the Supreme Court held that scientific evidence did not have to pass Frye's general acceptance test as a precondition for admissibility, that Rule 702 was the pertinent rule to apply, but that an interpretation of the rule requires trial judges to act as reliability gatekeepers to ensure that an expert's testimony rests both on reliable methods and conclusions and that it is actually relevant to the specific case. How the Supreme Court got there is another story, and an important one to understand in order to mount a pretrial attack on evidence that the State claims is "science," or in order to choose one's own scientific evidence wisely to meet the standards for admissibility.

The main question for the Court was, when Rule 702 speaks of "scientific evidence" what does that term mean and by what criteria should a trial court judge the "science" in deciding the question of admissibility?

Educated with help from 22 *amici* briefs filed by several of the nation's scientific leaders,[8] the Supreme Court cobbled together a definition of what it believes the term "scientific knowledge" means in Rule 702. "Scientific knowledge," said the Court:

- Implies a grounding in the methods and procedures of science;
- Implies a body of known facts, accepted on good grounds;
- Implies that an inference or assertion is derived by the scientific method; and
- Does not imply that the subject of scientific testimony must be "known to a certainty" for, arguably, there are no certainties in science. ("Science is not an encyclopedic body of knowledge about the universe. Instead it represents a process for proposing and refining theoretical explanations about the world that are subject to further testing and refinement.") Daubert, 590-591.

Under Frye, the scientists told us what was reliable and what was not. But, under Daubert, the Court told the trial judges that the determination of reliability was their duty. They, and they alone, were to be the gatekeepers of scientific truth at the bar. If the trial judge determines that proffered scientific testimony is reliable, the gatekeeper should admit it for the trier of fact to consider. Otherwise, the gatekeeper must exclude it. Federal Rule of Evidence 104(a) states in part: "Preliminary questions concerning the qualification of a person to be a witness . . . or the admissibility of evidence *shall* be determined by the court…." [Emphasis added] This was a mandate, not a choice.

The Court explained that, in acting as gatekeepers of scientific knowledge, there are some hallmarks of reliable science for which to look:

1. Is the "thing" that is being proffered as scientific capable of being tested, and has it been tested? Quoting 19th Century philosopher Karl Popper,[9] the Court wrote: "Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry.";
2. Has the theory or technique been subjected to peer review and publication? (Publication is not the *sine qua non* of admissibility and, indeed, is only one component of peer review. Sometimes what is published is not reliable, and sometimes that which is reliable is not published. But, publication ensures "submission to the scrutiny of the scientific community," and is a "component of good science, in part, because it increases the likelihood that substantive flaws in methodology will be detected.");
3. Is there a known or potential rate of error? (Do standards exist and are these maintained in the testing? This is important in relation to the validation of test results through the use of consistent standards and is part of good science.); and
4. Is the "thing" generally accepted? (Depending on the case and type of proffered scientific testimony, it *may* be relevant to identify a relevant scientific community and to determine the degree of acceptance within that community.) Daubert, 593-4.

The Daubert Court emphasized that the trial judge's "focus, of course, must be solely on principles and the methodology, not on the conclusions they generate."

Determining whether the proffered science is reliable is only one part of the gatekeeper's job, said the Court. The trial judge must also ensure that the scientific testimony is, in fact, relevant to resolve a disputed issue in the case. There must be a "fit" between the science proffered and the facts of the case. Wrote the Court: "Rule 702's 'helpfulness' standard requires a valid scientific connection to the pertinent inquiry as a precondition of admissibility." For example, in a case in which a child is found at autopsy to have died solely as a result of blunt impact to the head, evidence that the child died as a result of violent shaking is not scientifically relevant.

Under Daubert, reliability and relevancy (fit) are the two guiding principles in the determination of whether to admit scientific evidence.

Daubert's "check list" of four factors to consider in determining if proffered scientific testimony should be admitted contemplated the specific type of science offered in that case. On the plaintiff's side were meta-analyses, in vivo and in vitro studies while, on the defendant's side were epidemiological studies, that is, the study of large groups of people from which statistical correlations may be shown. Daubert's specific fact situation involved relatively esoteric areas of science.

Daubert did not answer the questions of whether, and how, it applied to other areas of expert knowledge, such as, engineering and other applied sciences and technology. And, it was not clear by what standard appeals courts were to review a trial judge's decision to admit or exclude scientific evidence. Was it a mere "abuse of discretion" standard under which reliance by the trial judge on any facts to admit or exclude expert testimony would be upheld? If so, that would mean that the decision of the trial judge, who already had the power to act as reliability gatekeeper, would be virtually untouchable on appeal. Or was the standard a more stringent one in which the appellate courts would completely review the trial judge's evidentiary ruling to see whether they disagreed with the lower courts' findings?

The Court answered these questions in Joiner and Kumho.

JOINER AND THE NOT-SO, SAY-SO OF THE EXPERT

Joiner claimed that while working as an electrician for General Electric, he developed small-cell lung cancer because of his exposure to polychlorinated biphenyls (PCB's) and their derivatives, furans and dioxins, found in the coolant fluid in transformers.

Joiner's experts on causation relied on studies performed on infant mice that developed tumors in their small air sacs after highly-concentrated, massive doses of PCB's were injected directly into their stomachs and abdominal walls. In contrast, Joiner's human exposure was indirect and on a much lower scale. Joiner's experts also relied on two studies, the authors of which, themselves, were unwilling to suggest a link between PCB's and lung cancer. They also relied on a third study in which a link between lung cancer and a specific mineral oil—to which Joiner had not been exposed—was found. The trial judge ruled that this testimony was inadmissible because it did not show that Joiner's small-cell lung cancer was caused by his exposure to PCB's and his experts' testimony to the contrary, and their insistence that causation was shown, did not rise above subjective belief or unsupported speculation.

The Court of Appeals for the Eleventh Circuit disagreed. Rather than simply deciding whether the trial court had abused his discretion in coming to a manifestly erroneous factual conclusion, the Eleventh Circuit applied a stringent review. It held that, in light of the fact that the rules of evidence display a preference for admissibility, the trial judge had incorrectly excluded the plaintiff's proof of causation. The trial judge had incorrectly played "science" judge by reaching a different conclusion about the research than the plaintiff's experts reached. The jury, not the judge, should decide between competing views of science, the Eleventh Circuit said.

General Electric appealed to the U.S. Supreme Court, arguing that the standard of review on appeal when a trial judge excludes scientific evidence is whether they abused their discretion, that is, whether the decision was manifestly erroneous in that it lacked any reasonable, factual foundation. Looking at the trial judge's findings in that light would mean that it was not manifestly erroneous for him to exclude the plaintiff's proof of causation because there was a factual basis for the finding that this failed to make the link between Joiner's exposure to PCB's and his cancer.

The U.S. Supreme Court took the case and ruled that the appeals courts were to apply the abuse of discretion standard to review a trial judge's rulings to admit or exclude scientific evidence. Unless there was no factual basis for the trial court's decision, it was to be left untouched. This development meant that not only was the power to determine scientific reliability in the hands of the trial judge, but the trial judge's ruling was to be pretty much impregnable to attack on appeal.

The other question in Joiner was whether a trial judge had to take the word of an expert that scientific evidence was reliable. The Court held that the trial court does not have to rely on the *ipse dixit*—the say-so or bare assertion—of an expert exerting his authority as such. Rather, the trial judge must be the reliability gatekeeper, scrutinizing not only conclusions, but also the methods used by experts in reaching those. Affirming the trial judge's exclusion of the plaintiff's proofs, the Court reminded trial judges who holds the reins of reliability and relevancy:

> "[C]onclusions and methodology are not entirely distinct from one another. Trained experts commonly extrapolate from existing data. But nothing in either Daubert of the Federal Rules of Evidence requires a [trial] court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert. A court may conclude that there is simply too great an analytical gap between the data and the opinion offered." Joiner, 146.

Both Daubert and Joiner involved "scientific knowledge." The question still nagging the federal trial bench was whether the trial judges had to be the reliability gatekeepers in all areas of knowledge, such as skills-oriented, applied sciences like engineering. Did the duty extend to testimony based on "technical" or "other specialized" knowledge, as included in Rule 702?

KUMHO AND THE TIRED-TIRE

In Kumho, the U.S. Supreme Court answered the question. In that case, one plaintiff had been killed and others injured when a tire of the minivan they were traveling in blew out. The plaintiffs' tire-failure-analysis expert inspected the tire and opined that the blow out was consistent with a tire-manufacturing defect and not due to wear and tear. But, he also conceded that the tire was old and worn and that it had twice previously been punctured and inadequately repaired.

Applying all four of the Daubert factors, the Kumho trial judge found that the testimony of the tire-expert witness on the cause of the blow out was not reliable, and dismissed the suit on grounds that the plaintiffs could not prove the element of causation. The plaintiffs moved for reconsideration on the grounds

that the trial judge too rigidly applied the Daubert factors. They argued that these factors were to be used for areas of science akin to that in Daubert and that the real focus was not on the four factors but on reliability. The Kumho trial judge reconsidered using reliability as the standard, after which he wrote that "the component of [the expert's] tire failure analysis which most concerned the Court [was] the methodology employed by the expert in analyzing the data obtained in the visual inspection, and the scientific basis, if any, for such an analysis." He did not buy that the plaintiffs' expert's tactile or hands-on inspection of the tire was reliable.

On appeal, the Eleventh Circuit Court of Appeals held in the plaintiffs' favor on grounds that the Daubert factors only apply to "scientific knowledge" and not to "technical" or other areas of "specialized knowledge" in Rule 702. Kumho Tire appealed to the U.S. Supreme Court.

The Supreme Court took the case "in light of the uncertainty" about whether Daubert applies to all areas of knowledge listed in Rule 702, i.e., science, technology and other specializations. The Court held it did. And, as for Daubert's four factors, the Court wrote:

> "[I]n our view…we can neither rule out or rule in, for all cases and for all time the applicability of the factors mentioned in Daubert, nor can we now do so for subsets of cases categorized by category of expert or by kind of evidence. Too much depends upon the circumstances of the particular case at issue."

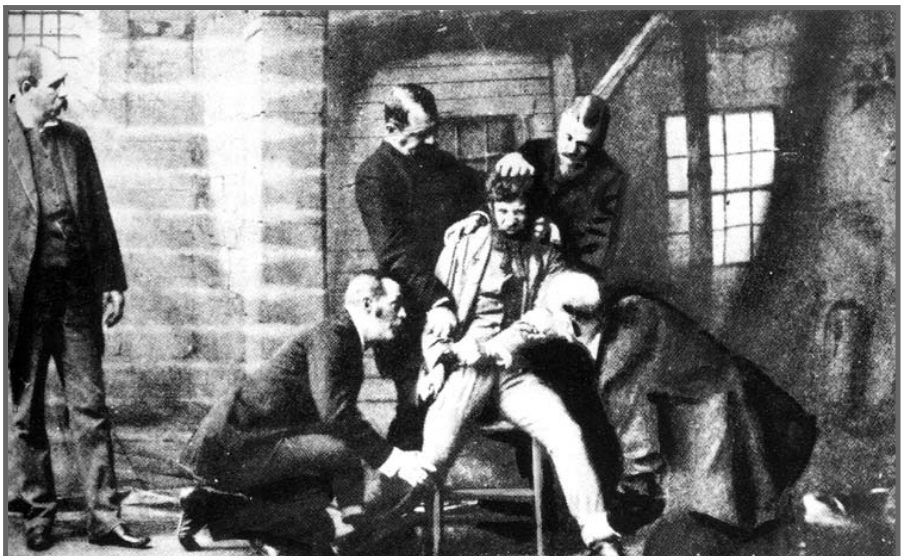
* * * * *

> [That] list was meant to be helpful, not definitive. Indeed, those factors do not all necessarily apply even in every instance in which the reliability of scientific testimony is challenged. It might not be surprising in a particular case, for example, that a claim made by a scientific witness has never been the subject of peer review, for the particular application at issue may never previously have interested any scientist. Nor, on the other hand, does the presence of Daubert's general acceptance factor help show that an expert's testimony is reliable where the discipline itself lacks reliability…" Kumho, 150-1.

The Kumho Court also warned against litigation bias in tests. Expert witnesses must use the same degree of intellectual rigor in testing for a court case as they would in the laboratory, or any other area of their practice. Methods and conclusions had to be reliable in all contexts.

Here's the sum and the substance of expert witness testimony:

- Daubert: The trial judge is the mandated gatekeeper of scientific reliability and relevancy. "Scientific knowledge" has certain hallmarks that make it reliable. All scientific expert testimony must be relevant to resolve an issue in the case (fit).
- Joiner: A trial judge's decision to admit or exclude scientific expert testimony will only be upset by the appeals court if it is an abuse of discretion, i.e., without facts to support it. The trial judge will usually have the last word. The trial judge does not have to take the word, i.e., the *ipse dixit*, of the expert that a conclusion is correct if the methodology is



Especially in the past 150 years, courts have become the proving ground for many new areas of science used to determine civil and criminal liability. They have answered questions about whether there are unique identifiers such as anthropometrics (unique facial measurements) fingerprints, and DNA. In the construction of reliable identification systems, "arrested men often put up such resistance to photography that they had to be held by force [before] the camera." Thorwald, J., "The Century of the Detective."

unreliable (fit).

- Kumho: Reliability analysis applies to all areas of expert testimony—scientific, technical or other specialized knowledge—and also to methods used, and conclusions reached, in testing for litigation.

Daubert, Joiner and Kumho together represent the search for reliability as part of the quest to do justice, and are included in the now-modified Rule 702:

> "If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case."

In evaluating whether scientific or other expert testimony is admissible, most states follow some version of the principles found in either Rule 702, or a mix of the principles in Rule 702 and Frye. None of the cases in the Daubert trilogy or in any state on the admissibility of scientific evidence are shrouded in mystery because two unifying threads are woven throughout their fabric: reliability and relevancy. No matter what the name of the case or the number of the rule of evidence governing the admissibility of scientific evidence in your state, without exception, they can all be boiled down to reliability and relevancy (fit).

When science or other areas of expert knowledge are in the courtroom, reliability is, on a fundamental level, about trying to find the truth. It is about proof beyond a reasonable doubt, clear and convincing evidence or proof by a preponderance of the evidence. It is about being able to trust a result.

Some Ideas to Use Where Frye is Not Retired

In states where the Frye test is still the standard, developments in science may allow testimony to be challenged that was previously generally accepted. For example, our evolving understanding of some of the pitfalls in fingerprint collection and analysis led one court to exclude evidence of latent prints as "unreliable."[10]

There have been many advancements in the field of so-called shaken-baby-syndrome showing that the theory of causation (pure shaking) is not reliable in many cases. In some cases, such as pure impact cases, evidence of shaking is not even relevant. Showing that the science, as it once was accepted, is no longer generally accepted because it is no longer deemed reliable, is an avenue open to those in Frye states.[11]

One can also argue that biomechanicians—who are experts in head injury causation—comprise the relevant scientific community, and not physicians. In most states, including those that use the Frye test, expert testimony about shaking causing subdural hematomas (SDH's) and retinal hemorrhages (RH's) was admitted before the science of traumatic brain injury causation was robust. The current situation is that the science of head injury causation is not taught in medical schools and is not understood by prediatricians and other physicians who treat children with traumatic brain injury. The Daubert Court made it clear that the notion of general acceptance is an evolving one. Science changes and advances. This is so in the area of pediatric head injury. Therefore, one could challenge old beliefs about general acceptance and shaken baby syndrome using the argument that the model of shaking as a cause of traumatic brain injury (TBI) in babies has never been validated by biomechanicians who now comprise the relevant scientific community. The main points to focus on are those of reliability and relevancy (fit).

Mounting the Challenge to Junk Science in SBS Cases

There are several points of attack in so-called SBS and impact cases. First, one may challenge the mechanism or causation of injury and the timing of injury. Secondly, one may challenge the application of the science to the case and the expert's qualifications.

There are a number of ways to structure a challenge, but here are two suggestions. First, one can mount a general attack by summarizing all the faulty assumptions of causation and timing in the medical literature, and filing a Daubert or Frye pretrial brief forcing the State to respond with it's own list of articles, to which one can then reply. Alternatively, or in addition to this, one can mount a specific attack by requesting all the articles upon which the State's experts' depend for their opinions about causation and timing. Here are some of the building blocks for this type or phase of the reliability challenge.

The Building Blocks for Challenging Reliability

In the area of SBS literature and opinion, there is more that is unreliable than that which is reliable. Here are a few suggestions for gathering the information you need in order to tell the judge the story of scientific truth in your specific case.

Using state rules of criminal discovery or by motion, request the following:

1. Any and all articles upon which the State's experts, including treaters who will testify, rely for their opinion on causation and timing. (Use the language of the indictment, grand jury testimony, and statements by the prosecutor in open court about what the experts are expected to say.); and
2. Names of cases, names of courts and docket numbers of cases in which the State's experts have previously testified about causation and/or timing of pediatric head injury; and
3. List of all articles and publications experts have written as first author, second author, co-author, contributor, etc.

As with any meaningful discovery, prepare yourself for an ordeal. The State, seeing you're coming, will try to hide the ball. Once you have obtained as much as you can from the State, sit down and read it all.

Note the flaws in the literature that make it unreliable.

Note any inconsistencies in the experts' positions in articles compared to the facts in the medical records or other materials, such as EMT reports, that make their scientific positions irrelevant. Read every article with the doctor's name on it. This does not require you to be a scientist or a physician. It does require some quiet time and some common sense about what is reliable and what is not. Here are a few tips about how to read and critique the medical literature.

FIGHTING FIRE WITH FIRE: DIAGNOSING THE ILLS OF UNRELIABILITY IN SBS MEDICAL LITERATURE

Medical experts for the State frequently base their opinion that a child was shaken and/or violently slammed, and/or that the injuries were "inflicted" within a specific time frame upon some of the hundreds of medical articles that have been published about SBS or impact cases in the past half-century.

Getting behind these opinions in so-called SBS/impact cases is an exercise in unraveling the literature upon which the doctor bases his or her opinion. Recall that the Kumho Court observed that an expert's testimony that a proposition is accepted is not helpful to determine reliability "where the discipline itself lacks reliability." Where the doctor's clinical judgment regarding causation and timing of injury is informed by medical literature and medical school child abuse courses which, themselves, are based on the same faulty literature, the doctor's clinical judgment is subject to a reliability attack. Similarly, where clinical judgment is based solely on a physician's own experience, it is not reliable or relevant to draw conclusions about the general forensic issues of head injury causation and timing of injury.

INCONSISTENT CRITERIA FOR DIAGNOSIS OF SBS

The easiest way of attacking the opinion of the State's experts is to show that there is no general acceptance among physicians about what constitutes the diagnostic criteria for so-called shaken-baby-syndrome, or shaken-impact-baby-syndrome "S-IBS."

Some articles claim that SDH's, with retinal hemorrhages, without external evidence of injury, are diagnostic of the syndrome. Some authors claim that either a SDH, alone, or RH's, alone, are sufficient to make the diagnosis of SBS. Some authors claim that SDH's, RH'S and broken bones, are diagnostic of the "syndrome." Other authors claim that any "constellation" of the above symptoms is enough to diagnose SBS, or non-accidental impact.

The fact that there is no general acceptance among physicians as to what diagnostic criteria exist for the syndrome should be included in all challenges to the State's science of causation. Remember that "shaken baby syndrome" is both a statement of diagnosis *and causation*. If there is no consensus about the diagnostic criteria, how can there be consensus about the cause?

In reading the medical literature in this area, note the lack of consensus on issues relevant to the client's case. For example, in the area of timing of injury, neuroradiologists cannot reach consensus about how to date subdural hematomas on CT or MRI scans for purposes of establishing when injury happened. Similarly, neurosurgeons cannot agree on how to date subdural hematomas, nor can they agree on the definitions of "hyperacute," "acute," "subacute," "subchronic," and "chronic." Consistency and general acceptance do not exist.

BIAS: THE NEMESIS OF RELIABILITY

Some judges may need to become accustomed to an attack on SBS literature. "Why," they might ask, "am I seeing this at this point in time?" It may be wise to explain in your motion and brief that there has been a shift to evidence-based medicine in the past ten years in the medical profession, so much so that the validity and reliability of medical literature in all fields of medical practice is now being re-evaluated and questioned.

To put this in further context, you might explain that much of the medical literature across the board, not just in the child abuse field, that has been reevaluated suffers from the malady of bias. Indeed, bias of one sort or another has been found to be the nemesis of reliability in most medical studies. A bias in a study is an error that can lead to an exaggerated or false conclusion. Bias occurs due to faulty selection of patients or subjects, the way data are collected, or the way in which the authors reach their conclusions. In the world of scientific evidence, bias equals unreliability and exclusion.

Such problems have been amply documented. In a 1986 review of some 4,235 research reports on the efficacy of drug trials, surgical, psychotherapeutic and diagnostic procedures—the information upon which our doctors rely to diagnose and treat our ills—three researchers concluded that only about 20 percent were valid studies.

And, don't be fooled by the prestigious glitter of the journals. Among others, the 4,235 reports appeared in the *New England Journal of Medicine* (NEJM), the *Journal of the American Medical Association* (JAMA), the *British Medical Journal* (BMJ), *The Canadian Medical Association Journal*, *The Lancet*, *The American Journal of Psychiatry*, *Annals of Internal Medicine*, *Archives of Neurology and Psychiatry*, *The Journal of Nervous and Mental Disease*, and *Psychiatric Quarterly*.

In 1985, other researchers uncovered the fact that of more than 200 articles in two anesthesia journals, only 15 percent were without major errors in design and/or analysis. Some improvements have been made, with peer review journal boards using qualified statisticians and biostatisticians to review study design and check conclusions.[12] But change is slow to come, and the mountain of bad science is still formidable.

If the medical literature upon which our doctors rely to diagnose and treat us is flawed or selectively presented to us by journal editors, what of the literature upon which a doctor relies in determining that a child was shaken or a victim of intentional blunt trauma, or that the injury was "inflicted" within a certain time frame that implicates your client? Mark Donohue, M.D., an Australian-based physician, recently reviewed the SBS medical literature and concluded that "the commonly held opinion that the finding of SDH and RH in an infant was strong evidence of SBS was unsustainable…from the medical literature."[13] You will need to find specific examples relevant to your case to

prove this to the court by analyzing the medical literature you obtain in discovery or through your own research.

STUDY DESIGN: PITFALLS IN RELIABILITY

Once you have the articles, either from your own research and/or as provided in pretrial discovery, crucial for evaluating the reliability of the methods and conclusions is to first understand how medical studies are classified. They are either (1) observational or (2) experimental. Each type of study has its own set of potential pitfalls. In the area of SBS, fortunately, the job is relatively simple because there are a very limited number of experimental studies. Most are of the less reliable, observational variety.

Observational studies consist of descriptive or case-series studies (collections of case reports), case control studies, cross-sectional studies and cohort studies.

CASE-SERIES STUDIES: THE LOWEST FORM OF EVIDENTIARY LIFE

A case-series study is a simple, descriptive account of interesting characteristics observed in a group of patients. It is one in which the "researchers" look at a group of cases and ask, "What happened?" Because the study looks back in time, it is often called a "retrospective study." For example, John Caffey, M.D., collected cases of children who presented to the ER with SDH's and fractures of their arms and/or legs. Based on his observations—not experiments—Caffey wrote an article about these cases. He concluded they had been victims of shaking without impact.[14]

Most of the "observational" medical studies in the field of pediatric head injury and child abuse fall into this category. According to the AMA, a mere case-series study, which involves unsystematic clinical observations, is the least scientifically significant. (Please see, "Table 1 - A Hierarchy of Strength of Evidence for Treatment Decisions, page 44.")

Case-series studies do have a role in that they help medical researchers to form hypotheses, but that is as far as their usefulness goes. One physician recently commented:

> "A scientist or physician who has only case reports listed in their curriculum vitae is not given much credit as a researcher. All it means is that they were fortunate enough to get a group of reportable cases, had the ability to write in the currently fashionable style and the instincts to pick and choose the journal most likely to publish their work."[15]

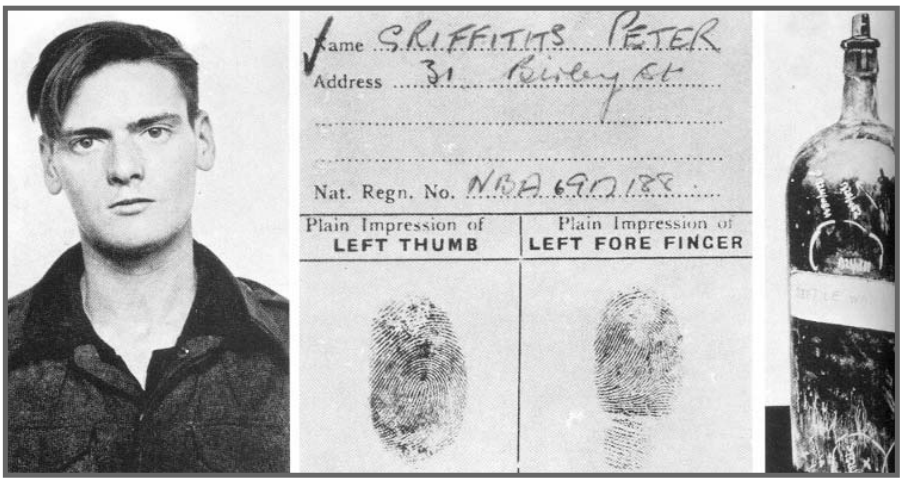
Case-series studies should be greeted with great skepticism if they claim to offer conclusions about pediatric head injury causation and timing of injury. Remember, case reports and studies based upon them only pose hypotheses. They do not contain science, and they are not a reliable basis for an opinion on causation or timing of injury. Each article should be separately critiqued in the pretrial challenge to the State's science, and each article critiqued should be attached as an exhibit to the brief so that the judge, who is the reliability gatekeeper, can do his or her job.

SELECTION BIAS

Selection bias occurs when a researcher only selects for the study those examples or subjects that will support their hypothesis. For example, a child with SDH's and RH's is assumed, without scientific proof, to have been a victim of shaking. (Other causes are not ruled out.) By contrast, random selection is far more reliable.

Australian physician Mark Donohue writes of the selection bias problem in the SBS area:

> "Studies and reports rely on either indirect or disputed evi-



Peter Griffiths, pictured above with his prints matching the prints on a bottle of distilled water found in the ward of a children's hospital in Blackburn, Lancashire in 1948. Griffiths kidnapped four-year-old June Devaney from her hospital bed, raped and murdered her. Griffiths was caught in the first mass registration of fingerprints in the world. More than 45,000 sets of prints were collected before Griffiths was caught, after which he confessed. Once thought to be a foolproof method of identification, now under Daubert, even the reliability of latent (partial) fingerprint identification has come under attack.

dence of the occurrence, severity, or type of trauma. Many studies lacking these critical data make the obvious logical error of selecting cases by the presence of the very clinical findings and test results they seek to validate as diagnostic [of the syndrome]. Not surprisingly, such studies tend to find their own case selection criteria pathognomonic [i.e., diagnostic] of SBS."

The problem of "indirect" or "disputed evidence" is one of shaky-data-syndrome. Case-series studies rely on faulty data, such as, the confessions of parents and babysitters to shaking and, in some cases, so-called "demeanor evidence." Confessions to shaking are not reliable because a person might confess to shaking a child in an attempt to revive the child, and not with the force required to cause injuries, such as, SDH's and RH's. A person might even confess to shaking because in a plea negotiation, arguably, it may be viewed as a less culpable and understandable stress response compared to slamming a child's head against a wall or other hard surface.[16]

Health care workers often include demeanor evidence in the medical record, such as, "Dad appeared to have no reaction and had flat affect." The theory that a flat affect is a sign of guilt is not only unreliable, it is ludicrous. All people respond differently to the shock of a loved one being suddenly rushed to the hospital. In some of the older SBS articles, demeanor evidence is included and some physicians still believe it is valid and may even "select" the "culprit" based upon this.

"Data-dredging" is another form of selection bias and occurs when a researcher takes a small part of a study and uses it to "prove" a point that the study was not designed to show. Data dredging is often used to form hypotheses but, unfortunately, these are then frequently presented as scientific fact. Watch for this common problem in SBS literature.

Another form of data dredging is where an author takes a series of case reports and tries to prove a hypothesis with these. It is particularly rampant in "review" articles that summarize the contents of other SBS case-series studies. A good example of this is a study published by the NEJM in 1998 in which four physicians tell us, for example, that short falls are unlikely to cause serious trauma, citing to nine case-series studies to prove that proposition.[17] We are also told that "risk factors" for nonaccidental injury include "young parents, unstable family situations [and] low socioeconomic status" and that "fathers and boyfriends" are the two groups most likely to abuse children. These propositions are supported by citations to three articles, each of them a case-series study, that is three articles that merely pose hypotheses, and give no reliable scientific data.

That most people lie when there are cases of blunt impact and serious injury is supported by cites to two case-series studies. Every single one of the conclusions in this NEJM article is a product of data dredging. There is not one original piece of research here in the text or, for the most part, cited in the footnotes. If any of the State's experts in a SBS case depend on this summary article, or others like it, for causation or timing of injury, this should be a ground for challenging their testimony, for both their methodology and conclusions are scientifically unreliable.

Table 1 - A Hierarchy of Strength of Evidence for Treatment Decisions

N of 1 randomized trial
Single Randomized trial
Systematic review of observational studies addressing patient-important outcomes
Single observational study addressing patient-important outcomes
Physiologic studies
Unsystematic clinical observations

This table shows the relative strengths (validity) of types of evidence presented in medical articles utlizing evidence-based medicine (EBM).

Looking at the hierarchy in the table, one can see that the strongest form of evidence is "N of 1 randomized trial," i.e., a double-blind study to test, for example, the treatment efficacy of a drug, in which one patient is tested alternately using a placebo and a real drug, but in which neither the patient nor the physician know which is which.

At the bottom of the rung us the lowest form of evidentiary life, i.e., "unsystematic clinical observations," of which case-series studies are the main example.

The reason case-series studies are on the bottom rung is that these are too unreliable for physicians to base their diagnoses upon and, therefore, their treatment decisions. Despite this, case-series studies are used as the main basis for physicians and, utlimately, prosecutors, to accuse others of deliberately causing injury to children by violenty shaking and/or slamming their heads on hard surfaces.

Table from: Article XXV, "A Users' Guides to Medical Literature," by American Medical Association, reprinted in JAMA, September 13, 2000, Vol 284, No. 10

One last point of definition: A study riddled with the flaws of selection bias is said to suffer from confirmation bias. The easiest way to distinguish selection bias from confirmation bias is to appreciate that *confirmation bias* is the *end result* of the *process of selection bias*. In addressing such issues in your brief, you might want to point out that the Daubert Court relied on Karl Popper for the modern day approach to scientific inquiry: It should seek to falsify, and not confirm a hypothesis.

INSUFFICIENT DATA

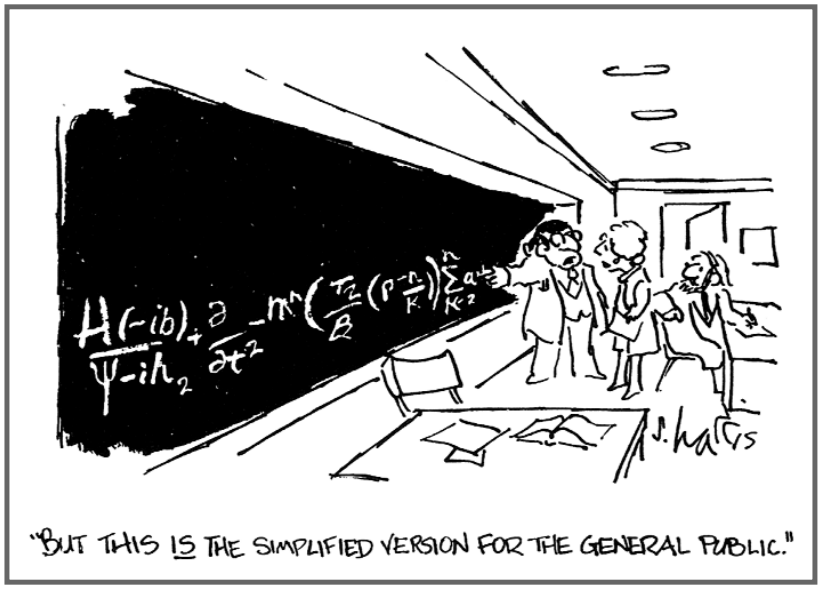
Case-series studies rarely provide enough clinical data to allow the reader to get behind the scientific conclusion and make his or her own determination about reliability. Therefore, "peer review" by fellow physicians and by pediatric head injury biomechanicians, is precluded. This flaw alone makes any study vulnerable to attack. The question is: "Is there enough information here that a physician would be able to agree with the factual foundation for the diagnosis of intentionally-inflicted injury?" Work with

your own expert or experts in determining the answer to that question and, if there is insufficient data, obtain an affidavit in support of your reliability challenge.

Forensic Neuropathologist Jan E. Leestma, M.D., reviewed 57 articles on SBS published between 1969 and 2000. He found that of 324 cases in the 57 articles, only 54 had enough data for a reader to make his or her own judgment about the scientific reliability of the medical conclusions regarding causation. If a study does not include enough data to allow other physicians to analyze the validity of the methodology and conclusions, then it has not been subject to genuine peer review, even though it has been published.

STATISTICS: A TOOL FOR THOSE WITH NO PROOF?

With increasing frequency, physicians who write up case reports as case-series studies are attempting to breathe scientific life into these lifeless corpses by applying statistical analysis. However, the sample sizes in case-series studies are often too small to have any statistical power, i.e., "statistical significance." Statistical significance means that a correlation between two or more factors is, allegedly, established to a high degree of certainty.[18] Statistical significance in a study is related to odds ratios. But these vary wildly depending on the size of the sample.



Science can bedevil us with its complicated terminology. But, to tell the story of scientific truth in your case, it is crucial to become conversant with the relevant areas of science so that you can empower the judge to be the reliability gatekeeper in your case. Copyright, 2004, Sidney Harris. Reprinted with permission.

Most SBS articles are based on case-series studies of small groups. Watch what happens to the odds ratios when the numbers are small, and then when they are large, in the following hypotheticals about whether estrogen causes breast cancer:

Hypothetical Study 1: Twenty women are divided into two groups of ten. The first group of ten women who are post-menopausal take estrogen and, over their life times, two of them develop breast cancer. The women in the second group of women (the control group) do not take estrogen at all, and over their life times, only one develops breast cancer. Here the odds ratio (OR) is 2 to 1 or 2.0 that estrogen use is related to the development of breast cancer.

Hypothetical Study 2: Twenty women are divided into two groups of ten. The first ten take estrogen after menopause, and only one in this group develops breast cancer. In the second group of ten, none of whom take estrogen, only one develops breast cancer. In Study 2, the odds ratio (OR) that estrogen is implicated in breast cancer is 1.0, that is, it is no more likely in the first group of ten women than in the second.

Hypothetical Study 3: Twenty women are divided into two groups of ten. In the estrogen group, one develops breast cancer, and in the control group, two develop breast cancer. By dividing two into one, the odds ratio is 0.5. In this study, it appears that women who take estrogen are only half as likely to develop breast cancer.

Hypothetical Study 4: Two hundred women are divided into two groups of 100 each. One hundred take estrogen and 100 do not. In the estrogen group, 14 women develop breast cancer, and in the control group, only 10 develop it. According to this study, the odds ratio is 1.4 that a woman who takes estrogen will develop breast cancer. (Compare to Study 2 where the odds ratio of developing breast cancer was 1.0. In Study 4 there is a $^{4}/_{10}^{ths}$ of a chance greater of developing breast cancer, i.e., 1.4.)

Hypothetical Study 5: Two hundred thousand women are each divided into groups of 100,000. In the estrogen group, 11,000 women develop breast cancer, while in the control group, only 10,000 women develop breast cancer. According to this study, the odds ratio of developing breast cancer if one takes estrogen is 1.1.

The significance of the odds ratios in Hypothetical Studies 1-5 is that when study samples or groups are small, the odds ratios jump all over the place. In small groups, the results are much more likely to be random. But, according to statisticians, the larger the study group, the more likely it is that the odds ratio is real. Epidemiological data that form the basis for statistics in

medical articles depend on sample sizes that are large enough to be statistically significant.[19]

Any reliance by a physician in an SBS case on a case-series study in which the 'researchers' applied statistics, should be attacked as unreliable.

DATA-POOLING TO CONJURE UP THE 'STATISTICS BOGEY MAN'

Many physicians who are proponents of shaking as a cause of SDH's and RH's have recognized the dilemma of too few subjects in too many studies and, as a way of covering their work with a scientific gloss, have "pooled" what they claim is "data" from several different studies. This is called "meta-analysis" or "re-analysis." While that might work in some areas of medical research in large, epidemiological studies, it does not work in SBS case-series studies. The simple reason for this is that, unlike some areas of epidemiological research, the study designs in the SBS literature are an inconsistent alphabet soup of mostly case-series studies of different specialties, such as radiology, neurosurgery and pediatrics, all of which assign different clinical meanings to the same medical terms. These studies have different group sizes, different selection criteria, different analytical methods, and different conclusions and, of course, no controls.[20]

The strength of a particular meta-analysis depends upon the validity of each of the studies included in the meta-analysis. The faulty methodology of combining cases from several case-series studies into a larger case-review series still cannot lead to a reliable conclusion because case-series studies, whether large are small, are designed to form hypotheses, and not conclusions. In other words, more voodoo.

If, when reading an article about so-called SBS, or any related area like retinal hemorrhages, you see words like "statistics," "statistical significance," "odds ratio," or "confidence interval," get ready to have some fun. It is pure junk science, through and through, for statistics, at least in the context of case-series studies, truly is for those who have no real proof. (Again, think about how ludicrous it is to apply statistics to a bunch of unproved hypotheses, i.e., case-series studies.)

Statistics are also not a reliable way to determine causation in reported cases of so-called SBS. Statistics do not, in fact, prove causation. All statistical conclusions do is to report correlations. "Statistics," said Aaron Levinson, are like bikinis. What they reveal is suggestive, but what they conceal is vital." While "correlations" may be reliable enough for a toxic tort lawyer in cases where social policy creates "proximate causation" to boost the causation-in-fact proofs, statistical correlations should not help the State in criminal cases of alleged SBS or in family court proceedings to terminate custody or parental rights. While it is the case that statistics are not admissible in criminal cases to prove causation, the fact is that the State's physicians often rely on faulty SBS statistical conclusions from SBS case-series studies in forming their opinions in a case about causation and timing. As lawyers, we need to get behind these opinions to expose their reliance on faulty, statistical foundations.

CASE-CONTROL STUDIES

As with case-series studies, case-control studies are retrospective. They involve two groups: one group with the condition and one without, that is, the "control group." This type of study looks back in time and attempts to determine what risk factors, if any, existed that caused the condition in one group, but not in the control group. For example, in the area of SBS, consider the following study:

Group 1 consists of 10 children all of whom are said to be victims of SBS because they have SDH's and RH's. Group 2, the control, are not victims of SBS. The study concludes that the children who were victims of SBS came predominantly from single-parent homes where the parent was on welfare. However, the study did not consider the other adults in the home with access to the children. The conclusion of the study is that the "risk factors" for children being shaken are (1) a single parent who (2) is on public assistance.

In this example, the conclusion that single parents on welfare pose a risk to children is flawed for several reasons including the failure to consider that other adults in the home may have been responsible for the alleged abuse. Studies that claim to be "case-control" studies should be analyzed for these types of biases.

CROSS-SECTIONAL SURVEY STUDIES

Cross-sectional studies analyze data collected on a group of subjects at one time, rather than over a longer period. They are snapshots in time as to, "What is happening now?" For example, a cross-sectional study might be of all defendants who confess to shaking a child in a particular year. Assuming they are convicted for this, they complete a survey as part of a pre-sentencing investigation. This type of study does not contribute to defining SBS. The confessions of pure shaking without impact (which may or may not be truthful in cases of plea bargaining) bear little or no relationship to the scientific evidence on the subject presented in those biomechanical studies that are scientifically reliable.

Cross-sectional studies frequently focus on issues of socio-economic status, gender and (in some of the older ones) even race or ethnic heritage. If the client is poor, African American, or Hispanic, some physicians may have relied on stereotypes in such studies in 'selecting' the suspect. Other studies, however, "prove" that child abuse is not peculiar to poor people or people of any particular race or ethnicity. Any conclusion based on cross-sectional studies like these is unreliable.

COHORT STUDIES

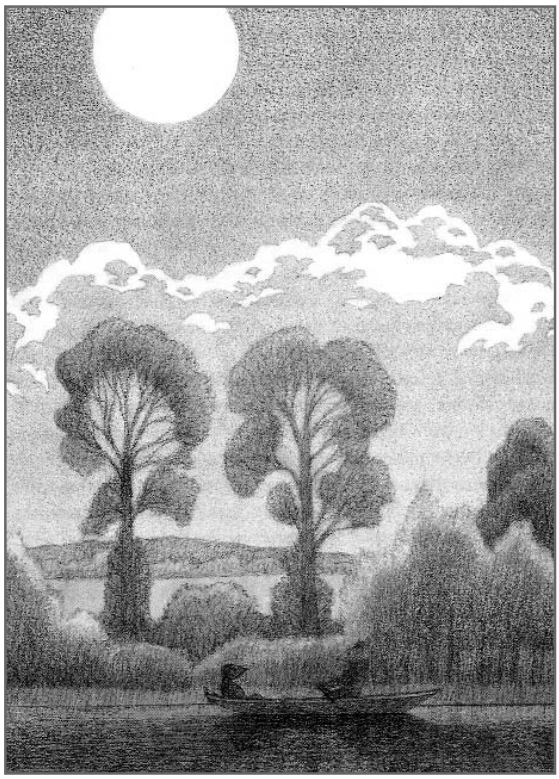
While there are no cohort studies in the SBS timing and causation literature, it may be helpful to know what one is so that, in the future, if you see one, you'll be able to classify it. A cohort is a group of subjects that have something in common and who remain part of a group over an extended period of time. In medicine, the subjects in cohort studies are selected by some defining characteristic, such as one that is suspected of being a precursor to a disease.

A cohort might be a group of children whose genetic history

puts them at risk for developing SDH's and fractured bones as a result of Glutaric Acidemia Types I or II. Cohort studies ask the question: "What will happen?" And, because they look forward, they are called "prospective" studies. They are still only observational studies and they are still subject to all the same biases as with others of their genre. Beware of SBS studies that claim to be prospective. This is often a description that tortures the underlying definition in order to add a gloss of (pseudo) reliability to the conclusions.

PUBLICATION BIAS

State experts sometimes claim that there are relatively few studies in the medical literature that challenge SBS as a cause of SDH's and RH's. This is frequently used as a way of diminishing and characterizing the defense experts as those who are mere "mavericks" or who are on the fringe of the medical establishment.[21]



Ratty and Mole venture into the night to find the lost child treasured by all as a gift. Do not be afraid to venture into the night to recover and seize the scientific truth for your client.

This is the time to raise the issue of publication bias. Many articles that present unreliable science still slip through the cracks of the boards of peer review journals and those that present science contrary to popular belief are more likely to be rejected. Take Dr. Atkins and his low-carb diet, for example. His efforts to get his message out were hindered by the American Medical Association (AMA) and the American Heart Association (AHA) that saw his diet as dangerous to their low-fat dogma.

One medical examiner recently lamented:

> Even though human biases are inevitable, they become dogma because the medical elite accepts, appropriates and then perpetuates them. Research to the opposite is not funded, cannot be published and, if it does see the light of print, ridiculed and pooh poohed 'til at times it causes the loss of the researcher's career. Journals are not pure and scientific, but only reflect the biases of the editorial boards, which means the power elite in the field."[22]

It is still extremely difficult to get any of the major journals to publish articles that refute the common and erroneous beliefs in so-called shaken baby syndrome.

In 2002, Ommaya and colleagues attempted to publish a major review article that included some of the medical and biomechanics data that challenged shaking as the cause of SDH's over the convexities of the brain.[23] A major U.S.-based neurosurgery journal rejected the work. In the United States, the issue of child abuse is highly politicized, with supporters often receiving government funding to "recognize" cases of SBS as part of a prevention program. In the UK, however, this is less the case. Following some shuttle diplomacy by an American neurosurgeon, the article was published in 2002 in the British Journal of Neurosurgery. And, in the less politically charged atmosphere of the U.K., authorities there recently announced that all cases of so-called SBS are now to be reviewed to determine if some people were wrongly convicted.[24]

EXPERIMENTAL STUDIES

These are studies in which researchers, having formed hypotheses from observational studies, set out to test them. One fact that speaks volumes about this area is that there are a limited number of these in the SBS literature. This state of affairs should be brought to the trial judge's attention to underscore the fact that case-series studies and other observational studies merely present unproved and untested hypotheses about causation and timing of injury. As discussed in Parts I and II of this series, two of the best known experimental studies show that by shaking, alone, a human being is not capable of generating the forces necessary to produce SDH's, diffuse axonal injury (DAI) and concussion.

THE PIPER AT THE GATES OF DAWN

In Graham Greene's "The Wind in the Willows," two of the story's charming little river creatures, Ratty and Mole, are just about to turn in for the night after a long, summer's day. But Ratty cannot sleep, plagued as he is with worry about the son of his friend, Otter. The child, Portly, has been missing for too long to be just off playing somewhere in a self-possessed and absent-minded way. Otter is afraid, and now Ratty is, too, that the baby Otter—who hasn't yet learned to swim—has been sucked up in the fast flow of a local ford, a place for which the child has an insatiable fascination. Instead of going to sleep, Ratty and Mole set off into the night in their rowboat along the moonlit river in search of the missing child.

Their journey is long and treacherous through the shadows and unfamiliar shapes of the now-darkened land. Then, as the pinks

and oranges of the sky declare themselves on the dawn's horizon, Ratty and Mole hear the magical and enchanting sound of pipe music. Drawn hypnotically to this, they row their boat closer and, as they do so, are drawn to a small island midstream from which a bright light is getting blindingly brighter by the second. They pull to shore and, blinking in the blaze of light, step onto the island where they are greeted by a heavenly being with pipes in hand. Between the toes of this great shining, holy being, the baby otter, Portly, slumbers peacefully in safe oblivion. Then the Piper at the gates of dawn smiles kindly at the brave little river creatures that came out in the night to find the lost baby otter. And with that smile, Portly, the precious otter child, is returned to those that love and treasure him as a gift.

As with a gift that has been returned to us, the power to decide what is reliable is now returned to the hands of trial judges. But that means we lawyers must work hard to discover the nuggets of scientific truth, and to expose unreliable science in our opponent's case, to become so conversant with the scientific issues that we are able to help guide the judge, by briefing, by proficient direct and cross examination of experts in pretrial Daubert hearings or through affidavits, to reach a just result based on reliable and relevant science. This article suggested a few guidelines to get started on that journey in SBS cases. You will no doubt discover many areas of faulty methodology and conclusions in your own cases. Do what you can to keep it simple, focusing on reliability and relevancy to get at the scientific truth in each case. The piper at the gates of dawn has returned to us the lost power to do justice. It is time now to use it for our clients.

ENDNOTES

1 Part I of this series, "The Elephant on the Moon," *The WARRIOR*, Fall 2003, and Part II, "A Matter of Gravity," may be obtained directly from the journal. Go to www.triallawyerscollege.com for the current address. Both Parts I and II may also be obtained as PDF's (and viewed using Adobe Acrobat Reader) from the author's web site at www.sharplaw.biz. All rights reserved.

2 This is part of a book currently being written by Elaine Whitfield Sharp. Copyright by Elaine Whitfield Sharp, January, 2004. Permission to publish given to *The WARRIOR*.

3 Daubert v Merrell Dow Pharmaceuticals, 509 US 579, 113 S.Ct. 2786 (1993). Daubert's progeny includes: General Electric Company v Joiner, 522 U.S. 136, 118 S. Ct. 512 (1997) ("Joiner"); and Kumho Tire Co. v Carmichael, 526 U.S. 137, 119 S.Ct 1167 (1999) ("Kumho"). Further citation to these cases is to the official reporters.

4 Frye v United States, 54 App. D.C. 46, 293 F. 1013 (D.C. Cir. 1923), ("Frye").

5 Mr. Frye was ultimately exonerated and released when exculpatory evidence came to light. However, the systolic blood pressure test, which spawned the polygraph, did not fare so well. In 2002, the National Academy of Sciences issued a report concluding that the lie-detector test was not reliable science.

6 See, e.g., United States v Llera Plaza, 2000 W.L 27305 (E.D. Pa., Jan, 2002), in which the court notes that the Frye test was applied to all science in the federal courts.

7 See, e.g., State v Porter, 241 Conn. 57 (1997), in which the Connecticut Supreme Court discusses the history of the Connecticut judiciary's independence from the scientific world in evaluating the reliability of evidence proffered as "scientific."

8 *Amici* briefs from some of the heavy hitters in the scientific community included: American Association for the Advancement of Science (AAAS) with the National Academy of Sciences, The New England Journal of Medicine (NEJM), Annals of Internal Medicine, the Carnegie Commission for Science, Technology and Government, and Physicians, Scientists and Historians of Science.

9 A discussion of the development of scientific thought up to and beyond Popper is included in Part I of this series, "The Elephant on the Moon," *The WARRIOR*, Fall 2003. Please see endnote 1, above, for information about how to obtain this.

10 See, endnote 6, above, in which one court discusses this development.

11 See, Part I, "The Elephant in the Moon," and Part II, "A Matter of Gravity," and endnote 1, above.

12 See, generally, Dawson-Saunders, B., and Trapp, R.G., "Basic and Clinical Biostatistics," Appleton & Lange, Chapter 1, (hereafter "Biostatistics").

13 Donohue, M., M.D., *Am J For Med & Path*, Vol 24, No. 3, pp 239-242, Sept. 2003, at 241.

14 See, Part I, "The Elephant in the Moon," and endnote 1, above.

15 Confidential communication.

16 This is discussed in Part I: "The Elephant in the Moon." See endnote 1, above.

17 Duhaime, et al, "Nonaccidental Head Injury in Infants—The 'Shaken Baby Syndrome,'" NEJM, Vol. 338, No. 25, pp 1822-29, June 18, 1998.

18 This degree of certainty is called a "confidence interval" or CI and, as it involves some complicated calculations that are not helpful here, this will not be discussed further. For further reading in this, see, Milloy, S.J., *Junk Science Judo*, Cato Institute, and *Biostatistics*, cited *supra*.

19 At least one author claims that big numbers result in big lies. See, Milloy, *supra*.

20 A "control" is a group that does not share the condition being studied.

21 Let us not forget that, in their own time, Hipprocrates, Galileo and Sir Isaac Newton were all mavericks.

22 Confidential communication.

23 Ommaya, A.K, M.D., et al, "Biomechanics and neuropathology of adults and paediatric heady injury," BJN, 2002; 16(3): 220-224.

24 The Times, (London) January 26, 2004, p 4, "Shaken-baby death cases to be reviewed."